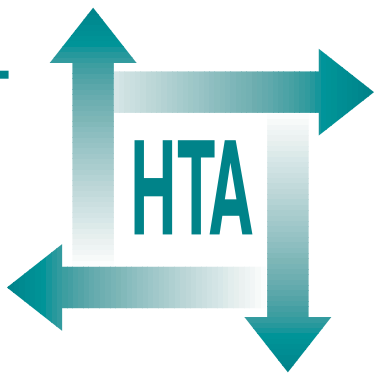
Health Technology Assessment 2001; Vol. 5: No. 4

Methodology

Quality-of-life measures in chronic diseases of childhood

C Eiser
R Morse

Health Technology Assessment
NHS R&D HTA Programme



How to obtain copies of this and other HTA Programme reports.
An electronic version of this publication, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (http://www.hta.ac.uk). A fully searchable CD-ROM is also available (see below).

Printed copies of HTA monographs cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our Despatch Agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per monograph and for the rest of the world £3 per monograph.

You can order HTA monographs from our Despatch Agents:

– fax (with **credit card** or **official purchase order**)
– post (with **credit card** or **official purchase order** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you **either** to pay securely by credit card **or** to print out your order and then post or fax it.

Contact details are as follows:

HTA Despatch
c/o Direct Mail Works Ltd
4 Oakwood Business Centre
Downley, HAVANT PO9 2NP, UK

Email: orders@hta.ac.uk
Tel: 02392 492 000
Fax: 02392 478 555
Fax from outside the UK: +44 2392 478 555

NHS libraries can subscribe free of charge. Public libraries can subscribe at a very reduced cost of £100 for each volume (normally comprising 30–40 titles). The commercial subscription rate is £300 per volume. Please see our website for details. Subscriptions can only be purchased for the current or forthcoming volume.

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *Direct Mail Works Ltd* and drawn on a bank with a UK address.

Paying by credit card
The following cards are accepted by phone, fax, post or via the website ordering pages: Delta, Eurocard, Mastercard, Solo, Switch and Visa. We advise against sending credit card details in a plain email.

Paying by official purchase order
You can post or fax these, but they must be from public bodies (i.e. NHS or universities) within the UK. We cannot at present accept purchase orders from commercial companies or from outside the UK.

How do I get a copy of *HTA on CD*?

Please use the form on the HTA website (www.hta.ac.uk/htacd.htm). Or contact Direct Mail Works (see contact details above) by email, post, fax or phone. *HTA on CD* is currently free of charge worldwide.

The website also provides information about the HTA Programme and lists the membership of the various committees.

Quality-of-life measures in chronic diseases of childhood

C Eiser*
R Morse

Department of Psychology, University of Sheffield, Sheffield, UK

* Corresponding author

Competing interests: none declared

Published February 2001

This report should be referenced as follows:

Eiser C, Morse R. Quality-of-life measures in chronic diseases of childhood. *Health Technol Assess* 2001;**5**(4).

Health Technology Assessment is indexed in *Index Medicus*/MEDLINE and *Excerpta Medica*/EMBASE. Copies of the Executive Summaries are available from the NCCHTA website (see opposite).

NHS R&D HTA Programme

The NHS R&D Health Technology Assessment (HTA) Programme was set up in 1993 to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and provide care in the NHS.

Initially, six HTA panels (pharmaceuticals, acute sector, primary and community care, diagnostics and imaging, population screening, methodology) helped to set the research priorities for the HTA Programme. However, during the past few years there have been a number of changes in and around NHS R&D, such as the establishment of the National Institute for Clinical Excellence (NICE) and the creation of three new research programmes: Service Delivery and Organisation (SDO); New and Emerging Applications of Technology (NEAT); and the Methodology Programme.

Although the National Coordinating Centre for Health Technology Assessment (NCCHTA) commissions research on behalf of the Methodology Programme, it is the Methodology Group that now considers and advises the Methodology Programme Director on the best research projects to pursue.

The research reported in this monograph was funded as project number 96/48/02.

The views expressed in this publication are those of the authors and not necessarily those of the Methodology Programme, HTA Programme or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for any recommendations made by the authors.

Criteria for inclusion in the HTA monograph series
Reports are published in the HTA monograph series if (1) they have resulted from work commissioned for the HTA Programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

The editors and publisher have tried to ensure the accuracy of this report but do not accept liability for damages or losses arising from material published in this report. They would like to thank the referees for their constructive comments on the draft document.

ISSN 1366-5278

Published by Core Research, Alton, on behalf of the NCCHTA.
Printed on acid-free paper in the UK by The Basingstoke Press, Basingstoke.

M

Contents

List of abbreviations

ACTG	AIDS Controlled Trial Group*
A-FILE	Family Inventory of Life Events and Changes for Adolescents*
AMA	About My Asthma
BASES	Behavioral, Affective, and Somatic Experiences Scale
CAQ	Childhood Asthma Questionnaire*
CBCL	Child Behaviour Checklist
CCTR	Cochrane Controlled Trials Register
CDI	Child Depression Inventory
CDLQI	Children's Dermatology Life Quality Index
CES-D	Centre for Epidemiological Studies Scale
CHQ	Child Health Questionnaire
CHIP-AE	Child Health and Illness Profile – Adolescent Edition*
COOP	Co-operative Information Project
CF	cystic fibrosis
CQOL	Child Quality of Life Questionnaire*
DQOL	Diabetes Quality of Life
EHRQL	Exeter Health Related Quality of Life Measure*
FILE	Family Inventory of Life Events and Changes for Parents
FSIIR	Functional Status (II) Revised*
GHQ	General Health Questionnaire
HAY	How Are You?*
HRQoL	Health-Related Quality of Life
HTA	Health Technology Assessment
HUI	Health Utility Index
ICC	intraclass correlation
ICD-10	International Classification of Disease-10
IQ	intelligence quotient
LSIA	Life Satisfaction Index for Adolescents*
OM-6	Quality of Life for Children with Otitis Media
PAQLQ	Pediatric Asthma Quality of Life Questionnaire
n.s.	not significant*
PCQL	Pediatric Cancer Quality of Life*
PedsQL	Paediatric Quality of Life Questionnaire
PGI	Patient Generated Index
PIE	Perceived Illness Experience*
PRQLQ	Paediatric Rhino-conjunctivitis Quality of Life Questionnaire*
QALY	quality-adjusted life-year
QLH-Y	Quality of Life Headache in Youth*
QLSD	Quality of Life Profile for Spine Deformities*
QoL	quality of life
QOLIE	Quality of Life In Epilepsy
QWB	Quality of Well-Being Scale
SF-36	Short Form 36
SIP	Sickness Impact Profile
TACQOL	TNO-AZL (Netherlands Organisation for Applied Scientific Research) Children's Quality of Life*
VAS	visual analogue scale*
WISC-R	Wechsler Intelligence Scale for Children-Revised*

* Used only in tables

Executive summary

Scope of the report

This report is concerned with the evaluation of measures broadly designed to measure quality of life (QoL) in children and adolescents, either by self-report or proxy raters. Four research questions were identified.

- To what extent are adult measures used in the evaluation of healthcare interventions in children?
- How appropriate are adult measures for use with children?
- To what extent do child self-reports correspond with assessments made by parents and carers?
- How feasible and reliable are proxy measures of QoL in different disease contexts?

Objectives

- To review the state of the art with regard to measurement of QoL for children.
- To make recommendations regarding the value of currently available measures for different purposes.
- To identify further research needs.

Method

Electronic databases were searched for the period 1980–July 1999 for articles relating to measures of QoL, health status or well-being in children (under 18 years) with chronic disease. Handsearching of relevant journals and cross-referencing with reference lists in identified articles was also carried out. Key workers in the field were contacted for additional information, and the Internet was searched for relevant websites.

Results

Forty-three measures were identified (19 generic and 24 disease-specific). Sixteen measures allowed for completion by children and parent/caregiver; seven only allowed for completion by a proxy, and the remainder ($n = 17$) allowed only for child completion.

The measures were described as QoL ($n = 30$), health status, ($n = 2$), perception of illness ($n = 1$), life satisfaction ($n = 1$) and quality of well-being ($n = 1$).

To what extent are adult measures used in the evaluation of healthcare interventions in children?

Three studies were identified where adult measures were used with very few changes made for children. In 11 studies involving nine separate measures of QoL, adult measures were used as a model for work with children.

How appropriate are adult measures for use with children?

Adult measures may fail to tap the specific aspects of QoL that are important to the child. Measures based on adult work impose considerable response burden for children, in terms of length, reading skills and response scale. Wording and format of adult measures may need to be modified to account for children's cognitive and language skills. More basic research is needed to determine the level of response burden that children of different ages can manage. Assessments of difficulty (e.g. reading age) need to be routinely included with information about new measures.

To what extent do child self-reports correspond with assessments made by parents and carers?

Fourteen studies were identified in which concordance between child and parent was investigated, often as part of the development of a new measure. There was some evidence for greater concordance between child and parent for physical functioning compared with social and emotional domains, but greater heterogeneity in the latter measures may contribute to inconsistent results. There was no simple relationship between concordance and moderating variables such as age, gender and illness, but this conclusion was addressed only very rarely.

How feasible and reliable are proxy measures of QoL in different disease contexts?

Only five papers fulfilled the review criteria. Evaluation is difficult because authors fail to justify their choice of measures, and do not report critical information such as completion rates or missing data.

Use of existing measures can potentially eliminate the time and expense required to develop a comprehensive measure of QoL, but a full battery of standardised tests may be expensive in terms of time for administration and scoring. In addition, battery measures tend to be lengthy and therefore demanding for sick patients. They are not recommended for work with children.

Recommendations for research

Minimum criteria for new measures

A set of procedures needs to be established for the development of new measures. These need to draw on the experience gained in development of child and adult measures to date. Basic research to enhance understanding of how children interpret questions in QoL measures is recommended. We need to understand the differences in meaning of items between children and adults, and between children of different ages. Some attempt to develop measures for children of 6 years or more have been reported, and these should be further developed.

Development of new measures should:

- follow established procedures for the development of measures
- take into account theoretical knowledge of children's understanding of illness, emotion, and ability to complete rating scales
- include facility for child and proxy report
- include developmentally sensitive age-appropriate sections
- include generic core and disease-specific modules.

Clinical appropriateness

There is a need to develop measures that are appropriate for the kind of questions to be answered in practice. Measures are frequently justified in terms of the value in clinical trials, comparing alternative treatments or assessing interventions. In more everyday contexts, QoL measures may potentially help health professionals and children's families evaluate clinical care. Outcome measures that are sensitive to changes in the child's QoL have considerable value, particularly in children with long-term illness.

To determine how far assessments of QoL can contribute to improved care, we need to move beyond the development of new measures. In order to encourage greater use in clinical practice, it is recommended that:

- developers of new measures need to be clearer about the procedures adopted for identifying the item pool, and more extensive information about their psychometric properties should be provided
- those developing new measures should work more closely with clinicians in order to ensure both the quality of the measures and their appropriateness in different clinical settings
- families should be encouraged to be more involved in the development and application of measures, in order to improve the face validity of measures, and to challenge criticisms that QoL measures impose an unnecessary burden.

Chapter 1

The history and scope of quality-of-life measurement for children

Aims and scope

The evaluation of adult quality of life (QoL) is well established, and measures of QoL are routinely included in many clinical trials. QoL has been a category in Index Medicus since 1966,[1] but interest in children's QoL did not gain momentum until the 1980s.

It is possible to identify a range of situations in which QoL decisions are important. One of the most critical catalysts in promoting measurement of QoL in children is the changing epidemiology of childhood disease. Where it is possible to manage but not cure a disease, we must determine how far treatment and disease compromise the child's QoL. In this way, informed judgements can be made about whether or not treatment is appropriate, and critically, where there is a choice, which might be the best option for the child. These considerations might apply when thinking about self-limiting illness (which still accounts for most childhood morbidity), chronic illness, or palliative care. Decisions of this kind make it essential that QoL is an integral component in clinical trials, or in any assessment of the outcomes of new treatments. At one extreme, QoL measurement may be useful in routine audit work or it may simply be useful to understand the child's perspective. At the other extreme, QoL may assist decisions about the rationing of resources. There is also an assumption that QoL needs to be considered when making 'end of life' decisions – when it is appropriate to withhold treatment because the anticipated QoL is so poor. QoL then is central to paediatric practice.

Our purpose in this chapter is to place the assessment of child QoL in historical perspective and identify the potential ways in which QoL assessment may be useful. This lays the foundation for the evaluation of QoL. In subsequent chapters, we address the four questions defined by the Health Technology Assessment (HTA) Programme.

- To what extent are adult measures used in the evaluation of healthcare interventions in children?
- How appropriate are adult measures for use with children?
- To what extent do child self-reports correspond with the assessments made by parents and carers?
- How feasible and reliable are proxy measures of QoL in different disease contexts?

Conceptual approaches to understanding QoL

The scientific study of QoL is complicated by the diverse ways in which the term is used in everyday language. To many people, QoL is synonymous with happiness. To others it is about material wealth. For others, QoL is about relationships with family and friends. Notions of citizenship, involvement in the community and contributions to charity work, have also been implicated.[2] Government Departments now acknowledge the need to include a wide range of concepts in any evaluation of the QoL of the nation. In a report in *The Guardian*,[3] John Prescott suggested that economic output alone was inadequate and more appropriate assessments of QoL need to incorporate a wider range of QoL measures including education, pollution and the number of birds in the countryside.

QoL, however, is not purely the preserve of politicians or environmentalists. Social historians emphasise the way in which perceptions of QoL differ over time and between cultures. In terms of material wealth, QoL is, of course, significantly better today compared with the turn of the last century. However, in terms of more socially orientated ideals, including perceptions of family life, or incidence of drug taking in young people, it may seem that QoL now is significantly poorer than in previous times. Thus, definitions and ideals are dependent on the specific social, cultural, spiritual and historical circumstances in which we find ourselves. In Western Europe, we take for granted running water, good housing and adequate food. These are a basic part of our QoL. With these fulfilled, we can think about our QoL in terms of social and psychological factors. In other

countries, however, QoL may be very dependent on the availability of the basic supplies.

In practice, a number of formal approaches to measuring QoL can be identified.[4]

The philosophical approach asks: 'What is QoL?' This question has attracted interest from ancient times. Does a good QoL involve some degree of conflict and challenge as Aristotle argued? His views might be seen to be precursors of ideas that QoL is about rising to challenges and coping with adversity.

The economic approach reflects the traditional concerns of Western European governments with the emphasis on economic growth measured in terms of gross national product. These measures are thought to reflect QoL in terms of acquired wealth. To some extent, this is an acceptable approach. QoL is considerably improved when a family gains a fridge, for example; less time is needed for shopping and there is less risk of food poisoning from contaminated food. Other perfectly defensible indicators that are also more health orientated include infant mortality, life expectancy and literacy rates. Social and health-related variables such as infant mortality, life expectancy and general literacy have been included in the Physical QoL Index.[5] To make the index more appropriate for children, two other variables have been included: the number of children in the labour force and female literacy rates.[6] Similar variables are now included in reports by UNICEF[7] for example, and allow for comparison between nations. Perhaps though we have become disillusioned with the concept of material wealth. Higher incomes are not necessarily synonymous with a high QoL. Material wealth without happiness is not QoL.

The sociological approach emphasises the social and environmental aspects of QoL. According to this tradition, QoL is less about material wealth and more about the individual's assessment of their circumstances. QoL is bound up with our relationships with others and our views about how far we are loved by others. The sociological approach emphasises the subjective nature of individual experience, and the interdependence between QoL and cultural experience. Underlying this approach are questions about who will judge QoL. What is the standard against which QoL judgements are made?

The psychological approach emphasises the role of individual appraisal in our conceptualisation of QoL. According to this approach, an individual with a good QoL has high self-esteem, is able to make decisions, is active, happy and fulfilled. Such an ideal is perhaps achieved by very few, but leads to the assumption that the nearer we feel we are to attaining our life goals the better our QoL.

The medical approach emerged in response to advances in medical care. As soon as we appreciate that cure may not be possible, it becomes important to establish that treatment really makes patients feel better. Thus, quantity of survival is no longer perceived to be the only end-point. How to measure this 'feel better' factor, particularly with children, remains a challenge. The concept of **health-related QoL** draws on ideas from all of the traditions described above and refers specifically to the impact of health and illness on the individual's QoL. It is often distinguished from the more general and popular meanings of the term.

In the following review, our discussions centre on questions of health-related QoL, but for simplicity the abbreviation QoL is adopted.

A brief history of child QoL

Early efforts to describe QoL in children were invariably focused on functional problems, and usually relied on assessments made by clinicians. Some of the earliest attempts to assess QoL were provided by Ditesheim and Templeton,[8] Herndon and co-workers,[9] and Henning and co-workers.[10] Ditesheim and Templeton[8] assessed QoL in infants following surgical repair of high imperforate anus. Their assessment of QoL was based on questionnaire information concerning school attendance, social relationships, and physical abilities. Herndon and co-workers[9] reported a follow-up of 12 survivors treated for major burns. They included assessments of physical functioning, degree of scarring and psychological adjustment. One-third of the children were very fearful, showed regressive behaviour or neurotic symptoms. However, many children showed "remarkable energy in adapting to their disabilities". Henning and co-workers[10] measured degree of function, height and attitudes in children with end-stage renal disease. Although average height was 'normal' most children were disappointed with their height.

The findings of these early studies paved the way for modern QoL work in two ways. First, they emphasised that children can adapt to their situation following major stress or injury, and

second that children's views about their disability (or height) are important. Thus, these early authors might be credited with anticipating the direction in which QoL measurement was to go (even though their measurement of QoL was very simple.)

Much of the impetus for more formal measurement of child QoL has come from work in paediatric oncology and neonatal intensive care. The first formal attempt to measure QoL is often credited to Lansky and co-workers,[11,12] and their very simple clinician- or parent-completed measure remains in use today (see chapter 4). (Parents are asked to choose one of 11 descriptions of their child's play activity.) From these isolated and simple studies has emerged the increasingly large and sophisticated literature concerning child QoL.

Two related but separate lines of work can be identified. One approach involved efforts to develop generic measures or 'health scales' for use in population surveys of children. Prior to 1990, a number of such instruments had been developed,[13,14] but as noted by Landgraf and co-workers[15] the definition of health outcomes employed in these approaches was rooted in clinical outcomes of morbidity and mortality. It was relatively common practice to combine children with different chronic diseases into a single cohort and compare them with 'healthy' samples.

A second approach involved disease-specific assessment of children with chronic disease, or those undergoing innovative treatments. These early studies did not attempt to define QoL or measure it in any systematic way. Often, QoL was used interchangeably with 'social and psychological problems' or was restricted to demographic or clinical indicators. Studies that attempted to evaluate the impact of a disease on 'lifestyle'[16] or compare marriage rates and number of children between groups differing in disease status[17] might be considered to be precursors of the contemporary interest in QoL.

Increasing interest in QoL is clearly reflected in a review of the literature reported by Bullinger and Ravens-Sieberer.[18] In attempting to review the state-of-the-art with respect to child QoL, these authors conducted a comprehensive literature search using key electronic databases: MEDLINE, EMBASE, Psyindex, PsycINFO, PsyCom, CANCERLIT, AIDSLINE, BIOSETHICSLINE and Somed. They identified over 20,000 publications related to QoL in medicine, of which 13% were relevant to children. Of these, 320 (in German and English) were concerned with QoL related to medical aspects of disease, of which 136 involved empirical work. Most were in oncology or transplantation medicine. Other conditions that received attention included asthma, epilepsy, diabetes and rheumatism. The conditions most frequently studied included those with high mortality rates and diseases in which treatment requires high cost and care. Other often more prevalent conditions were included less frequently. In addition, they noted that little attention had been given to any effects of acute disease on QoL. Bullinger and Ravens-Sieberer[18] also noted that age differences were considered in only 19% of studies.

Three of their findings are of special relevance to this review. The first relates to **respondent**, or the issue of who provides information about a child's QoL. Over 50% of studies used parents as reporters of the child's QoL and 40% used clinic staff. Second, only 19% of studies addressed **age differences** in QoL. QoL was more often assessed for older children (between 13 and 18 years) than for those between 6 and 12 years. The third involves the **assessment of QoL**. "A closer look at how QoL is assessed shows that a multitude of methods are available including interviews, questionnaires, clinical and sociological indicators. Very often, measures originally constructed to assess other variables, such as mood or functions are used."[18] Thus, "instruments that have been developed for another research area are now used to make statements about QoL without taking into account the multidimensionality of the QoL construct".[18] The authors concluded that despite the increasing interest in QoL measurement, most published work may best be described as 'opinion pieces' or reports about the development of new measures.[18]

More recently, a number of reviews of both generic[19,20] and disease-specific measures[21,22,23] have been published. These reviews highlight the lack of consensus regarding definition of QoL. While some include general measures of cognitive development, temperament, vulnerability or social activities as indicators or proxy measures of QoL, others adopt more strict criteria, limiting their reviews to a finite sample of comprehensive measures of QoL for which psychometric data are available.

This current report is both more general and more specific than previous reviews. It is more general to the extent that we attempt to consider not only issues of measurement of QoL but also the

application of QoL in clinical context. It is more specific in that we also attempt to answer the four questions specified by the HTA, and consider both the problems and merits involved in grounding the study of child QoL in adult work.

The rationale for measuring QoL in children

In the broadest possible sense there are only two outcomes that really matter to any paediatrician: the **quantity** of a child's life, and the **quality** of a child's life. All health outcomes are only of value because they have an impact on the quality or quantity of life or they are a proxy measure of such an outcome. The distinction is illustrated by four examples.

In the treatment of diabetes, measurement of haemoglobin is of importance because it indicates the tightness of control of blood glucose. The tightness of control of blood glucose is important for two main reasons. First, it is correlated with the probability of having an acute crisis (keto-acidosis) which reduces the QoL by necessitating hospital admission and intervention, and increases the chance of death – reducing the quantity of life. Second, it is correlated with the premature development of chronic complications such as kidney and eye damage, which reduce both the quality and quantity of life.

In the treatment of epilepsy, the use of more than one anticonvulsant treatment may reduce the frequency of night-time convulsions, but at the same time, prevents participation in afternoon sport (because the school refuses to administer treatment and social services have enforced supervised drug administration because of a previous drug error). Any decision about treatment must take account of all factors and not just the achievement of minimum fit frequency. In order to achieve the best overall QoL, treatment decisions must be informed by fit frequency and the child's social and family life.

In the treatment of babies that have suffered oxygen-starvation at birth, the use of anticonvulsant treatments can reduce the number of fits in the first day or two. Anticonvulsant treatments have been widely used on the assumption that fewer fits means less brain damage (quality) or better survival (quantity). Subsequent studies suggest that fewer babies may survive who have been given anticon-vulsants even if treatment is randomly allocated to babies who are identical in every other respect.[24] Thus, although the treatment achieves a measurable and desirable outcome, any real value collapses if the assumption that it improves life quality or quantity is in error.

In the treatment of acute croup, most children are now treated with steroids. Steroids reduce the apparent severity of symptoms at 6 hours after treatment, and shorten slightly the length of stay in hospital. These may or may not be important improvements in QoL. The main reason that the treatment has become established is that these improvements have been assumed to be a proxy indicator of potentially fatal complications. We assume that if they are treated they are less likely to become life-threateningly unwell. In fact mortality, and even the need to incubate and artificially ventilate such children are not influenced by the use of steroids at all.[25] An unexpected downside may be a change in illness behaviour, with parents more likely to return with the child to hospital in similar circumstances.

These examples have implications for clinical practice and research. In **clinical practice** our aims must constantly be critically evaluated in the light of quality and quantity of life. Much of what we hope to achieve fails by this test. It is easy to focus on symptoms or behaviours that are easily measured but do little or nothing to improve the quality or quantity of life. In **research** we must take care to choose outcomes that are readily reflected in improved **quality** of **quantity** of life, or that are proven proxy measures of the same.

In most instances the problem boils down to this: an advantage is demonstrated in terms of a measurable short-term outcome, but when more important long-term outcomes relating to the quality or quantity of life are examined, disadvantages outweigh advantages. There are many established treatments whose use is now controversial because they appear to fail by these standards. Important examples include: human albumin solution in shock, cisapride in gastric reflux, and high frequency oscillation in premature babies.

Surely then QoL measures (along with mortality measures) are the 'gold standards' against which all other health outcomes must be assessed. Both clinical decisions and the conduct of research can only be assisted by the development of good measures of the QoL in childhood.

Epidemiology of chronic disease

Advances in medical research have changed the emphasis in healthcare from diagnosis and management of infectious disease to prevention and control of chronic conditions. Improvements in medical care mean that survival rates for children with a wide range of chronic conditions have improved significantly.[26,27] The most dramatic improvements are frequently recognised in paediatric oncology, where survival rates have increased from less than 50% in the 1960s to over 70% in the 1990s,[28] but survival rates have also considerably improved for cystic fibrosis (CF), heart disease and many other conditions. However, while major improvements in medical and surgical care have resulted in improved survival, this is sometimes associated with disability, emotional problems or learning difficulties for some children.[29]

While improvements in survival must be welcome, this needs to be set against the apparent increased incidence of some conditions. Examples of this include asthma,[30] obesity,[31] Attention Deficit Hyperactivity Disorder[32] and diabetes.[33] In 1960, parents reported that 2% of their children had a limitation in activity due to a chronic health condition, but this had increased to 6% by 1995.[29] It is not always clear that these figures reflect a true increase, and undoubtedly some may be attributable to more sophisticated diagnostic techniques. In addition, there may have been some change on the part of paediatricians and parents to report symptoms or behaviour problems.

These changes in the epidemiology of childhood disease have led to a call for new outcome measures that reflect a more holistic approach to management and recognise the undesirable side-effects associated with many treatments. The demand for such measures also reflect contemporary views about the relationship between mind and body, and acknowledge the critical link between physical and psychological health.[34] Outcome measures are needed which reflect the fact that mortality is no longer an appropriate end-point when considering the efficacy of medical intervention.

It is critical that treatment should not only increase life expectancy but also improve QoL. While survival statistics have long been considered the gold standard as far as paediatric medicine is concerned, it has to be acknowledged that many children who survive disease or injury as a result of 'heroic' medicine subsequently experience considerable morbidity. Treatments for chronic disease affect children's QoL directly to the extent that they need to attend hospital appointments and undergo painful procedures. Where lengthy hospitalisation is necessary, opportunities to participate in normal activities can be very limited. In addition, there are indirect effects. Physical appearance and body image can be affected by the disease or its treatment. Many conditions including CF, cancer and renal disease can affect growth. Treatments may be associated with long-term complications; survivors of cancer are vulnerable to cardiac, endocrine, and fertility problems among others.[35] Treatment may also make the child feel listless and tired, aggressive, and learning may be compromised.[36] Children may experience limited physical skills and consequently be unable to take part in everyday social and physical activities. The way in which children respond to such adversity is difficult to predict. For every child with cancer who refuses to go outside the house until her hair has regrown, another impresses everyone with her determination to carry on as usual.

Comparing outcomes in clinical trials

The changing epidemiology of childhood illness resulting from both improved living conditions and innovations in management and treatment of disease have led to improvements in the survival for children with a wide range of diseases. As survival rates have improved, particularly in oncology, it has been recognised that simple survival statistics are no longer adequate. More sensitive and comprehensive measures of outcome are needed. The question is whether a child on one treatment has a better QoL than a child undergoing a different treatment. We might want to know about any differences in physical activities, and whether families find it easier to integrate one treatment into family life more than another. Comprehensive measures of QoL that might allow us to distinguish between different treatments, where there are no implications for survival, are potentially of real value.

Evaluating interventions

QoL measures may also be of value in the community setting, where they have been used, for example, to evaluate the impact of nurse-led interventions.[37] This emphasis on the child's perspective, may give a broader picture compared with relying on more traditional indicators such as school absence. For children with significant learning or behavioural difficulties, intervention programmes may focus on improvement in intelligence quotient (IQ) or academic attainments. Even where this cannot be achieved, it may be possible to improve the child's QoL by increasing self-esteem or confidence to deal with

social situations. For patients with neurological conditions (e.g. epilepsy, brain tumours), families may feel that improvements in emotional status and social functioning are as critical as measures of physical or cognitive functioning.

Assessing the outcomes of new treatments

Intensive care medicine has made important contributions to the survival of critically ill patients, but requires expensive equipment and a large staff. There are important quality assurance issues that have to be resolved involving equity of access to care in different regions, how the distance to a Paediatric Intensive Care Unit may affect referral patterns, and the social and financial impact on families who have to travel long distances to receive services. As regards neonatal intensive care, there is a need for further longitudinal work to look at the outcome beyond the age of 2 years and well into childhood if the more subtle consequences of being born too soon are to be understood.

Changes to services are frequently introduced with no attempt to determine how far these changes improve QoL for the individual concerned. Cochlea implantation is a case in point. The benefits of this treatment need to be distinguished from the confounding effects of the intensive rehabilitation programme that accompanies surgery. The dissemination of relevant QoL research in this field would help to re-align existing service provision to accommodate such innovations in practice.

Palliative care

There are times when the distinction between active and palliative care becomes blurred. If QoL is important during treatment, it follows that QoL must also be important during palliative care. The most immediate concern may be about symptom control, but both achievement of symptom control and the means by which this is achieved, have implications for QoL. For adolescents, some degree of control over pain may contribute to QoL. There are other QoL issues, however: what do children understand? how fearful are they about the future? Perhaps more than in any other situation, the need is for a holistic concept of QoL, which includes symptom control, but also recognises the social and emotional needs of the child and family.

QoL as a reflection of the child's perspective

Recent Government reports emphasise the importance of involving children in their own healthcare and taking their views into account.[38] We have identified a range of situations where it may be important to measure the child's QoL. The justification for including assessment of QoL when evaluating new treatments is clear, but we demand more than this of QoL measures. We require also that they reflect the child's view, that they allow us insight into the child's experiences. Previous recommendations that research with children cannot be justified unless it offers direct benefit have been moderated.[39] Research that emphasises children's perceptions may be justified in itself. Clinicians, and perhaps to a lesser extent parents, may recognise late effects such as infertility, cataracts or growth impairment as an undesirable but almost inevitable consequence of cancer care, but these may be less acceptable to children. For them, the benefits (extended survival) may not be seen to justify the costs, in terms of compromised fertility or vision. In any assessment of the value or efficacy of medical care, it has to be recognised that traditional indicators of medical functioning do not necessarily reflect patient's reports about their own health. Thus, the impact of modern medicine is not limited to change on a finite number of physical parameters, but must be considered in relation to the child's overall QoL.

For many children with a chronic condition, medical follow-up is routinely encouraged even though there are no real expectations of any significant change in the child's health. Thus, it is recommended that survivors of childhood cancer attend annual check-ups. Expectations about the reasons for these appointments may differ between patients and clinicians. From the clinicians' point of view, follow-up may be justified on more psychological grounds. It is important to help survivors adjust to the consequences of surviving a life-threatening condition, as well as inform them about their past disease and the possible future consequences. Clinicians may see the focus of these follow-ups on health promotion, rather than conventional disease detection. Epidemiologists may see the rationale in terms of justifying change in design of subsequent protocols, by linking specific late-effects with different treatment protocols. From the patient's point of view, the purpose of follow-ups may not be clear, or understood only vaguely in terms of detecting relapse. Indeed, for patients themselves, it may be unclear that any tangible benefits are associated with the visit. There may even be some costs. For example, evidence about the potential toxicity associated with chemotherapy resulted in some survivors of childhood cancer being asked to

undergo cardiac function tests to determine any impact of previous treatments on cardiac function. Such tests, even when negative, may compromise QoL by raising questions about current and future health status, while offering little if any practical advice. Similarly, parents of multiply handicapped children may routinely keep appointments with the paediatrician, but leave with little practical help, apart perhaps from some reassurance that the child is not deteriorating. In the current climate of healthcare costs, it is important to ask how far such visits are justified. Even if the paediatrician is unable to suggest any new treatment, does the visit benefit child or family QoL at all? A strictly cost-orientated Health Service may need to provide such justification for follow-up of this kind. If it is not possible to cure a condition, exactly what is achieved from the consultation?

In all these situations, it is reasonable to also ask how the child feels about treatment. This very psychological orientation enhances the value of QoL in paediatric medicine, but adds significantly to its complexity. Attention to arguments of this kind have resulted in a 'paradigm shift' in criteria to evaluate outcomes.[18] Changing attitudes in the Health Service and the increasing power of support groups have both been influential in forcing a sea change in how information is communicated and shared with patients.

While most of this activity has been conducted with adults, there is a vocal and powerful lobby making similar demands for child patients. A very limited body of work suggests that children want information about their disease[40] and surgical treatment.[41,42] Research with healthy children suggests that from approximately 9 years of age they are able to consider the consequences of medical treatment and give valid consent.[43] Accurate measures that reflect the impact of disease and treatment from the child's perspective are urgently needed and could become a useful additional measure of outcome of individual randomised studies. Critically, if such measures are advocated, the aim must be to take this information into account when planning future randomised studies. The measurement of child QoL must not be seen to be a limited academic exercise, but one that has 'real-life' implications.

In the review that follows, it may be important to bear in mind the range and diversity of situations in which QoL is used at a popular and more academic level. There is a need to consider QoL in conjunction with clinical measures and survival statistics, in order to make comprehensive decisions about treatment options. However, there are many other contexts where QoL considerations are crucial. Increasingly it is argued that QoL is a relevant outcome in itself. In evaluating this review, the reader is asked to consider how far it is possible for any single concept to fulfil all that is asked of QoL measures.

Summary

- From a historical perspective, the measurement of children's QoL has received less attention than that given to adult QoL.
- The concept of QoL is discussed loosely in the popular press, but also has been considered by different professional groups. The concept of health-related QoL draws on ideas from philosophy, economics, sociology and psychology, as well as medicine.
- Two approaches to measuring child QoL can be identified; one subsequently associated with generic and the other with disease-specific measures.
- A number of reasons can be identified to account for the emergence of QoL as an important outcome measure in Health Services work. These include increased survival rates, particularly in chronic and life-threatening conditions and following traumatic injury; recognition that treatment should not only increase life expectancy but also improve QoL; limited correlation between morbidity and patient satisfaction; and demands for patients to be more involved in decision-making and self-care.
- QoL measures may be of potential value in comparing outcomes in clinical trials, evaluating interventions, commissioning programmes of care, assessing the outcomes of new treatments, and simply to aid understanding of the child's point of view.
- It may prove difficult to identify any single way of measuring QoL that is appropriate for all of these situations.

Chapter 2

Issues in measuring QoL in children

Aims and scope

A number of decisions needed to be made before beginning the review. As discussed in the previous chapter, a critical question involved our definition of QoL. The literature suggested some overlap between QoL and related concepts, including health status or well-being. Our first task therefore was to distinguish between these related concepts, with a view to making decisions regarding inclusion in the review. Second, having established the relationship between QoL and related concepts, we consider the relative merits of generic and disease-specific measures. Third, as outlined in chapter 1, there are different ways in which QoL measures may be used. We argue for the need for measures to discriminate (in clinical trials), to evaluate (different treatments) and to predict (outcomes for survivors). In evaluating QoL measures, we need to take into account the different purposes for which they were developed. Fourth, these qualities are not the only standards against which a measure can be judged. We draw on the work of a number of authors who consider the qualities, or **performance characteristics** of different measures. As a result, we are able to define a set of inclusion and exclusion criteria to guide the literature searches for this review.

Determining inclusion and exclusion criteria for the review

Drawing on the research literature concerned with definition and measurement of adult QoL, we identified some key themes that need to be considered before conducting the review. These include:

- the distinctions between QoL and related concepts
- differences between generic and disease-specific measures
- defining the qualities or performance characteristics of measures
- clarifying the purpose of measures.

These distinctions have been discussed extensively in the past,[44] but for current purposes we adopted the framework outlined by Feeny and co-workers[45] because these were more sympathetic to the issues as they relate to measurement of children's QoL. These themes are also reviewed in order to identify the inclusion and exclusion criteria to be adopted for the review.

Distinctions between QoL and related concepts

The terms health-related QoL, health status and functional status are often used interchangeably.[46,47] To the extent that QoL is in part a reflection of functional status (inability to take part in daily activities will be associated with poor QoL), measures of functional status, ability to perform activities of daily living, and health status are all related to QoL.

Functional status

Functional status has been defined as an individual's ability to perform normal daily activities that are essential in order to meet basic needs, fulfil usual roles, and maintain health and well-being.[48] An individual's functional status may include two concepts of related interest: **functional capacity** and **functional performance**. Functional capacity represents an individual's maximum capacity to perform daily activities in the physical, psychological and social domains of life, whereas functional performance refers to the activities people actually do during the course of their daily lives.[49] Both functional capacity and performance can be influenced by biological or physiological impairment, symptoms, mood and other factors,[48] as well as by health perceptions. For example, individuals who are well but who view themselves as ill may have a low level of functional performance in relation to capacity.[49] Critically, measures of functional status focus on functional ability and overt behaviour, as opposed to subjective experience; they assess what individuals can actually do, rather than what they feel able to do.

Health status

Health status is a broader concept, reflecting an individual's relative level of wellness and illness, taking into account the presence of biological or physiological dysfunction, symptoms, and functional impairment. Most measures include some key dimensions such as physical function, sensation (e.g. vision, hearing, taste), self-care

and dexterity, cognition, pain and discomfort, and emotional and psychological well-being. These dimensions have been defined as 'narrow, within the skin', as the focus is on health status and the capacity for living accorded to the individual as a consequence.[50] Others have objected to this narrow, 'within the skin' approach, arguing that health status measures should also include assessment of social interaction, resilience and vulnerability.[48,51,52]

Health status measures can potentially be used to:

- characterise the health of communities and allow comparisons in terms of how well healthcare systems meet the needs of the population
- compare major subgroups of the population in order to detect systematic differences in health
- monitor the impact of health systems on the health of the population
- assess the adequacy of interventions designed to improve health, and
- serve as a screening tool to detect otherwise unrecognised pathology in individuals, and monitor changes associated with health service interventions.

Many of these goals might also be fulfilled by QoL measures. However, in addition to these goals, QoL measures have the potential to allow assessment of the patient's perspective of the disease or treatment. Thus, QoL measures add a dimension of personal judgement, which is not an integral part of health status measurement.

Well-being

Functional and health status measures tend to focus on assessment of negative health states. As a consequence, patients may not have the opportunity to report the 'good' things in their lives, and conclusions may therefore fail to account for the complete spectrum of behaviour. In recent years, there have been attempts to develop more comprehensive measures, which allow for description of patients' coping resources and positive well-being. Proponents of these approaches emphasise that measures of 'well-being' that focus on positive aspects of health are needed. Many of these draw on psychological theories of personality, social learning, happiness and optimism.[53]

QoL

Several key ideas define the concept of QoL. First is the idea that individuals have their own unique perspective on QoL, which depends on present lifestyle, past experience, hopes for the future, dreams and ambition. Second, when used in a medical context, QoL is generally conceptualised as a multidimensional construct encompassing several domains.[13, 54–56] This follows from the widely accepted definition of health put forward by the WHO as the state of complete physical, mental and social well-being and not merely the absence of disease or infirmity.[57] The Group goes on to describe QoL as the individual's perception of their position in life, in the context of culture and value systems in which they live, and in relation to their goals, expectations, standards and concerns.[58]

Third, QoL can include both objective and subjective perspectives in each domain.[59] The objective assessment of QoL focuses on what the individual can do, and is important in defining the degree of health. The subjective assessment of QoL includes the meaning to the individual; essentially it involves the translation or appraisal of the more objective measurement of health status into the experience of QoL. Differences in appraisal account for the fact that individuals with the same objective health status can report very different subjective QoL: "The patient's perceptions of, and attributions about the dysfunction are as important as their existence."[60]

To the extent that individuals with significant health impairment may report as good QoL as those with no health problem, inclusion of this subjective component adds a dimension of individual difference to QoL. Great caution needs to be adopted in applying this idea of the uniqueness of individual QoL, because it is inappropriate to define QoL differently for some groups in society: "It is imperative that all definitions of QoL be referenced to the general population both in their conception and operational measures".[61]

At the other extreme, there is considerable literature which challenges the view that if an individual is unable to perform a normative task, QoL is inevitably compromised: "To imply that physically disabled or elderly persons have a poorer QoL than younger or able-bodied individuals is a reinforcement of stereotypes that underlie discriminatory practices".[62]

A good QoL can be said to be present when the hopes of an individual are matched and fulfilled by experience. The opposite is also true: a poor QoL occurs when the hopes do not meet with the experience.[63]

These themes have also been taken up by Bergner[1] who states that: "QoL is enhanced when the

distance between the individual's attained and desired goals is less". It is clear that an individual's goals must be realistic and that the gap can be realistically narrowed. From a therapeutic point of view, it is as possible to improve QoL by helping a patient give up some dreams and accept reality (and restrictions) or to work with another patient to achieve more realistically set goals. A major limitation of this approach may be in the failure to specify when goals are realistic and achievable compared with those that are dreams or lacking in substance. Hayry[64] is particularly critical to the extent that such an approach to QoL may not only be used in rehabilitation but also form the basis of QoL decision-making.

Gill and Feinstein[65] distinguish three ways in which QoL may be assessed in the health context. These include:

- more objective measures, such as clinical indices, that patients would not themselves use or necessarily be aware (such as blood sugar or peak flow)
- functional performance (the ability to perform daily activities about which patients are aware, for example climbing stairs)
- patients' own evaluation of the subjective experience of being able to complete a given activity.

This latter subjective rating of health status is now frequently considered to be the defining characteristic of QoL, in contrast to concepts such as functional or health status.

Questions about the uniqueness of health status and QoL measures

While the absence of a unique definition has been criticised,[1,65] the value of the construct of QoL is not necessarily diminished. While a clinician may assess the efficacy of treatment in physiological terms (e.g. are blood sugar levels within population norms?), patients may place greater weight on how they feel. Thus, traditional indicators of success following growth hormone therapy may be increased height, but for children and their families, the success may be as much related to increased self-esteem or opportunities for work from which the child might otherwise have been excluded. Thus, the issue may be less about how many centimetres have been gained and more about whether children can keep up with daily activities and enjoy life. QoL is difficult to define precisely because the way in which a child or family evaluate the success of the treatment will be unique to their own special circumstances. Height can be measured so easily; changes in a child's self-esteem much less so. Thus, QoL and its measurement can seem nebulous or unscientific compared with traditional end-points. One of the goals of this review is to determine how far measurement of QoL meets the standards, and can be expected to reach the standards, that would be demanded from more traditional measures.

Although there has been considerable debate about the distinction between QoL and health status measures, there has been very little in the way of empirical work. An exception is a study reported by Smith and co-workers[66] who undertook a meta-analysis of the relationships between QoL and health status. This was determined in relation to three domains (mental, physical and social functioning). The analysis was based on 12 previously published studies (seven of which were conducted by the authors and their colleagues). All studies involved adults. These articles were identified from searching four key journals (*Quality of Life Research, Medical Care, Journal of Clinical Epidemiology* and *Social Science and Medicine*).

The meta-analysis indicated that, from patients' perspectives, QoL and health status are distinct concepts. When rating their QoL, patients place greater emphasis on mental, compared with physical functioning. However, when rating their health status, patients place greater weight on physical functioning. Neither QoL nor health status were affected by social functioning.

The authors conclude that at least from patients' perspectives, QoL and health status are different concepts. The results raise important issues for evaluations of medical treatments based on QoL or health status measures, to the extent that results will differ depending on the specific measure employed. Comparable analyses based on studies involving children would be useful to confirm these findings.

Inclusion criteria

The concepts of health status, functional status and QoL are closely related, but have not so far been distinguished conceptually. From adult work, there is some suggestion that patients use health status and QoL measures differently, suggesting that choice of measure is critical. However, given the lack of work aimed at distinguishing between these overlapping measures in paediatric work, we decided to include in this review studies involving any measures of health status, functional status or QoL.

Disease-specific and generic measures

In considering the measurement of QoL, a question arises with respect to the standards against which comparisons are made. Comparisons may be needed within a group of children with the same disease. Thus, we may wish to assess the efficacy of an intervention to improve QoL among children with CF. In these cases, **disease-specific** measures may be appropriate. In other instances, the issue may be about comparisons across different conditions, when **generic** measures may be preferable.

Disease-specific measures

Disease-specific measures include domains that are designed to be valid only for a specified condition.[67] Therefore they maximise content validity and provide for greater sensitivity and specificity. Deyo and Patrick[68] argue that disease-specific measures have greater salience for clinicians, better focus on functional areas of particular concern, and may possess greater responsiveness to disease-specific interventions.

However, disease-specific measures are not appropriate in a number of situations. They cannot be used to compare QoL in children with different conditions. In addition, because of their specificity they may lack sensitivity for the specific condition.[69] Disease-specific measures may have additional limitations. Children may have more than one condition. In these cases, it may be necessary to ask them to complete multiple disease-specific measures. This would be time-consuming, and inevitably result in some duplication in that disease-specific measures tend to include some standard or core-type information. In addition, disease-specific measures are inappropriate for children with rare conditions, for which no disease-specific measure is available. It is not a solution to advocate development of new measures for every condition that can be defined. This would be impractical. Apart from anything else, relatively large numbers of children are necessary in order to establish the psychometric properties of any new measure, and for the rarest conditions, this would simply not be feasible.

Generic measures

If we want to know the QoL of a child compared with the normal healthy population of similar-aged children, generic measures are preferable. This type of comparison is likely to be more acceptable to parents, who do not wish to know how the QoL of their child with asthma compares with that of a child with epilepsy. They want to know how their child's QoL compares with that of the child next door, who has no health problems. Generic measures are designed to be broadly applicable across conditions regardless of severity or treatments.[70]

At least in the adult context, generic measures tend to have been through a more rigorous process of development including item selection, reliability and validity testing.[71] Generic measures permit comparisons across interventions and diagnostic conditions, potentially of relevance when making decisions about resource allocation. They may also have the advantage of offering population norms for comparison across disease states. They also allow dysfunction to be quantified for an individual experiencing several disease conditions.[72] Generic measures may be further categorised into health profiles, or preference based measures.

Health profiles

Health profiles include multiple items, which are grouped into different domains of functioning and can be used for most populations. The domains of health and functioning represented in the profiles may or may not allow for aggregation into single summary scores. Examples of profile measures for adults include the Sickness Impact Profile (SIP),[73] the Nottingham Health Profile[74] and the Short Form-36 (SF-36).[75]

Preference-based measures

Preference-based measures may involve:

- direct assessment of preferences for health states
- use of multi-attribute health status classification systems.

The **direct assessment of preferences** for health is advocated for specific applications, allowing the analyst to incorporate, as with specific instruments, items of particular importance or relevance in that setting. Health economists have concentrated on developing single-score scales of QoL. This is known as the utility approach. It was developed from theories of decision-making[76] to explain how a rational individual makes decisions under conditions of uncertainty. In addition, the conceptual foundations of utility assessment are consistent with the theoretical foundations of cost–benefit analysis in economics.[77] Basic to this approach is the idea that outcomes of healthcare can be expressed in terms of quality-adjusted life-years, or QALYs. QALYs integrate mortality and morbidity data to express health status in terms of equivalents of well years of life. A year of complete

wellness is assigned a weight of 1.0 and weights between 0 and 1.0 are assigned to years in which life quality is reduced. There are a number of methods of assigning these weights (e.g. time trade-off, standard gamble). QALYs are not pure measures of QoL but measures of units of benefit from a medical intervention, combining life expectancy with an index of, for example, disability and distress.[54] The value of QALYs in paediatric as compared with adult medicine is not established.

The use of **multi-attribute health status classification systems** focuses on the patient's level of satisfaction with their health in various domains. This approach differs importantly from traditional multidimensional systems by integrating morbidity and mortality.[50] For example, Feeny and co-workers[78] developed the Multi-Attribute Health Status Classification System, which covers seven attributes: sensation, mobility, emotion, cognition, self-care, pain and fertility. This approach yields a score of the child's QoL for each health state, and is based on judgements of a general population sample of children and their parents' ratings of the relative importance of each of these attributes to children's QoL.[79]

Disadvantages of generic measures

Given their comprehensiveness, generic measures may not be responsive to small changes in children's conditions and clinically relevant aspects of children's lives related to a specific disease condition may be overlooked.[80, 81] In addition, in that generic measures can be longer than disease-specific measures, they may place increased burden on patient time, and may therefore be less acceptable.

In QoL work, generic measures are assumed important in the context of decision-making. They are considered preferable to disease-specific measures where decisions are made involving allocation of resources in a public health context. In contrast, disease-specific measures are needed to provide greater sensitivity and are of value in comparisons of clinical trials or alternative treatments. In practice, there have been few attempts to compare the efficacy of generic versus disease-specific measures in adult work, and less for children.

Battery measures

Measures of related concepts could be used as substitutes for QoL. Indeed, there are currently measures that could be used to assess physical function (e.g. physical symptoms and pain), psychological function (e.g. anxiety, depression or body image), and social function (e.g. employment, social relations or marital status). In the absence of more comprehensive measures, it is therefore possible that a combination of previously validated measures may be considered preferable to a single measure of QoL, particularly where few psychometric data are available. This **battery** or **modular** approach refers to collections of specific measures that are scored independently and reported as individual scores.[82]

Inclusion criteria

The relative advantages and disadvantages of generic and disease-specific measures depend primarily on the specific objectives of measurement in clinical research, practice, or policy analysis. Although it has been assumed that generic and disease-specific measures are preferable in different contexts, there has been no empirical work with children in support of this. In recognition of the potential value of both generic and disease-specific measures depending on the specific goals of the study, both generic and disease-specific measures are included in this review. In the absence of comprehensive measures of QoL, some workers use batteries of related measures. In recognition of this, and in order to answer our fourth question, we also decided to include any study involving a battery approach to measurement.

Defining the qualities or 'performance characteristics' of measures

Paediatricians have traditionally been interested in the QoL of their patients. This is reflected in standard enquiries about general health, which often open doctor–patient consultations. In fact, the ubiquitous greeting 'How are you?' is the quintessential QoL measure! Specifically in paediatrics, consultations will include questions about general functioning such as the child's activities in and outside school. These informal questions have their limitations, because differences in how the questions are asked may result in slightly different responses. 'Is school OK?' is likely to result in a fairly uninformative 'Yes' or 'No' response, whereas 'How have things been at school lately?' should only be asked if more time is available, or if the questioner is really interested in the child's answer!

It is therefore necessary to adopt more standardised approaches particularly if we want to understand changes in the child's QoL over time. In these situations, where we want to be able to understand

changes over time, compare patients with the same disease, or those with different conditions, we need to adopt a more standardised approach to measurement. In considering the inclusion criteria for this review, it was necessary to consider the performance characteristics or requirements for a good measure of QoL. We want to be able to distinguish between informal and more formal measures of QoL in order to include only those which might be expected to fulfil some of these criteria of reliability and reproducibility.

The issue of reliability has been addressed by a number of authors. There is considerable consensus regarding some requirements; for example, measures should be reliable and valid, brief and include facilities for proxy ratings. Other requirements are less frequently cited. A summary of the performance characteristics identified is shown in *Table 1*. In this section, we consider in some detail the most frequently cited requirements necessary for considering the quality of a measure of QoL.

Reliability and validity are the most frequently cited requirements of an acceptable measure of QoL (*Table 1*). In the most simple terms, it is important to know that a measure is reliable (i.e. that children will respond similarly on different occasions), and valid (i.e. that we are measuring QoL rather than some other concept).

Reliability

Reliability may be assessed in three ways: internal consistency, test–retest reliability and inter-rater reliability. Single-item measures (e.g. 'How are you?') may not be satisfactory to the extent that individuals can respond generally (e.g. 'I'm OK') without giving details about different aspects of their lives that are far from OK. Therefore it is

TABLE 1 *Performance characteristics for child QoL measures identified by previous authors*

Performance characteristic	Source
Reliable and valid for the groups of patients for whom it is used	Bradlyn *et al.,* 1996;[83] Eiser, 1997;[84] Jenney *et al.,* 1995;[86] Koot (in press);[23] Mulhern *et al.,* 1989;[85] Spieth & Harris, 1996[20]
Employ self-report whenever possible	Bradlyn *et al.,*1996;[83] Eiser, 1997;[84] Koot (in press);[23] Mulhern *et al.,*1989;[85] WHO, 1993[87]
Allow completions by proxy	Bradlyn *et al.,* 1996;[83] Eiser, 1997;[84] Jenney *et al.,* 1995;[86] Koot (in press);[23] Mulhern *et al.,* 1989[85]
Brief, easy to administer	Eiser, 1997;[84] Jenney *et al.,* 1995;[86] Mulhern *et al.,*1989;[85] Spieth & Harris, 1996[20]
Child-centred and developmentally appropriate	Bradlyn *et al.,* 1996;[83] Eiser, 1997;[84] Mulhern *et al.,* 1989;[85] WHO, 1993[87]
The usability of the instrument must be considered (e.g. the language, reading level, and burden to staff); parents and families should be consistent with the stated objectives of the study	Bradlyn *et al.,*1996[83]
Age-corrected, general population norms	Koot (in press);[23] Mulhern *et al.,* 1989;[85] Spieth & Harris, 1996[20]
Reflect the agreed upon definition of QoL	Bradlyn *et al.,* 1996;[83] Koot (in press)[23]
Cover the full spectrum of behaviours thought to contribute to Qol (e.g. psychological, physical, social functioning)	Eiser, 1997;[84] Koot (in press);[23] Mulhern *et al.,* 1989[85]
Include both objective and subjective assessment	Koot (in press);[23] WHO, 1993[87]
Include a generic core and disease-specific items	Jenney *et al.,* 1995;[86] WHO, 1993[87]
Allow for supranormal performance	Mulhern *et al.,* 1989[85]
Cross-culturally compatible	WHO, 1993[87]
Emphasise health-enhancing aspects of QoL	WHO, 1993[87]

hoped that greater reliability can be achieved by using multiple items to measure the same concept. For example, we might ask children how things are in school, with their friends, and in their relationships with their families. The question then arises as to how far these multiple items really reflect the same underlying concept.

Internal consistency refers to the extent to which the items of a domain or scale assess the same dimension, and is normally measured using Cronbach's alpha. This is a statistical assessment of the correlation between items within a dimension,[88] and tests whether items within a scale correlate positively (i.e. measure the same thing). An internal consistency of 0.70 has been recommended for measures used to detect between-group differences in clinical trials or outcomes research, and greater than 0.90 for interpreting individual scores.[89] In practice, values greater than 0.50 may be considered acceptable.[88] Internal reliabilities for QoL measures included in this review are shown in appendix 4 and 5 and clearly frequently fall short of recommended ideals.

Test–retest reliability is established where individuals complete a measure on two separate occasions, and the two sets of scores are positively correlated. In practice a measure can have high test–retest reliability even when the scores of individuals change over time, provided that individuals broadly retain their relative positions within the group. However, the assumption behind assessing reliability in this way is that individuals have not been differentially affected by anything that has happened to them between the two testing occasions. Measures of test–retest reliability in a clinical trial would provide reassuring evidence that changes in scores in the treatment group are reliable and not due to the chance result of an unstable measure that produces fluctuating scores under unchanging conditions. Low test–retest reliability may reflect actual change among some individuals and not necessarily indicate that a measure has poor reliability.[90] It has been argued that the importance of test–retest reliability is often exaggerated and given greater emphasis compared with sensitivity to change.[91]

Inter-rater reliability is an assessment of the consistency of an instrument when it is administered by different interviewers, and normally measured using the kappa statistic.[92] The time between interviews is usually relatively short to minimise the probability of a real change.

Validity

There is a very significant difficulty in establishing the validity of QoL measures. Typically, when a new measure is developed, validity is established in relation to a previously accepted measure of the same concept. Thus, any new measure of IQ has typically been validated against traditional measures. (This assumes that traditional measures really measure IQ, but this is a separate problem.) With a 'new' concept such as QoL, the absence of traditional measures or gold standards against which new measures can be compared, is a limiting factor in establishing validity in a conventional sense.

Construct validity or the extent to which an instrument is a good representation of the construct is central in the evaluation of any measure, and represents the key distinction between behavioural measurement as a science from non-scientific approaches. Cronbach and Meehl[93] argued that establishing the construct validity of a measure involves at least three steps:

- articulating a set of theoretical concepts and their interpretations
- developing ways to measure the hypothetical constructs proposed by the theory
- testing the hypothesised relations among constructs and their observable manifestations.

Some argue that factor analysis is the ideal method for establishing construct validity.[94] Multitrait-multimethod modelling, factor analysis and confirmatory factor analysis are sophisticated statistical techniques, which are yet to be exploited fully in the measurement of QoL. In many cases, given small sample sizes, factor analysis is not possible and decisions about the organisation of items into domains are made on a more *ad hoc* basis.

There are a number of disadvantages to using factor analysis. There may be many possible ways of subdividing the covariance matrix to produce groups of highly inter-correlated items, and thus to produce alternative, very different factorisations. In practice factors may be difficult to interpret and/or inconsistent across studies. There is a need to have a second, confirmatory sample in which to validate the model suggested by the factor analysis. Confirmatory factor analysis places emphasis upon the prior specification of a model.

Fayers and co-workers[95] argue that less emphasis should be placed on using factor analysis as a way of deriving construct validity. They argue that to a

large extent construct validation may be explored using a combination of less-numerical approaches including patient debriefing questionnaires and consensus interviews with patients and health-care professionals.

Criterion validity can be viewed as a special case of construct validity in which stronger hypotheses are made possible by the availability of a criterion or gold standard measure.[96]

Concurrent validity is the independent corroboration that the instrument is measuring what it is supposed to measure (e.g. the corroboration of a physical functioning scale with observable criteria). Criterion validity is usually explored using a gold standard, and correlations must be high (0.75 is a minimum). Thus, new measures of IQ are routinely validated against established measures such as the Wechsler Intelligence Scale for Children.[97] In the case of IQ, such an approach is possible, but there are difficulties for establishing concurrent validity for measures of QoL, given the lack of previous measures. Thus, concurrent validity is only useful where good criteria exist. Where criterion measures are not available, concurrent validity studies are best regarded as aspects of construct validity.[98]

The issue of determining validity in the absence of a gold standard requires a different approach. **Face validity** refers to subjective assessments of the presentation and relevance of the measure. Determination of **content validity** is more systematic than for face validity. It refers to judgements (usually made by a panel) about the extent to which the content of the instrument appears logically to examine and comprehensively include a full assessment of the characteristics or domains it is intended to measure. Qualitative methods can be particularly valuable in discovering the set of relevant dimensions or in assessing the content validity of a measure.

In order to demonstrate construct validity, it is necessary to show not only that a measure correlates highly with other variables with which it should correlate theoretically, but also that it does not correlate with variables from which it should differ.[99] Discriminant validity involves testing whether the data correlate with results from unrelated measures.

Item discriminant validity is demonstrated when each item on a hypothesised scale is substantially linearly related to the underlying concept being measured. An item-scale correlation, corrected for an overlap of more than 0.40, has been recommended.[99] The item discriminant validity does not directly support the overall ability of a measure to distinguish or discriminate across conditions; rather it provides evidence of the conceptual logic for placing an item within a particular scale relative to other scales within a measure.

Responsiveness is operationalised as the "change in QoL score due to a minimal clinical intervention divided by the fluctuation in QoL score due to error of measurement".[100] Hays and Hadorn[100] argue that responsiveness is an aspect of validity rather than a separate entity. If an instrument is responsive to a clinical intervention, this fact provides some support for the validity of the instrument.[100] A measure can be reliable, but unresponsive.[47,100] Common methods for evaluating responsiveness include comparing scale scores before and after an intervention, or comparing changes in scale scores with changes in other related measures that would be expected to move in the same direction as the target measure. Assessment of responsiveness often involves estimation of effect size or magnitude of change in QoL. Effect size translates the before-and-after changes into a standard unit of measurement.

The more detailed or specifically relevant the measure is to the particular intervention being evaluated, the more sensitive to change it is likely to be. While sensitivity to change is a valued characteristic of QoL measures, it is also important that the measure can produce stable results when there is no reason to expect change. Adequate variability of scores is necessary in order to ensure the sensitivity to differences between respondents.

The responsiveness of evaluative instruments may be compromised by ceiling effects in which patients with the best score may have substantial impairment in their QoL or floor effects in which patients with the worst score may deteriorate further.[47]

Inclusion criteria

The requirements for a satisfactory measure include reliability, validity and responsiveness. There is no single way to establish any of these properties particularly in the absence of a gold standard, as is the case with QoL measures. The criteria for inclusion in this review therefore included evidence that in developing a measure attempts were made to establish some of these properties of reliability, validity and responsiveness.

Purpose of measures

The need to develop a set of standardised procedures to assess hypothetical constructs is not unique to the field of QoL. The problem has long been recognised by researchers interested in measuring attitudes, personality or motivational variables such as self-efficacy or self-esteem. "The basic notions of measurement are not hard to grasp, but they become more difficult to achieve as we move from the natural sciences, through the biological and medical sciences, to the behavioural and social sciences".[101] The result is delineation of a standard set of procedures that are generally acknowledged to be appropriate.

The standard psychometric approach requires the individual to indicate the presence, frequency or intensity of symptoms, behaviours and feelings, usually on a series of Likert scales. Responses to individual questions are aggregated to create homogeneous scales (e.g. physical function) or global summary scores.[102] Results derived from psychometric measures are designed to arrange persons along a continuum of function or well-being. The general purpose is to discriminate levels of functioning between groups and to detect changes in function over time, the very same requirements as defined for measurement of QoL. Measurement qualities include reliability (i.e. the extent to which the measure produces scores that are internally consistent and/or stable over time), and validity, or how well it measures the construct under consideration.

Guyatt and co-workers[47] have proposed that QoL measures can be used for three main purposes. Measures can be used to **discriminate** between individuals or groups with respect to an underlying dimension (discriminative index), to **predict** or **classify** individuals into a set of predefined categories (predictive index), or to **evaluate** or **assess** the magnitude of longitudinal change in an individual or group (evaluative index).

To provide a more concrete example of direct relevance to this review, we draw on a report provided by Rosenbaum and Saigal.[103] If we wanted to know how cerebral palsy (CP) affected QoL, we might need to consider this in different ways. The functional limitations of CP are well-known, but individuals may vary in how far they are consequently restricted in terms of everyday activities. Those with the same apparent degree of functional limitation may differ in self-assessed QoL. We might therefore wish first to measure how far the QoL of the child with CP differed from other children. In this case, we would want to use QoL as a discriminative measure.

Second, QoL measures can also be used to evaluate the impact of different treatments or interventions with children with CP. It may be unreasonable to expect that any intervention would lead to changes in functional capacity for all children, but more realistically, an intervention might lead to broader benefits such as increased self-confidence or ability to enjoy a wider range of activities. The hope might be that QoL measures could fulfil a need in evaluating programmes at such a broad-based level.

Third, it may be important to be able to predict outcomes in order to provide appropriate education and social opportunities for the child. Given limited resources, there would be advantages in being able to predict which children would benefit from an intervention.

In more general terms, measures that reflect discrimination are useful in cross-sectional studies where there may be a need to distinguish the burden of morbidity between groups or individuals. For these purposes, a QoL measure is needed not only to show that a measure correlates highly with other variables with which it should correlate theoretically, but also that it does not correlate with variables from which it should differ.[99]

Different QoL measures may be required when the purpose is to predict the score on another measure at the same point in time. These measures, which are more suitable for prediction, are valuable when determining the relationship between a standard form of a measure and a new short form, or when developing simpler measures involving less patient burden. Finally, evaluation measures are needed for longitudinal or prospective studies involving clinical trials. Here it is essential that the measure can identify change in QoL over time (either improvement or deterioration) where this occurs, but also shows little variation in the scores of individuals whose condition remains stable over the period of study.

In addition to these main purposes, as argued in chapter 1, QoL measures may be useful to enhance understanding of the child's experience.

Inclusion criteria

QoL measures can be used to discriminate between individuals or groups, to predict outcomes, or to evaluate the impact of a treatment or intervention. These properties have most frequently been determined as part of the process of developing a measure. In deciding on the inclusion criteria for this review, we therefore included any study where attempts were made to determine discrimination, evaluation or prediction of an established or emerging measure of QoL or health status or to describe the child's perspective.

Respondent issues

To the extent that QoL is subjective, there are strong arguments in favour of eliciting data directly from children wherever possible. It is not possible for anyone else to have insight into the child's unique experiences, and there are circumstances (in school for example) where neither parents nor clinicians have the information that is necessary (see *Box 1*). From quite an early age, children may learn that certain topics upset their parents and are better not discussed at home. Some of the very earliest reports about how children cope with cancer suggested that they learned pretty quickly that it was better not to discuss this with their parents.[104] For any child, part of the normal growing-up experience involves establishing their own friends and values independently of their parents. Comprehensive evaluation of QoL is therefore dependent on developing methods that allow for self-completion whenever a child is old enough or well enough.

Although the individual is the 'best' reporter of their own QoL, it may not always be possible to elicit QoL information directly from children as they may lack the cognitive and linguistic skills necessary to complete measures (see chapter 5). In these cases, it may be necessary to seek information from 'proxies' usually clinicians, parents or teachers. While parents' views about the child's QoL may be useful, interesting and indeed valid, there are no reasons to suppose that they can be used as substitutes for the child's own views. Parents' views will be based on different information, different expectations and reflect different experiences. They will also be affected by their own mental and physical health. Despite these obvious differences that must exist between children and their parents, validation of new QoL measures have often been reported in terms of positive correlations between parents and children's reports.

***BOX 1* A case study of Jo**

Jo (16 years) was recovering from surgery to remove a bone tumour in her lower leg. The tumour had been successfully removed and the diseased bone replaced with a metal prosthesis. Six months later, according to Jo, she was making an excellent recovery. She was having difficulty running, but otherwise she was riding her bike, going out with friends and doing everything she wanted to do.

Her mother told a different story. Jo was often sick and had not returned to school full-time. She could ride the stable bike in the Physiotherapy Centre, but would be unable to ride a normal bike on a road. Yes, she was seeing her friends, but only when her father drove her around. The family had re-organised the house so that Jo could sleep downstairs and therefore not have to climb upstairs.

The dilemma created here is between the picture painted by Jo of someone coping well and 'back to normal' and that painted by her mother. Such discrepancies are not unusual but remain difficult to explain. How do we decide exactly who is telling 'the truth?' Jo who wants to convey the impression that she is fine, or her mother who emphasises the difficulties her daughter is experiencing? Part of the mother's concerns stem from her anxieties about the future and these may (or may not) be unfounded.

Inclusion criteria

Although QoL information should ideally be provided by the individual, it has often been assumed that children are unreliable informants and therefore adults, usually mothers, are asked to provide information for their child. There are also circumstances where children are too young or too ill to provide information themselves. While it may be necessary to rely on parents for information, there are no reasons to suppose that QoL data provided by parents and children will be the same. Even so, in developing new measures of QoL, attempts are often made to establish comparability between parents and children. It was therefore decided to include any study which reported QoL information provided by child or proxy respondent or both. This decision was also necessary in order to answer the third question of the review: To what extent do child self-reports correspond with the assessments made by parents and carers?

Additional criteria

Given the scope of our review, it was necessary to make further decisions regarding the definition of chronic disease and childhood. In the event, we adopted the traditional definition of chronic disease as one lasting for 3 months or more and for which there is no cure.[105] While recognising inter-relationships between disease, handicap and disability, we attempted to adhere to the requirements of the review, and focus on issues of QoL in physical disease. Thus, the review does not attempt to consider QoL issues in relation to disability or handicap.

Legal definitions of childhood may differ from that adopted in clinical practice. Even here, definitions may differ and be more flexible for children with long-standing or chronic conditions.[106] For simplicity, we adopted the legal definition, including research involving children of 18 years or less.

Summary

On the basis of the literature reviewed in this chapter, the following inclusion criteria were adopted.

- **Health status, functional status and QoL** Measures of health status, functional status and QoL were included on the grounds that some measures called QoL are in reality health status, and to some extent the reverse is also true. To the extent that no empirical work has been conducted to date that allows for satisfactory distinctions to be made between these concepts, any study claiming to involve assessment of health status, functional status or QoL was considered for inclusion. However, studies were excluded where QoL was used purely as an indicator of physiological function (such as blood sugar level). (See appendix 3 for other examples.)
- **Generic and disease-specific measures** In recognition of the potential value of both generic and disease-specific measures, the decision was taken to include both generic and disease-specific measures. We also included studies involving battery measures in order to answer our fourth question: How feasible and reliable are proxy measures of QoL in different disease contexts?
- **Psychometric properties** The requirements for a satisfactory measure include reliability, validity and responsiveness. There is no single way to establish any of these properties particularly in the absence of a gold standard, as is the case with QoL measures. The criteria for inclusion in this review therefore included evidence that in developing a measure attempts were made to establish these properties of reliability, validity and responsiveness. In addition, in recognising that QoL can be useful in enabling us to understand the child's perspective, we included any study where this was the goal.
- **Respondent issues** It was decided to include any study that reported QoL information provided by child or proxy respondent or both. This decision was also necessary in order to answer the third question of the review: To what extent do child self-reports correspond with assessments made by parents and carers?
- **Chronic illness** Criteria for inclusion were based on the International Classification of Disease class 10 (ICD-10) diagnoses. Following standard recommendations, a chronic disease is defined as one that lasts or is expected to last for at least 3 months in any given year, and for which there is no cure. In addition, survivors of illness diagnosed earlier in childhood (e.g. malignant disease, neonatal intensive care, end-stage renal disease) are included. This approach means that studies of long-term survivors of chronic disease are included.
- **Age of children assessed** Studies were included if the children assessed were aged 18 years or younger. Studies that include both children and adults without analysing data from children separately were excluded.

Chapter 3

Review methodology

Aims and scope

Based on the literature reviewed in chapter 2, we defined the inclusion and exclusion criteria for the current work. A search strategy was then devised with the aim to identify a broad range of papers potentially of relevance to the issues. The papers were coded for relevance to the key questions.

The search strategy was developed to provide a comprehensive list of publications, both published and unpublished, which have employed measures of QoL or proxy measures of QoL for children. Our initial overview of the history of measurement of QoL for children indicated that the first generally acknowledged measure of QoL for children was reported by Lansky and co-workers.[11] To cover the possibility of earlier work the starting date for retrieval was set at 1980. A general search strategy was employed, in order to identify any papers that were relevant to the definition, measurement or application of QoL measures in children.

It was not possible to review the papers identified in the conventional sense of applying an established methodology as used by the Cochrane groups. This was due to the heterogeneity of the studies identified, and the broad focus of the review questions. This is intended to be a comprehensive review, and one that presents an accurate balance of opinion, but inevitably includes our own judgements and opinions.

Method

Search procedure

Search strategies were devised using the appropriate keywords and combination of keywords (see appendix 1). These were applied in combination using the logical operators specified by each database. In order to ensure a comprehensive recall across a range of measures, the terms health status, well-being and functional status were included. For the same reason, we specified individual chronic conditions in addition to general terms such as chronic disease and illness. Adoption of these very broad concepts resulted in good sensitivity but poor specificity. The subject headings 'quality of life' and 'children' were used when searching the Cochrane Controlled Trials Register (CCTR).

Searches were restricted to English language papers as to do otherwise would greatly increase the costs of the review. However, papers published in English but which originated from non-English speaking countries were included. In these cases, English versions of the QoL measures were available, but the published data were usually based on the original version.

Electronic databases

The following databases were searched from 1980 to July 1999:

- MEDLINE via WebSPIRS
- BIDS ISI Science Citation Index
- BIDS ISI Social Science Citation Index
- PsycLIT via WebSPIRS
- CCTR
- meta Register of Controlled Trials (URL: http://www.controlled-trials.com).

Handsearches

The electronic databases were supplemented by handsearching relevant journals and cross-referencing with reference lists in identified articles. In addition, information was requested from key workers in the field (see Acknowledgements). This identified work in progress and papers that had been submitted for publication.

In recognising that issues of QoL have been of interest to clinicians and behavioural scientists, the following key journals in British and American paediatrics, as well as journals focusing on psychological issues of chronic childhood disease and QoL were identified. As a result, the following journals were handsearched:

- *Archives of Disease in Childhood* [Jan 1990 – July 1999]
- *Child: Care Health and Development* [Jan 1990 – July 1999]
- *Journal of Pediatrics* [Jan 1990 – July 1999]
- *Journal of Pediatric Psychology* [Jan 1986 – July 1999]
- *Pediatrics* [Jan 1990 – July 1999]
- *Quality of Life Research* [1992 – July 1999].

Internet search engines were used to identify links to QoL research by entering key words. The following websites were identified:

- Clinician's Computer Assisted Guide to the Choice of Instruments for Quality of Life Assessment in Medicine (URL: http://www.qlmed.org/index.html)
- International Society for Quality of Life Research (URL: http://www.isoqol.org)
- Measurement of Health Related Quality of Life (URL: http://www.fhs.mcmaster.ca/hrqol/qolintro.htm)
- American Thoracic Society: Quality of Life Resource (URL: http://www.atsqol.org/frmain.html)
- Health related Quality of Life: Mapi Research Institute (URL: http//www.mapi-research-inst.com/research/data2.htm).

Summary of inclusion and exclusion criteria

As a result of the literature review described in chapter 2, the inclusion and exclusion criteria adopted for the review are as follows.

Inclusion criteria

- Measures of QoL, health status or well-being
- Measures that possessed minimum psychometric properties (i.e. some reliability and/or validity data) or assess the child's point of view
- Single (generic or disease-specific) or proxy measures (batteries)
- Measures that include facility for completion by child or proxy or both
- The presence of an ICD-10 diagnosis of a chronic disease or condition
- Children aged 18 years or younger.

Exclusion criteria

- QoL measured only by clinical indicators (e.g. haemoglobin level)
- QoL restricted to demographic or environmental indicators
- Review articles or comments about the measurement of QoL in children or adolescents.

Data extraction and synthesis

Where possible, abstracts were screened in order to assess the relevance of articles. Papers that clearly met the exclusion criteria were rejected at this stage. Where abstracts were ambiguous the article was obtained. References that appeared relevant were downloaded into Reference Manager (Reference Information Systems, version 8) and the full article was obtained. Primary research papers that fulfilled the inclusion criteria were coded using a data extraction form, which was designed to preserve as much of the original information as possible (see appendix 2). Codings were made by two independent researchers who later cross-checked for errors and omissions.

Primary research papers

As a result of the initial screening the abstracts of 255 potentially relevant studies were identified, and these were downloaded into Reference Manager (*Figure 1*). An additional 14 references were obtained from other sources (requests for articles in press and identifying articles in references of obtained articles). Background and review papers that were identified during the screening process were obtained in order to inform the writing of the review.

The application of the inclusion criteria defined above resulted in a total of 137 papers being retained for the review. (Information on the excluded papers is given in appendix 3).

Results

Of the 137 papers included in the review, 43 involved the development of a new measure, 79 reported further development and application, and 15 assessed QoL based on a battery approach.

Identification of measures of QoL

Forty-three measures of QoL, health status or well-being were identified that met the inclusion criteria. Descriptive characteristics of these measures are provided in *Tables 2* and *3*. Further information in terms of their structure, format, and psychometric properties is given in appendices 4 and 5, as well as references for any publications associated with each of the measures (whether as part of development or the application of the measure.)

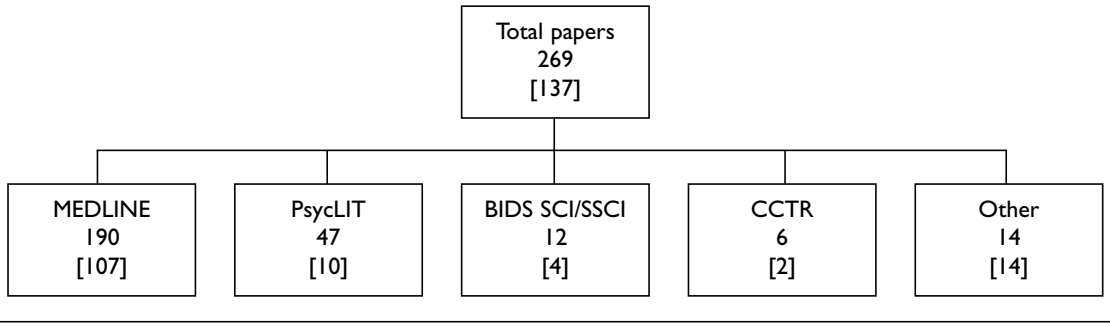
Generic and disease-specific measures

Of the measures retrieved 19 were generic and 24 were disease-specific (*Figure 2*). Multiple measures are available for some chronic conditions: asthma ($n = 4$), cancer ($n = 5$), and epilepsy ($n = 4$). Measures were also identified for arthritis ($n = 1$), Crohn's disease ($n = 1$), diabetes ($n = 1$), headache ($n = 1$), neuromuscular disorders ($n = 1$), otitis media ($n = 1$), rhinoconjunctivitis ($n = 1$), skin disorders ($n = 1$), spina bifida ($n = 1$), short stature ($n = 1$) and spine deformities ($n = 1$).

Proxy ratings

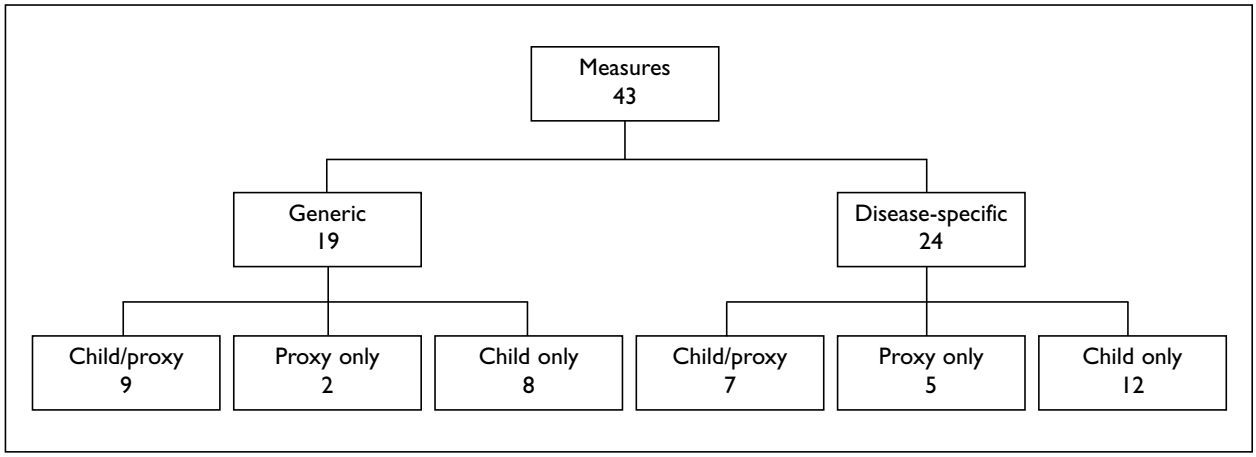
The number of measures are summarised in *Figure 2* in terms of whether they are generic or

FIGURE 1 *Results of the search showing number of papers abstracted (included papers shown in brackets)*



EMBASE was not used in the original search as access was not available. However, this search has subsequently been carried out and did not reveal any additional studies

FIGURE 2 *Breakdown of the number of measures retrieved, classified as generic or disease-specific and by respondent*



disease-specific, and whether they allow for completion by both child and proxy, proxy alone or child alone. There were 15 measures that allowed for completion by children and parent/caregiver. In some cases, these included parallel forms for both respondents. In other cases, wider changes were made in order to make the separate versions more suitable for completion by children.

Domains

The number of domains assessed ranged between 1[11,12] and 17,[110] with the number of domains being comparable for generic (mean, 6.8) and disease-specific measures (mean, 5.7).

Length of scale

The total number of items ranged between 1[11,12] and 153.[110]

Country of origin

Measures were identified that were developed in the USA (n = 18), the UK (n = 8); Canada (n = 8), and The Netherlands (n = 2). Single measures were developed in Germany, Israel, Spain, Sweden, Norway and Finland.

Classification of measures

The measures were described by their authors as QoL (n = 30), health status (n = 8), functional status (n = 2), perception of illness (n = 1), life satisfaction (n = 1), and quality of well-being (n = 1).

Age

Measures were categorised according to the chronological age of the child targeted. Among generic measures, one was targeted at children between 0–5 years, six at children across a broad age range, two at children in middle childhood (about 6–11 years), four at adolescents and four at children from 8 years to late adolescence. In addition, two measures were based on adult measures. Comparable figures for disease-specific measures were 0, 8, 1, 6 and 8, with one measure based on an adult measure.

Identification of papers that address the four review questions

The 137 papers included in the review were coded according to which, if any, of the review questions they addressed. Some papers addressed more than one of the review questions.

- The extent to which adult measures are applied in the evaluation of healthcare interventions was addressed in 14 papers, of which 11 involved modification of an adult measure.
- The extent to which child self-reports correspond with assessments made by parents and carers was addressed in 11 papers.
- Fifteen papers were identified which addressed the feasibility and reliability of proxy measures of various aspects of QoL.

The remaining review question, relating to the appropriateness of adult measures for use with children, was addressed non-systematically in a number of review and empirical papers. In considering this question therefore, we drew on review and background papers, as well as related literature identified from reference lists. These are included in the general references.

Summary

- A search strategy was devised from surveys of previous reviews, with the aim of identifying as broad a spectrum of papers as possible of potential relevance to the key questions.
- The search strategy was developed based on key words and combinations of key words using the logical operators specified by each database. Adoption of broad concepts resulted in good sensitivity but poor specificity. Computerised searches were supplemented by handsearches of key journals.
- As a result of the initial screening, hand-searches and requests to authors for papers in press, 269 articles were identified. Application of the inclusion and exclusion criteria resulted in 137 papers being included in the review.
- Of the 137 papers, 43 were primarily concerned with the development of a new measure of QoL, 79 reported subsequent development of these same measures, and 15 used a battery approach to measure QoL.
- Fourteen papers were identified concerning the extent to which adult measures are applied in the evaluation of healthcare interventions, of which 11 involved modification of an adult measure.
- Eleven papers were identified concerning the extent to which child self-reports correspond with assessments made by parents and carers.
- Fifteen papers were identified concerning the feasibility and reliability of proxy measures of various aspects of QoL.
- The remaining review question, relating to the appropriateness of adult measures for use with children, was addressed non-systematically in a number of review and empirical papers.

TABLE 2 *Summary of the generic QoL measures identified*

Measure and study	Respondent	Age range (years)	No. of items	No. of domains	Name of domains	Reliability	Validity	Origin
Child Health and Illness Profile (CHIP-AE) Riley *et al.*, 1998;[107,108] Starfield *et al.*, 1993,[109] 1995,[110] 1996[111]	Self	11–17	153	6	Achievement, comfort, disorders, resilience, risks, satisfaction	Test–retest Internal consistency	Criterion Convergent	USA
Child Health Questionnaire (CHQ) Landgraf *et al.*, 1997,[19] 1998[112]	Self Parent	10–19 4–19	87 98, 50, 28	12	Physical functioning, role/social functioning (physical), general health perceptions, bodily pain, discomfort, general behaviour, mental health, self-esteem, role/social functioning (emotional), parental impact (emotional), parental impact (time), family activities, family cohesion Global item: change in health	Internal consistency	Concurrent	USA
Child Quality of Life Questionnaire (CQOL) Graham *et al.*, 1997[113]	Self Parent	9–15	15	15	Getting about and using hands, school, out-of-school activities, friends, family relationships, discomfort due to bodily symptoms, worries, depression, seeing, communication, eating, sleep, appearance	Test–retest	Convergent	UK
Dartmouth COOP Functional Health Assessment Charts COOP adolescents: Wasson et al.. 1994,[114] 1995[115]	Adolescent	Adolescent	6	6	Physical fitness, emotional feelings, school work, social support, family communication, health habits	Test–retest	Construct	USA
Exeter Health-Related Quality Life Measure (EHRQL) Eiser *et al.*, 2000,[116] 1999[117]	Self	7–12	16	–	Dimensions yet to be proposed	Internal consistency	Clinical	UK
Functional Status (II) R (FSIIR) Stein & Jessop, 1990[14]	Parent	0–16	43 14	8	Communication, mobility, mood, energy, play, sleep, eating, toileting	Internal consistency	Construct	USA

continued

TABLE 2 contd *Summary of the generic QoL measures identified*

Measure and study	Respondent	Age range (years)	No. of items	No. of domains	Name of domains	Reliability	Validity	Origin
Generic Health Questionnaire Collier, 1997[118]	Self	6–16	25	5	General affect, peer relationships, attainments, relationships with parents, general satisfaction	Internal consistency Test–retest	Concurrent	UK
How Are You? (HAY) Bruil, 1999[119]	Self Parent	7–13	80	5	Physical functioning, cognitive functioning, social functioning, physical complaints, and happiness	Internal consistency Construct		The Nether-lands
KINDL Ravens-Sieberer *et al.*, 1999[120]	Self	8–16	40	4	Psychological well-being, social relationships, physical functioning, every-day life activities	Internal consistency Test–retest	Convergent Clinical Concurrent	Germany
Nordic Quality of Life Questionnaire for Children Lindstrom & Koehler, 1991,[4] Lindstrom & Eriksson, 1993,[121] Lindstrom, 1994[122]	Parent Self	2–18 12–18	74	4	Global, external, inter-personal, personal	Under evaluation	Under evaluation	Sweden
Pediatric Quality of Life Questionnaire (PedsQL) Varni *et al.*, 1999,[123] submitted[124]	Self Parent	5–18 2–18	30	5	Physical functioning, psychological functioning, social functioning, school functioning, well-being	Internal consistency	Construct Clinical	USA
Perceived Illness Experience (PIE) Eiser *et al.*, 1999,[125] 1995[126]	Self	Adolescent	34	8	Perceived impact on physical appearance, interference with activity, disclosure, integration in school, peer rejection, parental behaviour, manipulation, preoccupation with illness	Test–retest Internal consistency	Construct	UK
Quality of Life Profile – Adolescent Version Raphael *et al.*,1996[127]	Self	14–20	54	3	Being (physical, psychological, spiritual); belonging (physical, social, community); becoming (practical, leisure, growth)	Internal consistency	Construct	Canada

continued

TABLE 2 contd *Summary of the generic QoL measures identified*

Measure and study	Respondent	Age range (years)	No. of items	No. of domains	Name of domains	Reliability	Validity	Origin
Sickness Impact Profile (SIP) (Adult measure used by Iorio *et al.*, 1997[128])	Self Parent	3–14	135 136 (adult version)	12	Ambulation, mobility, body care and movement, communication, alertness, emotional behaviour, social interaction, sleep and rest, eating, work, home management, recreation and pastimes	Reported for adults not children		USA
TACQOL Theunissen *et al.*, 1998,[129] Vogels *et al.*, 1998[130]	Self Parent	8–11	108	7	Pain and symptoms, social functioning, motor functioning, autonomy, cognitive functioning, negative global emotional functioning, positive global emotional functioning	Internal consistency	Convergent Clinical	The Nether-lands
The Warwick Child and Morbidity Profile Spencer & Coe, 1996[131]	Parent	0–5	16	10	General health status, acute minor illness status, behavioural status, accident status, acute significant illness status, hospital admission status, immunisation status, chronic illness status, functional health status, quality of life	Test–retest Inter-rater	Construct	UK
Preference-based measures								
Health Utilities Index (HUI) – Mark 2 Feeny *et al.*, 1992[78]	Self Parent	X 6+	7 7	7 7	Sensation, mobility, emotion, cognition, self-care, pain, fertility	Test–retest Internal consistency	Clinical	Canada
Health Utilities Index (HUI) – Mark 3 Boyle *et al.*, 1995[132]	Parent		8	8	Vision, hearing, speech, ambulation, dexterity, emotion, cognition, pain/discomfort	Test–retest	Clinical	Canada

continued

TABLE 2 contd *Summary of the generic QoL measures identified*

Measure and study	Respondent	Age range (years)	No. of items	No. of domains	Name of domains	Reliability	Validity	Origin
16D Apajassalo *et al.*, 1996[133]	Self	12–15	16	16	Mobility, vision, hearing, breathing, sleeping, eating, elimination, speech, mental function, discomfort and symptoms, school and hobbies, friends, physical appearance, depression, distress, vitality	Test–retest	Clinical	Finland
17D Apajassalo *et al.*, 1996[134]	Self	8–11	17	17	Mobility, vision, hearing, speech, breathing, sleeping, eating, elimination, discomfort and symptoms, school and hobbies, friends, physical appearance, depression, anxiety, vitality, ability to concentrate, learning ability and memory	Test–retest	Clinical	
Quality of Well-Being (QWB) Scale Bradlyn *et al.*, 1993[135]	Parent Self	0–18 Adolescent	3 scales and 23 symptoms	3	Mobility scale, physical activity scale, social activity scale	Internal consistency		USA

TABLE 3 *Summary of the disease-specific QoL measures identified*

Measure and study	Respondent	Age range (years)	No. of items	No. of domains	Name of domains	Reliability	Validity	Origin
Asthma								
About my Asthma (AMA) Mishoe *et al.*, 1998[136]	Self	6–12	44	–	Unitary scale not grouped into domains	Test–retest	Construct	USA
Asthma Quality of Life Gibson *et al.*, 1995[137]	Self	Adolescents	30	4	Asthma symptoms, activities limited by asthma, environmental triggers, emotions	Adult based		Australia
Childhood Asthma Questionnaires (CAQ) Christie *et al.*, 1993,[138] French *et al.*,1994[139]	Self	4–7 8–11 12–16	CAQA: 14 CAQB: 22 CAQC: 31	1 2 5	A (distress) B (active QoL, passive QoL) C (active QoL, teenage QoL, distress, severity, reactivity)	Internal consistency Test–retest	Clinical	UK
Pediatric Asthma Quality of Life Questionnaire (PAQLQ) Juniper *et al.*, 1996[140]	Self	7–17	23	3	Activity limitations, symptoms, function	Internal consistency	Concurrent	Canada
Cancer								
Behavioral, Affective and Somatic Experiences Scale (BASES) Phipps *et al.*, 1994,[141] 1999[142]	Parent Self Nurse	5–17	38 14 38	5	Somatic distress, compliance, mood disturbance, quality of interactions, activity	Internal consistency Inter-rater	Clinical	USA
The Miami Pediatric Quality of Life Questionnaire (MPQOLQ) Armstrong *et al.*, 1999[94]	Parent	1–18	56	3	Social competence, emotional stability, self competence	Internal consistency Test–retest	Clinical	USA

continued

TABLE 3 contd *Summary of the disease-specific QoL measures identified*

Measure and study	Respondent	Age range (years)	No. of items	No. of domains	Name of domains	Reliability	Validity	Origin
Asthma								
Pediatric Cancer Quality of Life Inventory (PCQL) Varni *et al.*, 1998[143,144]	Self Parent	8–12 13–18	32	5	Disease- and treatment-related symptoms, physical functioning, psychological functioning, social functioning, cognitive functioning	Internal consistency	Construct Clinical	USA
Pediatric Oncology Quality of Life Scale (POQOLS) Goodwin *et al.*, 1994,[145]	Parent	Pre-school – adolescence	21	3	Physical functioning, emotional distress, externalising behaviour	Internal consistency Inter-rater	Concurrent Clinical	USA
Play Performance Scale for Children (PPSC) Lansky *et al.*, 1985,[11] 1987[12]	Parent	6 months – 16 years	1	1	Play	Inter-rater	Construct Clinical	USA
Epilepsy								
Impact of Child Illness Scales (ICIS) Hoare & Russell, 1995[146]	Parent	6–17	30	5	Treatment, impact on child, impact on parents, impact on family, cumulative impact	Internal consistency Test–retest	Construct	UK
Quality of Life in Epilepsy Scale (adapted from QOLIE-89) Wildrick *et al.*, 1996[147]	Self	8–18	25	5	Self concept, home life, school life, social activities, medicines		Face	USA
Quality of Life in Epilepsy (QOLIE-31) Keene *et al.*, 1997[148]	Parent or self	Not clear	31	7	Emotional well-being, social functioning, energy level, cognitive functioning, seizure worry, medication effects, overall quality of life	Adult measure		Canada

continued

TABLE 3 contd *Summary of the disease-specific QoL measures identified*

Measure and study	Respondent	Age range (years)	No. of items	No. of domains	Name of domains	Reliability	Validity	Origin
Quality of Life in Epilepsy (QOLIE-AD-48) Cramer *et al.*, 1999[149]	Adolescent	11–17	48	7	Impact, memory/ concentration, attitude to epilepsy, physical functioning, stigma, social support, school behaviour, health perceptions	Adult measure		USA
Other								
Juvenile Arthritis Quality of Life Questionnaire Duffy & Duffy, 1997[150]	Self Parent	1.66–18.33 (self-rated in children aged 9 and above)	74	4	Gross motor function, fine motor function, psychosocial function, general symptoms	Test–retest		Canada
Quality of Life in Children with Crohn's Disease Rabbett *et al.*, 1996[151]	Self Parent	8–17	88	6	Disease and its treatment, social, emotional, family, education, future aspects	Internal consistency Test–retest		USA
Children's Dermatology Life Quality Index (CDLQI) Lewis-Jones & Finlay, 1995[152]	Self (with help from parents if needed)	3–16	10	6	Symptoms and feelings, leisure, school holidays, personal relationships, sleep treatment	Internal Test–retest	Concurrent	UK
Diabetes Quality of Life (DQOL) Ingersoll & Marrero, 1991[153]	Adolescent	11–18	52	3	Disease impact, disease-related worries, satisfaction with life		Concurrent	USA
Quality of Life Headache in Youth (QLH) Langeveld *et al.*, 1997,[154] 1999[155]	QLH-Y (youths) QLY-P (parents)	12–18 12–18	71	14	Stress, harmony, fatigue, strength/vitality, depression, cheerful mood, optimism, impact on leisure activities, functioning at home and school, social interaction with siblings, social interaction with peers, general satisfaction	Internal consistency	Construct	Norway

continued

TABLE 3 contd *Summary of the disease-specific QoL measures identified*

Measure and study	Respondent	Age range (years)	No. of items	No. of domains	Name of domains	Reliability	Validity	Origin
Life Satisfaction Index for Adolescents with neuromuscular disorders (LSIA)	Adolescent	12–19	35	5	General well-being, interpersonal relationships, personal development, personal fulfilment, leisure and recreation	Internal Test–retest	Clinical	Canada
Reid & Renwick, 1994[156]								
Quality of Life for children with Otitis Media (OM-6)	Parent	6 months – 12 years	6	6	Physical suffering, hearing loss, speech impairment, emotional distress, activity limitations, caregiver concerns		Construct Responsiveness	USA
Rosenfeld *et al.*, 1997[157]								
Paediatric Rhino-conjunctivitis Quality of Life Questionnaire (PRQLQ) Juniper *et al.*, 1994[158]	Self	6–12	23	5	Nose symptoms, eye symptoms, practical problems, other symptoms, activity limitations	Internal Test–retest	Construct Sensitivity	Canada
		12–17	25	6	Practical problems, non-hay fever symptoms, nose symptoms, eye symptoms, patient-specific activities, emotions			
Quality of Life and Short Stature Pilpel *et al.*, 1995[159]	Self	8–16	45	5	Academic achievement level, leisure activities, physical self-esteem, emotional self-esteem, relationships with peers and family members	Not reported		Israel
Quality of Life in Spina Bifida Questionnaire Parkin *et al.*, 1997[160]	Parent	5–12	44	10	Social, emotional, intellectual, financial, medical, independence, environmental, physical, recreational, vocational	Test–retest	Construct	Canada
	Self	13–20	47	10				
Quality of Life Profile for Spine Deformities (QLSD) Climent *et al.*, 1995[161]	Adolescent	10–20	21	5	Psychosocial functioning, sleep disturbance, back pain, body image, back flexibility	Internal Test–retest	Construct	Spain

Chapter 4

To what extent are adult measures used in the evaluation of healthcare interventions in children?

Aims and scope

In this chapter, we report the results of our searches relevant to the issue of whether adult measures of QoL have been used with children. In addition, we consider the extent to which changes were made to account for developmental differences between children and adults in language and cognitive skills.

Background

There has been a much longer and more focused history of work concerned with adult compared with child QoL. This is reflected in the review reported by Bullinger and Ravens-Sieberer[18] in which they concluded that 87% of published articles were concerned with adults. Given this extensive literature concerned with adult measures, it is essential to consider how far this literature can inform developing work in child QoL. This approach, drawing on the experience of those working on issues related to adult QoL, might be justified on a number of grounds.

First, there are cost considerations. Development and validation of measures is expensive, both financially and in research time. Adaptation of an adult measure could potentially reduce costs. Second, those involved in assessing child QoL may be able to benefit from lessons learned in the development of adult measures. The standards against which the quality of an instrument are judged including basic psychometric criteria of reliability, validity and sensitivity, need to be satisfied whether the focus is on a child or adult measure. Third, there may be direct competition for the allocation of scarce healthcare resources between providers of care for adults and those who care for children. If these provider groups use completely different measures of outcome, it is not possible to compare their contributions directly.[162]

Results

Three studies were identified where adult measures were employed in the evaluation of healthcare of children with chronic illness, with little or no modification (*Table 4*). A further 11 studies were identified where measures for children were developed by modifying previous measures for adults (*Table 5*). These 11 studies drew on nine separate adult measures of QoL.

Adult measures involving no or minor changes

Three adult measures were identified that were used with children. These were the QWB Scale,[178] the SIP[167] and the QOLIE Scale-31.[175] These measures are summarised below.

QWB Scale

The QWB scale is a standardised interviewer-administered QoL measure. It includes three function scales that assess mobility, physical activity and social activity and a list of 27 symptoms or 'problem complexes' that assess symptoms on a specific day (e.g. general tiredness, weakness or weight loss, need for eyeglasses or contact lenses).[178–180] The QWB has been used in three studies to assess QoL in young people with CF.[163–165] In addition, Bradlyn and co-workers[135] used the QWB Scale with children (4–18 years) with cancer.

In the above studies, few changes were made to account for specific QoL issues relating to children compared with adults. In part this is justified on the grounds that the measure was developed to be interviewer-administered. Therefore issues about suitability in terms of length or reading age are less critical than where a measure is developed for self-completion.

However, there are a number of questions that need to be asked about the appropriateness of the measure for work with children. The first relates to the issue of preference weights. There is no evidence to date that the preference weights derived from an adult community sample are relevant for a group of younger children with a chronic disease (see chapter 2). Second, there are ceiling effects suggesting the measure lacks

sensitivity at least for children who are relatively 'well'. This was identified as problematic for children with CF.[164] Third, there are poor correlations between child and parent ratings (see chapter 6). Finally, there is little evidence of concurrent validity, to the extent that there are low or non-significant correlations between QWB scores and clinical indicators or treatment toxicity. Thus, no correlation was found between QoL and forced expiratory volume-1 for children with CF,[164,165] or between QWB and treatment toxicity for children with cancer.[135]

SIP

The SIP[167,168,181] is a 136-item **behaviourally-based health status measure**, developed for healthcare evaluation across a wide range of health problems and diseases, and designed to measure sickness-related dysfunction rather than disease. Scores can be calculated to yield a profile of dysfunction in 12 areas of everyday living, as well as a summary of overall dysfunction.[73, 182] It may be self- or interviewer-administered. It has demonstrated excellent reliability and validity in a number of populations,[183–186] although no data have been reported for children. The developers acknowledge that the length of the measure may compromise patient acceptability.[168]

Iorio and co-workers[128] used the SIP to determine the side-effects of alpha-interferon in 94 children (aged 3–14 years) affected by chronic viral hepatitis. Treatment schedules ranged from 3 to 12 months. The SIP was modified by omitting the 'home management' item because it was thought not applicable to children, and replacing 'work' with 'school'. Children and their parents completed the measure. Given the age range of the sample, many of the children must have required help to complete the measure, but this was not discussed. The SIP was reported to be an appropriate measure, in that it was sufficiently sensitive to detect changes in QoL during the first month of interferon therapy, and during post-therapy follow-up.

The limited use of the SIP as a measure of QoL in children can be attributed to its length, and the fact that, as described, it is a behaviourally-based health status measure. The SIP is not really a measure of adult QoL, and consequently fails to be an assessment of child QoL as well.

QOLIE-31 Scale

The QOLIE Scale[187] assesses seven domains (emotional well-being, social functioning, energy level, cognitive functioning, seizure worry, medication effects and overall QoL). The QOLIE-31 includes a number of items from the SF-36.[75] In that the SF-36 was developed for adult samples, it is important to question how far it is appropriate to use the items without change for work involving children below 16 years of age.

Keene and co-workers[148] did not consider any of these issues but employed the QOLIE-31 to assess QoL in 64 patients under 18 years of age with medically refractory epilepsy who had previously undergone either primary (n = 48) or repeat operative procedures (n = 16). No changes to the adult QOLIE-31 were reported. The QOLIE-31 was administered as part of a structured phone interview to patients or caregivers. No information was given about whether responses were obtained from the children or caregivers. Responses from children and caregivers were combined for the purposes of analyses. Thus, no account was taken of any differences between children and caregivers in their views about QoL, or their ability to complete the scale.

Adult measures that have undergone more substantial changes

Eleven studies involving five generic and six disease-specific measures: (asthma, cancer, dermatology, diabetes, epilepsy, hay fever) were identified where the main aim was the development of a child and/or adolescent QoL measure from an adult version. These 11 studies were based on nine separate adult measures of QoL. This approach contrasts with the studies reviewed above, as attempts were made to ensure appropriateness of the measure for children.

The usual practice here was that a group reported the development of a measure for adults, and then went on to develop a parallel measure for children. Similar procedures in terms of selection of items, choice of response scales are utilised. Given the overlap in philosophy and definition of QoL, and adoption of similar procedures, it is hardly surprising that the resulting measures tend to mimic the adult versions in characteristics such as format, response scale, and number of domains (see *Table 5*).

In evaluating the suitability of this approach, consideration needs to be given to the response burden imposed on children in completing an adult-style measure. This includes the overall length of the scale, as well as the length of individual subscales, the number of domains and items, wording, and the demands inherent in completing the rating scale. Consideration also

needs to be given to the quality of the final scale, particularly whether attempts were made to establish the psychometric properties separately for children rather than assume comparability with the adult measure. Critically consideration needs to be given to the question of how items and domains were developed. Specifically, were items from adult measures assumed to be appropriate for children, or were attempts made to determine developmental appropriateness?

Response burden

It might be expected that those involved in developing measures for children would pay special attention to questions of suitability, in terms of matching the demands of the questionnaire, in terms of its length, complexity, and so on, to the skills of the children. Sadly, developers often make only passing reference to these practical issues. It is worth drawing together, then, some suggestions regarding the 'child-friendliness' of measures.

Length

Several authors report a reduction in the length or specifically in number of domains or items following modification for children. In two cases[11,152] the final length of the measure for children is the same as the adult measure. Marginally longer measures for children and adolescents compared with adults are reported by Apajasalo and co-workers[133,134] and for the CHQ compared with the adult SF-36 by Landgraf and co-workers[15] and for the COOP charts.[114] In their diabetes-specific measure, Ingersoll and Marrero[153] reduced the number of domains from four to three, effectively combining the two separate general worries and diabetes-specific worries for adults into one worries scale for children. Others reduced the number of items and domains. This applies to the measure for epilepsy developed by Wildrick and co-workers,[147] for asthma by Juniper and co-workers,[140] and for hay fever by Juniper and co-workers.[177]

Lengthy scales can be a burden for adults and children. We must expect more missing data with increasing scale length.[188] Despite this, there has been little consistent attempt to reduce length for children. Even for those measures where length was reduced, the overall reduction tended to be quite small.

Changes to the response scale

Most authors use the same length of response scale for children as they do with adults. Thus, items on the Children's Dermatology Life Quality Index (CDLQI) were rated on the same four-point scale as used with adults. However, children were not given the opportunity to select the 'not relevant' option. A five-point Likert scale is used in most other measures, with the notable exception of the measures developed by Juniper and co-workers. Justifying this on the basis of greater sensitivity to change, Juniper has used a seven-point Likert scale throughout her QoL work.

Juniper and co-workers reported that children as young as 7 years had no difficulty in selecting response options based on these scales. Exceptionally, Juniper and co-workers[80] give instructions about how to administer the test with younger children. In these cases, the measures can be interviewer-administered rather than self-completed, with the proviso that the questions should not be paraphrased or explained. If a child appeared to have difficulty with a question it should be repeated verbatim.[80]

Few authors comment on how well children are able to use response scales. Some problems were reported by Apajasalo and co-workers,[134] who noted that children below 8 years had difficulties understanding some of the alternatives, or difficulties choosing between the five alternative responses. In addition, children were not always able to justify their choices. Further consideration of this issue is given in chapter 5.

Time scale

In two of the measures[133,153] no time scale was provided for either the adult or child version. Others favour 6 days or a week,[140,152,177,178] 2 weeks[140] or 4 weeks.[15,114] In adult work, there has been some discussion about the most appropriate time scale to use and suggestions that more accurate responses can be achieved where the time scale is defined and relatively recent.[71] A comparable discussion would be helpful in order to make recommendations about the most suitable time scales for work with children. An exception is work by Juniper and co-workers,[140] who reported that children had some difficulty remembering more than a 1–2-week period.

Quality of measures

Reliability, validity and acceptability

Making any changes to the structure of scales and subscales alters the psychometric properties. Thus, a scale developed to measure adult QoL may have good psychometric properties, but these cannot be assumed to be transferable to other populations including children. For this reason, most authors

make some efforts to determine some aspects of reliability and validity.

Authors typically report internal reliabilities for the total scale score or any subscales (see appendix 3). Test–retest reliability is less-frequently reported.[114,152]

The importance of test–retest reliability needs to be balanced against the need for scales that are responsive to changes in health or well-being during the course of treatment. This has been stressed by Juniper and co-workers in all their work. These workers have included information about responsiveness in their measures for work with children with asthma or hay fever.

Determination of clinical validity, or the extent to which a measure differentiates between children in different health states, is frequently reported. Lewis-Jones and Finlay[152] reported evidence of clinical validity to the extent that it was possible to distinguish between children with different skin conditions. Evidence regarding the clinical validity of the diabetes QoL scale (DQOL) is mixed. Ingersoll and Marrero[153] were unable to demonstrate clinical validity in that the scale did not distinguish between young people in 'good' or 'poor' diabetic control (e.g. blood sugar levels). However, Guttmann-Bauman and co-workers[189] found that QoL scores correlated with mean haemoglobin, suggesting perhaps that the measure reflected differences in physiological indicators. Wasson and co-workers,[114] using the COOP charts reported limited clinical validity to the extent that one of the six domains (health habits) distinguished between those 'at risk' on the basis of drug abuse or antisocial behaviour and normal children.

Expert versus patient-centered approaches

In determining the content of scales, different procedures were adopted. These related to the extent to which the items were taken directly from adult QoL measures, were based on the views of experts or derived from more child- or family-centered perspectives. Lansky and co-workers,[11] Juniper and co-workers,[140] and Cramer and co-workers,[149] especially argue that children are not little adults and that consideration needs to be given to their unique ways of viewing the world and the impact of disease.

Use of professional judges

A number of authors relied solely on professionals to review the items in adult scales and judge their suitability for work with children. In recognising the need for a measure of functioning appropriate for work with children with cancer, Lansky and co-workers,[11,12] modified the Karnofsky Performance Scale.[172] The resulting Play Performance Scale was developed to be suitable for completion by parents of children aged between 1 and 16 years. Like the adult measure, it was developed to be a graduated scale and is rated in deciles (0 to 100). Underlying the rationale for the scale is the assumption that children's ability to engage in age-appropriate play influences peer contacts, skill acquisition, and more general feelings of self-esteem.

The 16D instrument for adolescents aged between 12–15 years[133] and the 17D instrument for pre-adolescents aged 8–11 years[134] are both modifications of the 15D for adults.[170] The adult questionnaire includes 15 dimensions each being divided into five levels. Utility weights (see chapter 2) were elicited from responses from the general public through a three-stage procedure based on Multi-Attribute Utility Theory.

In adapting the measure for adolescents the views of experts in paediatrics, child psychiatry, neurology and health economics were sought. For the 16D, one of the questions in the adult measure was deleted (sexual functioning), some added (physical appearance, friends) and others reworded to be age-appropriate (usual activities became school and hobbies). The 17D (for pre-adolescents) was derived from the adolescent 16D. One of the questions on the adolescent measure was dropped (distress) for the 17D, others were added (ability to concentrate, learning ability and memory, anxiety) and others reformulated (vision, vitality, depression).

To obtain the dimension importance weights, parents (n = 115) of the children participating in the study were instructed to indicate the relative position of each of the 16,[133] or 17[134] dimensions on an importance scale (0–100) by placing the dimension considered most important at the top of the scale. For the 16D, similar procedures were used with children (n = 213) aged between 12 and 15 years.

Thus, in the development of the 16D and 17D changes were made to some of the dimensions assessed, and to the wording of the response scales. Cartoon line drawings were used to illustrate the 17D as an aid for younger children. In addition, the 17D is completed with the help of an interviewer.

Ingersoll and Marrero[153] reported that items in the DQOL for adults[190] were first reviewed by a paediatrician with an interest in diabetes, a nurse practitioner, a social worker and an educational

psychologist. Items that were thought to be of limited value for children were dropped (e.g. 'How often do you worry about whether you will be denied insurance?') and items related to school and peers were included. This revised measure was pilot-tested with 15 young people, aged 11–18 years. Views of children and adolescents were not sought in the development of items for this measure. Further work is therefore necessary to establish how far the measure reflects issues of concern for children rather than issues assumed to be of concern by adults.

Professional and child resources

Consultations with professionals and children were used in determining the items used in a number of measures. Both the Asthma QoL[140] and the Rhinoconjunctivitis QoL measures[158] are based on a methodological approach developed by the group at McMaster University in Toronto.

In both cases, initial discussions with patients, parents and health professionals were conducted, which sought to identify possible functional impairments in childhood asthma or hay fever. From these discussions, appropriate items were selected and 100 children were asked to identify which applied to them. These items were then rated on a scale (1 = 'Does not matter much to me' to 5 = 'Bothers me very much'). Items that were most frequently identified as bothersome were selected for inclusion. Changes were also made to the list of activities that might be affected by disease. For example, vacuuming and gardening were replaced by skipping and roller-blading, respectively.

Previous literature was reviewed and consultations made with professionals and adolescents in the development of the COOP charts for adolescents.[114] Unusually, these authors also asked adolescents their views about ease of responding and likelihood that they would give honest responses.

In the development of the CHQ, Landgraf and co-workers[15] recognised the 'seminal' work in development of adult measures (notably the SF-36), and argued that a similar 'architectural framework' can be appropriately applied to the development of paediatric measures. The authors do not suggest that the conceptual model used to develop adult measures is applicable for development of a measure for children: "Rather the architectural model is a useful platform upon which to build a robust child assessment tool".[15] In fact, Landgraf and co-workers (1996) reported extensive work involving both professionals and families in developing their measure.

Involvement of children

The CDLQI[152] was developed along similar lines as described for the adult measure. However, the initial stages did involve contributions from children. Children aged between 3 and 16 years (n = 169) were first asked to write down, with help from parents if needed, how their skin disease affected their lives. Subsequently ten questions were derived to reflect the most common responses. This draft questionnaire was piloted with 40 children and minor modifications were made. The resulting questionnaire resembles the adult version. There are ten items (e.g. 'Over the last week, how much have you changed or worn different or special clothes/shoes because of your skin?'). Ratings are again made on four-point scales ('very much' to 'not at all'), but unlike the adult counterpart the children's scale does not include a 'not relevant' option.

There are five questions that reflect the same QoL considerations as in the adult measure (itching, scratching, embarrassment or self-consciousness, choice of clothes, impact of treatment). For the children questions about the impact of the condition on 'shopping' or 'looking after the home/gardening' were changed to 'friendships', 'social or leisure activities' to 'going out, playing', 'sport' to 'swimming or other sports', 'sexual difficulties' to 'sleep', and 'problems with your partner' to 'being called names, teasing'. A question for adults concerned with the impact of the condition on work or studying is split into two parts for children; one concerned with the impact on school and the other appropriate for school holidays.

The authors seem to acknowledge the importance of involving children, and used data from children about the impact of their skin condition on their QoL. The information gained during the development phase was used both to change the specific language used and identify activities of greater importance to children. Given their commitment to involving children, it is perhaps surprising that the final measure so closely mirrors the format of that used for adults.

Cramer and co-workers,[149] reported the development of the QoL in Epilepsy-Adolescent version following publication of the QOLIE-89 for adults.[175] Items were adapted from generic instruments, including the SF-36 health survey[75] and the views of seven experts. Cramer and co-workers,[149] conducted some focus groups with adolescents, but essentially relied on items derived from adult measures

(e.g. SF-36), and consultation with 'experts'. Age-appropriate adaptations were achieved by changing phrases to reflect adolescent activities.

Reliance on professionals as informants about the suitability of items for use in children's measures may be limited, in that clinicians may have no way of knowing exactly how illness affects a child. Approaches that at least rely on some input from children are more defensible. However, where interviews or focus groups are conducted with children, they tend to be poorly described and little detail is given about how the content of interviews or focus groups is transformed into questionnaire items. This procedure needs to be formalised in order to ensure that information from children is not selectively retrieved to fit with a model of the measure defined previously for adults.

Simple extrapolation of adult measures is unlikely to be adequate for several reasons. Children are not just little adults. They have a unique perspective on the world, and do not necessarily share adult views about the meaning or implications of illness.[191] At a practical level, measures for adults may simply be too difficult for children to complete. They tend to be very lengthy. They may involve language beyond the child's comprehension, making it difficult for the child to understand the question and the response scale. Most importantly, there are no reasons to assume that the dimensions of QoL relevant for adults are the same for children (see chapter 5 for extended discussion of this). In practice, some modifications need to be made to adult measures for work with a child population. Disagreement exists as to how extensive these changes need to be. Some favour making minimal changes to adult measures.[178] Others argue that satisfactory measures can only be achieved by making more radical changes to take into account the child's perceptions and understanding of the situation.[22] In the following chapter, we attempt to document the kind of considerations that need to be made in order to develop a 'child-friendly' measure. We are then in a better position to judge the appropriateness of an approach based on modification of adult work.

Summary

- Three studies were identified where adult measures were used with very little or no modification. In these papers, there was little if any discussion about the appropriateness of an adult measure for work with children. Changes to adult measures were not made or were extremely few. Where changes were made, they involved very minor and non-systematic changes to improve face validity.
- A further 11 studies (based on nine measures) were identified in which modifications were made to an adult measure. In all cases, the final measure paralleled the original adult measure in format, response scale and dimensions of QoL assessed. This approach can be found in a range of measures from simple screening tools[114] to more lengthy measures.[15]
- The appropriateness of measures based on adult work were considered in relation to response burden. By this we mean the extent to which issues of scale length, or type of response need to be adapted to account for children's language and cognitive skills.
- The extent to which measures took into account children's views varied widely. In some cases, no changes at all were made to the content of adult measures, in others information was obtained from the research literature, professionals and children themselves. Where information was elicited from children, little information was given about how the data were analysed in order to determine specific item content. Although qualitative techniques (focus groups) may be useful in determining the changes needed to be made to make adult measures suitable for children, the way in which qualitative data are analysed is rarely specified.
- Insufficient information tends to be given about the age when children can complete a measure unaided, and responses from children and caretakers are often combined for the purpose of analyses.

TABLE 4 *Studies using adult measures for assessment of QoL of children and/or adolescents*

Study using adult measure	Domains assessed	Sample	Results	Comments
Quality of Well-Being (QWB) Scale Bradlyn *et al.*, 1993[135]	Mobility, physical, social List of 27 symptom–problem complexes	30 children aged 4–18 years who were currently being treated or who were within 2 years of having completed treatment for cancer Interviewer administered to parents in the presence of the child	Internal consistency: Cronbach's alpha ≥ 0.92 for each subscale and the total score No relationship between age and QWB	No modifications to QWB reported Some reliability data reported for this sample
Quality of Well-Being (QWB) Scale Munzenberger *et al.*, 1999[163]	As above	20 children aged 9–20 years with CF, admitted to hospital for treatment of acute pulmonary disease exacerbation Interviewer administered to parents of patients younger than 10 years, and in presence of the children over 10 years	Responsiveness was indicated by significant changes in QWB score, physical, social and symptom–problem complexes domains, and all pulmonary function tests from before to after treatment	No modifications to QWB reported Reliability of the QWB with this population not investigated Re-administered 7/8 days after hospital discharge. When possible follow-up scores were obtained at 6 months and 12 months after discharge
Quality of Well-Being (QWB) Scale Czyzewski *et al.*, 1998[164]	As above	191 caregiver–child dyads Patients were 1–18 years of age with CF	The QWB characterised almost 80% of the sample as having few functional limitations and described no increases in limitations in the older patients groups	No modifications to QWB reported Reliability not reported for this population The study compared the QWB scale with a battery of measures of QoL
Quality of Well-Being (QWB) Scale Czyzewski *et al.*, 1994[165]	As above	199 patients with CF Adolescents ≥ 12 years were interviewed separately from their caregivers	Low-to-moderate correlations between parent and adolescent reports on the QWB over 5 days	Showed no relationships that would allow one to argue for the clinical and practical utility of the QWB over more specific health and psychosocial measures within the child and adolescent CF population

continued

TABLE 4 contd *Studies using adult measures for assessment of QoL of children and/or adolescents*

Study using adult measure	Domains assessed	Sample	Results	Comments
Quality of Life in Epilepsy-31 (QOLIE-31) Keene *et al.*, 1997[148]	Seizure worry, emotional well-being, energy, cognitive function, medication effects, social function, overall QoL	64 patients < 18 years of age with medical refractory epilepsy QOLIE-31 was administered to the patient or a caregiver over the phone	No significant differences in QoL data between the various types of surgical procedure Significant differences were found between those who were seizure-free and those who were not seizure-free for each QoL of domain In addition, QoL was higher in those with a successful outcome than those with a poor outcome	No modifications were reported to have been made to the QOLIE-31 for its use with children Reliability of QOLIE-31 with this population not reported
Sickness Impact Profile (Modified) (SIP) Iorio *et al.*, 1997[128]	Physical, psychosocial	94 children aged 3–14 years with chronic viral hepatitis	Mean total SIP scores before interferon therapy were comparable with those found in adults	Minor medications were made to the SIP The home management item was omitted, and 'work' was replaced by 'school' Reliability with this population not reported

TABLE 5 *Summary of adult QoL measures and their child counterparts*

Measure	Characteristics	Adult version	Child version	Other child version
Quality of well-Being (QWB)	Study	Kaplan & Anderson, 1988[166]	Bradlyn *et al.,* 1993[135] Czyzewski *et al.,* 1994[165] Czyzewski *et al.,* 1998[164] Munzenberger *et al.,* 1999[163]	No other versions
	Domains	Mobility, physical, social activity	Mobility, physical, social activity	
	No. of domains	3	3	
	No. of items	3 27 symptom complexes	3 27 symptom complexes	
	Response scale	3–5 levels of function	3–5 levels of function	
	Time frame	Past 6 days	Past 6 days	
	Age	Birth onwards	Birth onwards	
	Child involvement	None	None	
	Comments	Population-based measure	Application with children reported in 4 studies	
Sickness Impact Profile (SIP)	Study	Deyo *et al.,* 1982,[167] 1983[168]	Iorio *et al.,* 1997[128]	No other versions
	Domains	Physical: ambulation, mobility, body care and movement Psychosocial: communication, alertness, emotional behaviour, social interaction Independent categories: sleep and rest, eating, work, home management, recreation and pastimes	Physical: ambulation, mobility, body care and movement Psychosocial: communication, alertness, emotional behaviour, social interaction Independent categories: sleep and rest, eating, school, recreation and pastimes	
	No. of domains	12	11	
	No. of items	136	135	
	Response scale	Tick if item applies	Not specified	
	Time frame	Not reported	Not reported	
	Age		3–14 years	
	Child involvement		None	
	Comments		No home management domain One item removed One item modified	
Dartmouth COOP charts	Study	Nelson *et al.,* 1996[169]	Wasson *et al.,* 1994[114]	No other versions
	Domains	Physical fitness, daily activities, feelings, social activities, change in health, overall health, social support, QoL, pain	Physical fitness, emotional feelings, school work, social support, family communication, health habits	
	No. of domains	9	6	
	No. of items	9	6	
	Response scale	5 levels of functioning	5 levels of functioning	
	Time frame	Past 4 weeks	During the past month	
	Age	Adult (age not specified)	Adolescents	
	Child involvement	None	At all stages	
	Comments	Illustrated with pictures	Items derived from the Child Health and Illness Profile Similar format to adult COOP Illustrated with pictures	

continued

TABLE 5 contd *Summary of adult QoL measures and their child counterparts*

Measure	Characteristics	Adult version	Child version	Other child version
SF-36/Child Health Questionnaire (CHQ)	Study	Ware & Sherbourne, 1992[75]	Landgraf *et al.,* 1996[15]	No other versions
	Domains	Physical function, role limitation – physical, role limitation – emotional, social functioning, mental health, energy/vitality, pain, general health perception, change in health	Physical functioning, role/social-physical, general health perceptions, bodily pain, parental impact – emotional, role/social emotional, role/social behavioural, self-esteem, mental health, general behaviour, family activities, family cohesion, changes in health	
	No. of domains	9	14	
	No. of items	36	98 (parent forms) 50 28 87 (child form)	
	Response scale	2-, 3-, 4-, 5- or 6-point Likert scales	4-, 5- or 6-point Likert scale	
	Time frame	4 weeks	4 weeks	
	Age	16 years and above	Child 10–18 years Parent 5–18 years	
	Child involvement	None		
	Comments		The CHQ adopts a similar framework to the SF-36, however, the resulting questionnaire is child-centred	
15D	Study	Sintononen & Pekurinen, 1993[170] Adult version	Apajasalo *et al.,* 1996[133] Adolescent version	Apajasalo *et al.,* 1996[134] Child version
	Domains	15D: Breathing, mental function, speech, vision, mobility, usual activities, vitality, hearing, eating, elimination, sleeping, distress, discomfort and symptoms, sexual activity, depression	16D: Breathing, mental function, speech, vision, mobility, school and hobbies, vitality, hearing, eating, elimination, sleeping, distress, discomfort and symptoms, friends, depression, physical appearance	17D: Breathing, speech, vision, mobility, school and hobbies, vitality, hearing, eating, elimination, sleeping, anxiety, discomfort and symptoms, friends, depression, physical appearance, ability to concentrate, learning ability and memory
	No. of domains	15	16	17
	No. of items	15	16	17
	Response scale	5 levels	5 levels	5 levels
	Time frame	Not specified	Not specified	Not specified
	Age	Adult (age not specified)	12–15 years	8–11 years
	Child involvement		Adolescent/adult version: Modifications based on adult views Preference weights obtained from adolescents	
	Comments		Final version similar to adult version	Line drawings included

continued

TABLE 5 contd *Summary of adult QoL measures and their child counterparts*

Measure	Characteristics	Adult version	Child version	Other child version
Asthma Quality of Life	Study	Juniper *et al.,* 1993[171] Adult version	Juniper *et al.,* 1996[140] Child version	Gibson *et al.,* 1995[137]
	Domains	Activities, symptoms, emotions, environment, overall QoL	Symptoms, activities, emotions, overall QoL	Symptoms, activities, environment, emotions
	No. of domains	5	4	4
	No. of items	32	23	30
	Response scale	7-point Likert	7-point Likert	7-point Likert
	Time frame	2 weeks	1 week	Not reported
	Age		7–17 years	
	Child involvement		At all stages	
	Comments			
Performance Status Scale: cancer	Study	Karnofsky & Burchenal, 1949[172]	Lansky *et al.,* 1985,[11] 1987[12]	No other versions
	Name of domains	Work	Play	
	No. of domains	1	1	
	No. of items	1	1	
	Response scale	Deciles from 0 = dead to 100 = normal; no complaint; no evidence of disease	Deciles from 0 = unresponsive to 100 = fully active, normal	
	Time frame	Not defined	Not defined	
	Age	Adult (not specified)	10–16 years	
	Child involvement	None	None	
	Comments		Same format as adult version Based on the Gessell Scales of Development Parent completed	
Dermatology	Study	Finlay & Khan, 1994[173]	Lewis-Jones & Finlay, 1995[152]	No other versions
	Domains	None	None	
	No. of domains	N/A	N/A	
	No. of items	10	10	
	Response scale	4-point-Likert	4-point Likert	
	Time frame	Last week	Last week	
	Age		3–16	
	Child involvement		At all stages	
	Comments			
Diabetes Quality of Life (DQOL)	Study	DCCT Research Group, 1988[174]	Ingersoll & Marrero, 1991[153]	No other versions
	Domains	Satisfaction with life with diabetes; impact of diabetes; worries related to social and vocational issues; worries related to long-term sequelae	Life satisfaction, disease impact, disease-related worries	
	No. of domains	4	3	
	No. of items	46	42	
	Response scale	5-point Likert	5-point Likert	
	Time frame		No information	
	Age		10–21 years	
	Child involvement		Not involved in initial modification of adult items Modified scale assessed by 15 youth	
	Comments			

continued

TABLE 5 contd *Summary of adult QoL measures and their child counterparts*

Measure	Characteristics	Adult version	Child version	Other child version
Quality of Life in Epilepsy (QOLIE)	Study	Devinsky *et al.,* 1995[175]	Cramer *et al.,* 1999[149]	Wildrick *et al.,* 1996[147]
	Domains	Epilepsy, cognitive, mental health, physical health	Epilepsy impact, memory/ concentration, attitudes towards epilepsy, physical functioning, stigma, social support, school behaviour, health perceptions, total summary	Self-concept, home life, school life, social activities, medication issues
	No. of domains	4	9	5
	No. of items	89	48	25
	Response scale	4-, 5-, and 6-point Likert scales	5-point Likert scale	5-point Likert scale
	Time frame		Past 4 weeks	
	Age	Adult	11–17 years	8–18 years
	Child involvement		197 adolescents and parents involved in initial testing	None
	Comments		Similar format to QOLIE-89	Very little relationship with the QOLIE-89
Hay fever (rhino-conjunc-tivitis)	Study	Juniper & Guyatt, 1991[176] Adult version	Juniper *et al.,* 1994[158] Adolescent version	Juniper *et al.,* 1998[177] Child version
	Domains	Sleep, non-hay fever symptoms, practical problems, nasal symptoms, eye symptoms, activities, emotional	Practical problems, non-hay fever symptoms, nasal symptoms, eye symptoms, activities, emotional symptoms	Nasal symptoms, eye symptoms, other symptoms, practical problems, activity limitation
	No. of domains	7	6	5
	No. of items	28	25	23
	Response scale	7-point Likert	7-point Likert	7-point Likert
	Time frame	Past week	Past week	Past week
	Age		12–17 years	6–12 years
	Child involvement		At all stages	At all stages
	Comments	General principles of scale development defined by group and adapted for all 3 measures		

N/A, not applicable

Chapter 5

How appropriate are adult measures for use with children?

Aims and scope

In chapter 4, we identified at least two approaches to measuring QoL in children. We can draw on established expertise in adult work, or place greater emphasis on methods developed more specifically to take into account children's perspectives. The argument in favour of using adult QoL measures as a basis for work with children is both rational and persuasive. There has been a much longer history of development of QoL measures for adults compared with children. Thus it is reasonable to look to adult work as a model for paediatric practice. It is financially attractive, on the assumption that some stages in the development of a child measure can be short-circuited.

In this chapter, we consider both the advantages and disadvantages involved in drawing heavily on adult work. These include differences in the nature and aetiology of diseases that affect adults compared with children, as well as developmental differences that limit children's ability to use measures developed for adults. In particular, we consider the appropriateness of different domains used to assess QoL in adults, for work with children. In addition to the content of adult measures, we need to consider how well children are able to respond to measures, in particular how well they are able to read and understand the items, and use the rating scale provided. Finally, we consider alternative approaches that might result in more child-friendly measures.

Background

The question of how far adult measures can be used for work with children is not restricted to issues of QoL. For example, clinical depression in children was initially thought to be rare,[192] but closely related to adult depression.[193] As the relationship between child and adult depression became better understood, the development of measures of depression in childhood benefited from the knowledge gained from advances made in understanding adult depression. However, it was also recognised that children, parents and teachers differed in their perceptions of the child's depression. These findings mirror current discussions particularly regarding differences between children and adults in perceptions of the child's QoL (see chapter 6).

Differences in aetiology and management of physical disease between children and adults

Despite the same nomenclature, conditions with the same name may differ in their implications for children and adults. Children with diabetes mellitus have the insulin-dependent form, while much greater variability is found among adults. Children with epilepsy may experience seizure disorders,[194] learning difficulties or have other impairments.[195,196] While this is also true for adults with epilepsy, there are fewer implications for education or self-care. Although asthma in children and adults is related, the causes and triggers may differ. Consideration of the aetiology of asthma also illustrates that the nature of disease changes with time. Other conditions, such as CF, affect children. Given the nature of the disease, survival in CF has only recently been such that a significant number of children reach adult life.

Parents have become increasingly responsible for the delivery of treatment (e.g. chest physiotherapy for children with CF or home dialysis monitoring for renal patients) and thus a significant burden of daily care has been taken up by families.[197] For children, this imposes a degree of dependence on parents and may compromise the attainment of age-appropriate autonomy. These differences between children and adults in the aetiology, treatment and management of disease suggests that adult measures may need to be modified or extended if they are to tap the domains of importance in assessing QoL in children.

Developmental considerations

Developmental change makes it difficult to apply any single measure to all age groups from infancy

through adolescence. "Children are a moving target for whom levels of function in various dimensions – and even the dimensions themselves – change with age and developmental stage".[198] Thus there are reasons to advocate that child QoL measures should be developed within the context of a developmental framework.

A child with a chronic illness faces the same developmental tasks and challenges as a healthy child. However, mastery of these tasks and successful coping with the common stresses of childhood are made more difficult by the continuing presence of a disease that can significantly alter the child's physical and mental functioning.[199] It can be difficult to predict the impact of a particular course of treatment on an individual child, because problems associated with prognosis and therapy may be a function of a number of variables, including the child's age at diagnosis, and social and family relationships. At any age, the greatest impact of the disease may occur when the demands of the disease or treatment prevent the attainment of life goals.[200]

Preschool children may have problems in both the short and long term. For example, a diagnosis of cancer in the preschool period may have immediate implications for the child's eating habits and nutritional status, and later implications for growth or learning. Changes to care arrangements and family life may also significantly disrupt the establishment of normal attachments during this period. The physical consequences may also be greater for the younger child. For example, irradiation of the spinal column in children with cancer will result in shortening of growth by an average of 2–3 cm if given at age 10 years, compared with 9–10 cm at 3 years.[201] Treatments that affect the CNS also appear to have greatest impact during this period.

At all ages, the child's understanding of illness may be an important determinant of their QoL.[202] Systematic changes in children's perceptions of health and illness, understanding of the causes of illness, beliefs about how the body works and impact of treatment have been extensively documented.[203] Preschool children can think that their illness has a magical explanation, or is a result of their behaviour, usually wrong-doing.[204] Consequently treatment may be viewed as a form of punishment. Preschool children have limited abilities to understand the reasons for treatment. While it was previously argued that children of this age could not provide reliable information about their pain, this view is now challenged. With developmentally appropriate questions, even 3-year-olds can give descriptive information, while older children can give rich descriptions of the type and severity of pain.[205]

In terms of their understanding of illness, younger children offer less complex explanations and rely less on internal bodily cues to indicate the presence of illness. With age, children offer a more organised description of process and cause. Although sick children often have a relatively sophisticated understanding of their own condition,[206, 207] this does not generalise to their understanding of illness. Thus, the challenge for those working with children is to understand how perceptions of illness change with maturity and experience, and the implications for QoL.

Given the central role of school during middle childhood, illness can have far-reaching consequences for learning and achievement, as well as disrupting social relationships.[22, 208] Long periods of illness, particularly when this is associated with fatigue and loss of energy, can reduce the child's opportunities to participate in sports activities and lead to weight gain, loss of friends and compromised self-esteem.

During adolescence, illness represents a major barrier to the attainment of autonomy and independence. Illness can create a dependency on parents, just at a time when others are exerting their independence. Thus, parent–adolescent relationships can be particularly strained. For adolescents there are the additional demands of the national examination system and the practical problems created when treatment interferes with school work and exam schedules. Although older children are more able to understand the reason for treatment, this in itself may be distressing, raising questions about disability and long-term survival.[22] Adolescents are more able to understand the implications of a life-threatening disease compared with younger children.

This developmental account suggests that the appropriateness of adult measures needs to be considered with respect to differences between children and adults in terms of disease and treatment, and consequently the implications for domains of QoL. This will affect both the content and organisation of the measure. With regard to the measurement of QoL, there are specific questions about the appropriateness of the domains and items. Specifically, these refer to how far domains defined for adults are appropriate for work with children. Consideration needs to be

given to any differences between QoL, and perceptions of QoL in children compared with adults. With regard to developmental level, it is necessary to determine whether the child possesses the cognitive and language skills that are necessary for the completion of the measure.

Appropriateness of domains and items

As argued in chapter 2, QoL is frequently defined as a multidimensional concept involving some combination of physical, cognitive, emotional and social domains. Additional domains, to cover 'spiritual' or 'environmental' issues are advocated by others.[44] Typically, adult QoL instruments include between five and 15 domains, derived from both informal or more philosophical considerations.[209] The process of measuring QoL usually involves some attempt to aggregate items into groups or 'domains' that reflect the chosen dimensions of QoL. In fact the two terms, 'domains' and 'dimensions' are often used synonymously and interchangeably.

At one extreme, Borthwick-Duffy[210] argued that similar domains should be used in the definition of QoL across the whole life-span, although consideration would need to be given to the way in which the characteristics of a domain may be expressed at different ages. In contrast, Rosenbaum and co-workers,[198] argue that "one cannot directly apply to children measures developed for use in adult populations. The dimensions of function, and the levels therein, are simply too distinct and different to be applicable". In practice, the appropriateness of adult domains for children needs to be considered in relation to the different contexts in which children and adults live and work. Children live in families, and responsibility for their health is usually taken by adults. It is generally adults who make decisions about the child's health. In addition, attention needs to be given to developmental differences in the importance of domains. For example sexuality and body image may be a crucial part of QoL in adolescence but may not be so important during earlier childhood. (However, recent reports suggest that children are becoming preoccupied with issues of body image at younger ages.)

Although questions have been raised about whether children would choose comparable dimensions to describe their own QoL as adults,[211] there is little empirical data of direct relevance. The basic domains identified for adults by the WHO[57] have been assumed relevant for children. How far this is justified for work with children needs to be considered.

Physical functioning

Physical functioning represents a domain of QoL with which clinicians are likely to be most comfortable. Physical functioning, at least in theory, may relate more closely to objective measures of disease progression (such as peak flow in asthma, or tumour response following chemotherapy) than psychological functioning. In adult QoL scales, questions about physical functioning tend to focus on the individual's employment and ability to perform self-care and domestic tasks, such as climbing stairs,[75] or vacuuming the house.[171] Emphasis on similar everyday tasks is also made in children's measures. For example, the physical functioning scale of the CHQ[19] includes questions about difficulties the child experiences 'walking one block or climbing one flight of stairs' or 'taking care of self, that is, eating, drinking, bathing or going to the toilet.' The measure described by Varni and co-workers[123] asks children about difficulties they have 'doing chores around the house'.

Thus, the emphasis is on the fulfilment of quite mundane aspects of everyday activity. There is little facility for parents or children to report better than average functioning or excellence in sports. For younger children, there is no mention of play, usually considered vital to development. These kind of questions also lack sensitivity to differences in physical functioning that might be attributed to age, gender or the physical environment in which the child lives. 'Feeling able to walk one block or climb one flight of stairs' may say less about the QoL of a 4-year-old, while it says a great deal about the QoL of a 10-year-old.

Social functioning

That illness affects an individual's social life is beyond question, yet social functioning has less frequently been considered to be a distinct domain in adult QoL measures compared with other domains such as physical functioning.[60] Inevitably, the implications of illness for social relationships and social activities will be different for children and adults. If illness restricts social activities, the consequences may also be greater for adolescents compared with younger children. Staying at home may not be a problem for middle-aged adults in stable relationships, but may have much greater implications for a young adult or adolescent who would, under other circumstances, go out on a regular basis. Friends have a major influence on

the social and emotional functioning of children in both the immediate and longer term.[212] Younger children (aged 9 years) make more new friends during the course of a school year compared with older children (16 years). This seems to be part of a shift to a smaller number of more intimate friendships during adolescence compared with childhood.[213] Critically, poor social relationships during childhood have been consistently linked with negative consequences during adult life.[214] Thus, measurement of QoL during childhood needs to be considered as much for the implications for adult QoL as for the immediate consequences.

Psychological functioning

Questions to assess mood, anxiety or depression are included in most QoL measures. These questions are not sufficient to make clinical diagnoses of anxiety or depression but serve more as broad indicators of overall psychological functioning. Age and gender changes in psychological functioning may need to be considered when interpreting these data. The incidence of depressive symptoms increases during adolescence and is particularly pronounced for girls.[215] As was emphasised in the discussion of the social domains of QoL, consideration of normal development is critical in determining the domains of QoL to be assessed in children's measures.

Potential additional domains

In addition to the domains initially attributed to the WHO[57] there are others that may contribute to QoL in childhood. At the least, there is evidence that domains to measure cognitive function, autonomy and body image are critical, particularly for work with adolescents.

Cognitive functioning

Cognitive functioning, or the ability to learn, understand and remember may well be compromised for children with chronic illness, although the causes are necessarily multifactorial. School absence, reduced motivation, changed expectations from teachers and parents and a combination of the illness and treatment have all been implicated. For example, there is evidence that some children with acute lymphoblastic leukaemia treated with CNS radiotherapy have lower IQ scores than children treated without CNS radiotherapy.[216] More pronounced effects have been reported for those treated when they are younger. Although recommendations about the youngest age when children are vulnerable have varied, a conservative estimate would be 4 years of age (Mulhern, 1994).[216] Age-related differences have also been reported for children with diabetes. Children who experience more frequent incidences of hypoglycaemia before 5 years of age have lower IQ than those who do not.[217]

The effects of theophylline, corticosteroids or beta-agonists and antihistamines on academic performance of children with asthma have been reviewed by Lemanek and Hood.[218] Evidence of subtle damage is apparent, though it appears that children differ in their susceptibility.[219] Effects of steroids on mood and memory are likely to be dose-dependent. Again preschool children seem more vulnerable.[219] Cognitive impairment has been identified in children with paediatric HIV infection,[220] sickle cell disease,[221] cardiac conditions[222] and seizure disorders.[223]

The implications of cognitive functioning for QoL are considerably greater for children compared with adults. Children will be aware of their poor performance and attainments through daily school activities and can also experience teasing or bullying. In contrast, adults are more likely to find work of an appropriate level. For all these reasons, we would suggest that cognitive function constitutes a significant factor in QoL for children.

Autonomy

According to key theories of child development[224] adolescence is characterised by attempts to establish autonomy and independence, close personal relationships with the opposite sex and the definition of work goals. Attainment of these goals is believed to be critical to a successful passage through adolescence. Illness potentially compromises all of these achievements. The adolescent with serious illness may require help in administering treatment, or be dependent on parents for transport to hospital. Treatment may compromise the chance to live independently or learn to drive. To the extent that illness can compromise these achievements, it may be argued that comprehensive assessments of QoL must include items that measure the developing autonomy of the young person and any compromise to the establishment of close relationships and work goals. These issues are unique to adolescence and therefore will not be adequately assessed in any QoL measure developed for adults.

Body image

Although less frequently included than other domains in measures of either child or adult QoL, strong arguments for inclusion of a body image domain have been put forward by Hopwood[225] for

adults, and by Kopel and co-workers[226] for children and adolescents. Some conditions have a direct impact on the child's physical appearance. Those with CF for example may well be small and physically poorly developed compared with their peers. In other cases, the treatment can have a major impact on physical appearance. Typically, children treated for cancer lose their hair and may put on a lot of weight.

Among all adolescents, there is a huge concern with body image and physical appearance. In recognising this, Elkind[227] introduced the concept of 'adolescent egocentrism', to account for adolescents' preoccupation with their appearance and assumption that they are the centre of attention. With some exceptions, measures for adults do not include more than passing reference to issues of body image, or indeed sexuality.

Family relationships

The family has considerable impact on the health behaviour of the young. Children who smoke are more likely to come from families where parents or siblings smoke compared with those who do not smoke.[228] Illness behaviour may also be linked to the behaviour of other family members. There is some evidence that children with unexplained pain come from families where adult members have experienced long-term illness, the suggestion being that children model their own behaviour on that of adults in the family.[229]

Clinically, few would argue that the way in which the child with chronic illness copes with the condition is intimately related to the coping strategies adopted by the family, and this is also borne out by the research literature. Children's coping with pain and behaviour in clinic is linked to specific parenting practices.[230] Child adherence to cancer treatment (measured in terms of cooperation with care and appointment keeping) was related to parental sensitivity and warmth.[231] Thus, for both population-based samples and children with chronic disease, there are indications that variables linked with child QoL are closely associated with family attitudes and coping styles.

Implications for the future

The health of children is important not only for their immediate functioning but also for the implications for future functioning.[191] Thus, poor nutrition and eating habits are associated with poor child health, but are also of concern because of the implications for adult health; overweight children tend to become overweight adults.[232] In the same way, health promoting behaviours, such as regular exercise seem to be laid down during childhood; children who exercise become adults who exercise.[233] Children with poor social skills tend to continue to have difficulties with social relationships as adults.[214] Similar arguments can be made about QoL for children. QoL implications are not limited to the present. Where illness limits a child's ability to participate in everyday activities, social, physical or educational, then the repercussions may be as much, if not more, for their future. For this reason as much as any other, it is important to question how appropriate it may be to draw too heavily on methods, concepts and definitions developed in an adult context.

Developmental level: are questionnaire measures appropriate given children's cognitive skills?

Children's ability to understand questionnaires

Self-report questionnaires are the most widely used and convenient method for assessing QoL in adults. This does not necessarily mean that they are also the ideal way to assess QoL for children.[130] The ability to comprehend and respond appropriately to self-report questionnaires increases throughout childhood and consequently, adaptations to test materials and administration procedures will be necessary for young children. The younger the child the more extensive these adjustments need to be.[234] Furthermore, as a consequence of their disease children may have physical or cognitive limitations that may compromise their ability to make written or verbal responses.

In considering how far adult questionnaire measures are appropriate, it is important to consider the basic skills involved when children are asked to rate their QoL. First, it is important that children can read and understand the question. Second, children need to decide how far the question applies to them or not. They may be asked to indicate how much they agree or disagree with the specific question. Typically ratings are made on a forced choice or continuous Likert scale, or less often as a multiple-choice response. This raises questions about children's abilities to use rating scales and understand the end-points. Third, it is generally recommended by test developers that judgements be made over a defined period of time (usually from 1 week through to one month). This raises questions about children's ability to recall their behaviour and emotions over these defined time periods.

In summary, if standardised questionnaires are to be used, children need to be able to read and understand the specific question, to make decisions about their QoL on a continuous or multiple-choice scale, self-report emotion or pain, and remember what happened over a defined period of time.

Can children read and understand the questions?

Most QoL questionnaires for adults require a relatively advanced level of literacy, estimated equivalent to the 13–14 year level.[235] This may be difficult for many adults, and even more so for children.

Although the level of difficulty of a questionnaire can be assessed by using readability formulas (e.g. Flesch, 1948[236]) reading age is rarely reported by scale developers. For children who are unable to read, a simple solution would be to design the measure so that an interviewer can read the question aloud to the child, in which case the situation is more like a structured interview. Such an approach is open to criticism unless care is taken not to introduce bias into the style of question asking. Where children have limited reading skills there may be strong arguments in favour of developing measures that include provision for completion by a proxy. This would be the easy solution however. An alternative, but equally persuasive argument, may be that where children have limited reading skills, different types of QoL measures are necessary.

Can children use Likert or other rating scales?

Typically, measures require forced choice (e.g. yes/no) responses or ratings on Likert scales. There is good evidence that children as young as 4 years of age can make forced choice responses. For example, 4- and 5-year-olds correctly predicted vulnerability to colds on the basis of proximity to an infected other when asked simple yes/no questions.[237] This contrasts with findings from interview studies where children do not seem to understand the concept of contagion until they are older.[204]

As might be expected, children can experience difficulties completing QoL measures that use Likert or visual analogue scales. They may fail to use the full range of scale available, and notably focus on the end-points, essentially reducing the Likert scale to a yes/no fixed alternative.[238] Despite the frequency with which Likert scales are used in QoL measures, there is little empirical data regarding how adept children of different ages are in using them.

There is also a need for systematic data concerned with children's ability to understand common verbal anchors used in rating scales. Adjustments to the verbal anchors that are used with adults are frequently needed, to ensure that the language and concepts involved can be understood by children at specific stages of development.[239]

There have been few attempts to compare systematically children's abilities to complete QoL measures differing in format. Juniper and co-workers[80] asked a sample of 52 children (aged 7–17 years) with asthma to complete the Pediatric Asthma QoL Questionnaire (PAQLQ)[140], the HUI,[50] the Feeling Thermometer[240] and the Standard Gamble.[240]

Both the PAQLQ[140] and the HUI[50] were completed as described in the original publications (see appendices 4 and 5 for description of these measures). The Feeling Thermometer looks like a thermometer with clearly defined end-points, where 0 is the least preferred health state (death) and 100 is the most preferred health state (perfect health). Children were given three examples of health states related to asthma (mild, moderate or severe) and were then asked to place a marker on the thermometer representing how asthma had affected their own health state during the previous week. Only this last rating was used in the analysis.

For the Standard Gamble, children were asked to think about a health state and then decide whether they would like to remain in that state or take a chance with a new imaginary treatment. It is explained that this new treatment may return them to perfect health immediately with no side-effects but may cause instant death. Initially, if they take the treatment, the probability of perfect health is set at 100%. All those who understood the concept chose to take the new treatment. In theory, the probability of perfect health is gradually decreased until the child decides to remain in the current health state rather than take the new treatment. In practice, the probability of perfect health and death is varied until a point of indifference is found. The indifference probability represents the value that the patient places on health. In this study, the Standard Gamble was used with each of three hypothetical asthma states. These depicted mild asthma ('you have asthma but it hardly bothers you at all. Occasionally you wheeze'); moderate asthma ('you have asthma and it bothers you quite a bit. You wheeze quite often and get out of breath'); and severe asthma ('you have asthma and it's really bothersome. You feel wheezy and it's

difficult to breathe a lot of the time. You can't join in any sports. You have bad asthma attacks and these are quite frightening'). Children also completed a measure of reading vocabulary and word recognition. Parents completed a measure of their own QoL,[140] the Impact on Family Scale[241] and a global rating of change.[242] This latter measure was used to determine any change in the child's asthma between clinic visits.

In terms of understanding and ease of administration, the authors reported few difficulties when using the PAQLQ measure.[140] Children from 7 years of age were able to understand the questions and response options, at least when the measure was administered by an interviewer. As required in this measure, children were able to volunteer individualised activities and appeared able to understand and use the seven-point rating scale. The only problem noted was that some of the younger children had difficulty with making ratings with regard to their feelings 'over the previous week'. Most children completed the HUI without assistance, although correlations with clinical data suggested that validity was poor. The youngest children had difficulty completing the Feeling Thermometer, which appeared more appropriate for children with grade 2 reading skills (aged about 7 years). As might be expected, the Standard Gamble task proved difficult for children with reading skills below grade 6 (equivalent to 12 years). It is possible that the task was made easier by the use of asthma-specific scenarios, and that healthy children asked to consider hypothetical states about which they had no experience might find the task even harder.

This study therefore provides some data regarding the task demands involved in completing different QoL instruments. The suggestion that children are able to use seven-point Likert scales is potentially valuable, and the observation that the youngest children had difficulty using the time frame needs to be taken into account in developing new measures. Given the importance of the issue, the study needs replication, and more detailed work concerned with children's use of rating scales and understanding of verbal anchors is needed.

Can children self-report about emotions, feelings or pain?

We need to question whether children are able to make the kind of judgements about emotions (happy, sad), which form an integral part of QoL assessment. The answer seems to be that they can, although it does depend on the specific way in which the questions are asked.

Children are able to answer simply worded questions about emotional issues, particularly if the response options are represented by pictures.[239] Children as young as 21 months can understand and produce emotion-description adjectives.[243] Ridgeway and co-workers[243] demonstrated over 75% of 6-year-olds use terms for feeling comfortable, excited, upset, glad, unhappy, calm, embarrassed, hateful, nervous and cheerful. Thus it seems that preschool children show substantial understanding of emotions. However, children younger than 6 years will typically assert that only one emotion can be felt at one time, so individuals can feel happy or sad, but not experience mixed feelings. Older children are more likely to examine the situation and its possible emotional consequences more exhaustively, and so acknowledge the possibility of ambivalent emotions.[244, 245]

Similar findings can be reported with regard to children's experience of their own pain. With the use of play and story-telling tasks, it has been shown that children from 18 months of age are able to say that a pain hurts, localise it and make efforts to alleviate it.[246] This literature involving assessment of children's pain, can almost be considered a model for assessments of children's QoL. From a situation in which it was believed that children could not reliably describe their pain, to one in which a number of methods have been satisfactorily developed, the message is that children can self-report, as long as the method is appropriate. In assessing the QoL literature for children, it is important that criteria are identified that are sensitive to the strengths and weaknesses of developing linguistic and cognitive skills. From this perspective, we may gain considerably from the experiences of those involved in developing measures to assess other aspects of children's development (e.g. self-esteem[247]).

Can children remember over the required time frames?

A wide variety of time frames, from a few days to several weeks, have been used in QoL measures. In assessing QoL in adults, there has been some discussion about the most appropriate time frame to adopt. Recall is likely to be more accurate when the time frame is as short as possible, although there are disadvantages to short time frames in that they may be affected by temporary fluctuations in health. Attempting to remember how much pain, anxiety or dizziness was experienced over the preceding month is considerably more difficult to recall than for the preceding 3 days.[71] For children and adolescents, time frames are not always defined, but may vary between 1 and 4 weeks.[15,140] An examination of children's understanding of

time is essential in designing tools that will elicit meaningful responses from children.[248] It has been shown that 3- to 4-year-olds have some understanding of the passage of time, and know about the usual duration of everyday activities.[249]

Innovations to improve child-centredness

Much has been written about how to promote rapport with the child in interview situations.[250, 251] In the case of questionnaires, an attractive layout is important. The environment in which the data are collected should help and not hinder the task. Questionnaire length, simplicity of format, comprehensibility of instructions, length and wording of individual items all contribute to the assessment burden. The time needed to complete a measure may have important consequences for whether children omit items and how appropriately they respond. In addition, young children need sufficient time to think about their response.

Response format

To sustain a child's motivation particular care needs to be taken in the choice of response format.[239] Attempts to reduce difficulties experienced in using rating scales have included the simplification of Likert scales (e.g. reduction in number of response categories), and the use of smiley faces. Reducing the number of points on the Likert scale (e.g. nine points to five points) may be at the expense of sensitivity and seven-point scales are specifically recommended for use in clinical trials.[150] The use of smiley, neutral, and sad faces as simple rating scales is in widespread use for the assessment of pain in childhood, and these have been adopted in measuring QoL in children with asthma.[138]

Pictorial support

Using pictures to support written text is a developmentally appropriate approach for work with children with limited reading skills. Advantages of the pictorial format include engaging young children's interest, increasing understanding, sustaining attention, and all of these may contribute to more meaningful responses.[247] With the aid of pictorial support, the youngest age at which self-report data may be obtained with acceptable reliability is about 4 years.[247, 238] In QoL measures reviewed in chapter 4, pictorial support was provided by Lewis-Jones and Finlay[152] and by Apajasalo and co-workers.[134]

Computer presentation

Given advances in technology and increased availability of computers, presentation of visual stimuli by means of portable computers is becoming possible, and is often assumed to be a method of great attractiveness to children. Measures identified in our search that used this approach have been reported by Eiser and co-workers,[117] and Ravens-Sieberer and co-workers.[120]

Props

The use of props or puppets is also useful for work with young children. Mize and Ladd[252] found that 4- and 5-year-olds produced a greater number of responses when using props compared with verbal accounts alone. However, no studies were identified that have used props in this way specifically to assess child QoL.

Summary

- There are advantages to basing new measures of QoL for children on those previously developed for adults. These include potential value in longitudinal studies where there is a need to determine changes in QoL across the life-span, and the opportunities to draw on established expertise of those working in adult QoL and economy. It costs less if some of the basic stages of item determination used in adult work can be assumed appropriate for children.
- The disadvantages of this approach also need to be recognised. QoL is usually regarded as a multidimensional concept, including physical, social, and emotional domains related to the early definitions of health proposed by the WHO. In adapting adult QoL measures for children, a central question relates to how well the domains assessed in adult measures are appropriate to assess QoL in children.
- The traditional domains, including physical QoL may not have the same meaning for children as adults. In adult work, there is an emphasis on attainment of everyday routine activities which may not fully reflect the range of physical activities typical of normal children.
- Adult measures may fail to tap the specific aspects of QoL that are important to the child. At the least, additional domains to tap autonomy, body image, cognitive skills and relationships with the family are needed. Failure to include these domains may inherently limit understanding of the child in a social context and therefore does not provide a robust framework for the development of measures for children.
- Response scales, wording and format of adult measures may need modifying to account for children's cognitive and language skills.

- A number of innovations were identified which may facilitate scale completion by children. These include the use of pictures or computers rather than paper and pencil measures, and modifications to rating scales to simplify the child's task.

Chapter 6

To what extent do child self-reports correspond with the assessment made by parents and carers?

Aims and scope

It is important to consider the role, or potential use of proxy ratings. First, proxy ratings can be used as substitutes for ratings made by children. This is particularly important when children are too young or too ill to provide their own ratings. Second, proxy ratings can provide important complementary information about children.

In either case, we might ask how far ratings made by children and proxies may be identical, or whether systematic differences might occur. In developing new measures of QoL, it is standard practice to determine the concordance between child and proxy rating. It follows that if the concordance between child and proxy rating is poor, then the measure is in some way inadequate. There are clear holes in this argument; proxies, by whom we really mean mothers, and their children may not agree about many issues.

The question 'To what extent do child self-reports correspond with the assessment made by parents and carers? may therefore need some refining. In this chapter, we attempt to clarify the circumstances in which parents are able to make judgements about their child's QoL, and identify variables that effect the degree of concordance between parent and child ratings. Conversely, we attempt to clarify circumstances in which proxies cannot answer for children. The ultimate aim is to determine how far, and in what circumstances, we can rely on parents for information about the consequences of illness and treatment for the child's QoL. Three further questions therefore seem critical.

- Is concordance greater for some domains (e.g. physical functioning) than others?
- Do parents rate their child's QoL to be better or poorer than their child's own ratings?
- How is degree of concordance effected by child age, gender and illness status?

Background

Advocates of measures that rely exclusively on parents as informants argue that these may better facilitate assessments of children across the age range, compared with multiple measures designed for child self-report at different age levels.[145] Against this, relying on one parent as informant may result in incomplete assessment to the extent that the child's subjective experience and perceptions of QoL may be overlooked.

However desirable it may be to obtain information from children, there are circumstances where such direct assessments are not possible. Children may be too young or too ill to answer questions or complete questionnaires themselves. Ironically, it may be in these particular situations that information about the quality of the child's life is most pertinent. While a 9-year-old may well be able to voice anxieties and concerns about treatment procedures, the 3-year-old is less able to understand explanations and may consequently experience greater distress. In these circumstances, there may be no alternative but to rely on proxy raters. These are usually parents, but other relatives, medical staff and teachers may also contribute valuable information.

The need for proxy raters is not confined to work with children, and there is a relatively extensive literature concerned with the advantages, and limitations of, proxy raters for work with elderly, disabled or terminally ill adults. It may therefore be useful to consider briefly what can be learned from this adult literature.

In their review of this adult literature, Sprangers and Aaronson[253] emphasise the potential value of proxy raters in both clinical and research settings. Clinically, proxy raters are important in considering the QoL of the elderly or terminally ill. In research, proxy raters may be able to play a unique role in resolving common methodological problems experienced in longitudinal studies. Clinicians may be reluctant to involve patients in lengthy assessments of QoL especially as their condition deteriorates. Consequently, missing data is a major problem in analysis of longitudinal work. Proxy raters can potentially provide such missing data.

QoL data, in collaboration with clinical information, may be useful when making medical decisions.

Thus, the decision to withdraw a child from a clinical trial may be made on the basis of information about both prognosis and QoL. Where children are very distressed by treatment, the decision to withdraw them from a clinical trial may be taken earlier compared with a child of similar clinical status but who does not seem to experience a similar level of stress. Conversely, decisions to initiate treatment may also be taken on the grounds of QoL information. Short stature by itself may not be considered enough to warrant growth hormone therapy, but if short stature is associated with compromised QoL, either for the child's current or future functioning, then recommendations to commence therapy are more likely to be made.

In the examples given above, satisfactory decision-making is dependent on concordance between clinician and patient perceptions of QoL. The doctor needs to understand the child's distress about being very short. Adult work suggests that in practice, clinicians have limited skills with which to assess patient QoL.[253] Most typically, clinicians underestimate pain intensity,[254, 255] and both over- or underestimate levels of anxiety and depression.[256] Relatives, too, have limited skills. Their bias is to underestimate both patient QoL and their ability to perform different activities.[257]

In recognising the discrepancies that can occur between patient and clinician ratings of QoL, or indeed between patients and their relatives, efforts have been made to identify characteristics of proxies who make more accurate ratings (where accuracy is defined in terms of match between patient and proxy). In general, accuracy increases where proxies live in the same house as the patient and tends to decrease as the patient's health deteriorate.[258] However, accuracy does not seem to be related in any simple way to demographic factors such as age or gender, socio-economic status or educational level of proxy or patient.[259]

As far as children are concerned, it may be that the accuracy of proxy ratings is dependent on the specific domains of QoL being considered. Tangential evidence for this comes from research involving parental ability to rate children's behaviour problems. Achenbach and co-workers[260] found that parents are more able to judge a child's externalising problems (acting out, aggressiveness), but are much less accurate when it come to judging the child's internalising problems (sadness, anxiety). This might suggest that parents are more able to rate the child's QoL in relation to domains of physical functioning or physical symptoms compared with less-visible domains such as social or emotional functioning.

There is also evidence of some systematic bias in parent's judgements of the child's functioning or emotional response. Parents tend to report that an illness has more negative consequences for their children than children themselves.[261] For example, in the study by Ennett and co-workers,[261] mothers of children with juvenile arthritis consistently rated their child's competence to be more effected by the disease than the children did themselves. In these cases, there are no answers to questions about who is more right. Parents may exaggerate the problem; children may deny it. Research has simply identified the discrepancy and put forward some *post hoc* hypotheses about the cause.

The ability of parents to rate the child's functioning may be dependent on demographic factors such as the child's age, gender or health status. Concordance might be expected to increase with the child's age, particularly as greater verbal skills may facilitate children's abilities to describe their experiences and emotions to parents. There is some evidence that parents are more likely to discuss emotional issues with daughters than with sons,[262] which may suggest that mother–daughter concordance would be greater than that for mother–son. Finally, given the increased dependence that can occur between parents and sick children, we might expect parent–child concordance to be greater for sick compared with healthy children.

Results

As a result of the searches described in chapter 3, fourteen papers were identified which examined concordance between child and proxy. These were based on ten separate measures of QoL and are summarised in *Table 6* (generic measures) and *Table 7* (disease-specific measures). In 11 of the 14 studies, the reason for establishing child and parent concordance was ostensibly to determine the validity of the QoL measure. Exceptions include work by Walker and Heflinger,[263] Glaser and co-workers.[264,265]

In all the studies identified, concordance was explored between child and parent. In addition, concordance between child and clinician or nurse was explored in three of these.[142, 265, 266] Concordance between child, parent and teacher ratings was reported in two studies.[264, 265]

Is concordance greater for some domains than others?

Parent–child concordance was rarely examined for children below 7 years of age. However, Glaser and

co-workers,[264] included some children below 7 years in their clinic (but not the home-based sample).

Good agreement ($r > 0.50$) was generally found between child and parent for domains reflecting physical activity, functioning and symptoms (see *Tables 6* and *7*). Thus, good agreement was reported for physical activities,[119,123,144] physical symptoms,[126] and somatic distress.[142] Bruil[119] reported good parent–child concordance for physical complaints for a sample of chronically ill children, but not for a comparable group of healthy children.

Poor agreement ($r < 0.30$) was reported for a number of domains that reflected more social or emotional QoL issues. These were variously described as 'global functioning-(negative)', 'global functioning-(positive)', and 'autonomy'[130] 'appearance', and 'communication with physician or nurse',[123] 'social functioning',[143] 'cheerful mood', and 'optimism about the future'.[154] Phipps and co-workers[142] reported poor agreement for compliance. Eiser and co-workers[126] reported poor agreement for 'disclosure' (willingness to talk about the illness), and impact of treatment. Poor agreement was also reported for cognitive functioning.[123,130]

Thus, in line with previous research, there is some evidence for greater concordance between child and parent ratings for observable behaviours such as physical functioning, and less for non-observable functioning such as emotional or social QoL. However, there are at least three exceptions. These are provided by Czyzewski and co-workers,[165] Theunissen and co-workers,[129] and Langeveld and co-workers.[154]

Czyzewski and co-workers[165] used the QWB scale[180] to compare ratings made by adolescents with CF and their parents. Parents and their children showed poor agreement in terms of symptom scores. Only low-to-moderate correlations over a 5-day period were reported. Even comparatively objective sections of the QWB ('dressed oneself') were only moderately correlated (30% common variance). When parents and adolescents were asked to identify 'the most undesirable symptom', only 5% common variance was reported (but see chapter 4 for discussion of general problems with this scale).

In the measure reported by Theunissen and co-workers,[129] distinctions were made between health status and QoL, where QoL was defined as the affective evaluation of health status. Thus, to rate health status, respondents first indicate the presence of a problem by choosing one of three alternatives ('never', 'sometimes', 'often'). If a problem was present respondents were then asked to assess the child's emotional reaction by choosing one of four alternatives ('very good', 'not so well', 'rather bad', 'bad'). In contrast to work described above, Theunnissen and co-workers[129] found that children and parents were least likely to agree about physical complaints, for both the health status and QoL scales, compared with any other domain.

Theunnissen and co-workers[129] further argued that concordance would be greater for the more observable health status items compared with QoL items. Overall, significant correlations between parent and child ratings were found. However, QoL correlations were significantly lower than those for health status, particularly for motor functioning and autonomy. As a consequence, Theunissen and co-workers (1998)[129] concluded that parents were less able to rate their child's QoL than they were to rate the more objective health status items.

Langeveld and co-workers[154] reported a large number of parent–child correlations. Of note is the very good correlation for social interaction with peers; a finding not predicted from other work.

This review therefore finds only limited support for the widely held view that parents are more able to judge the child's QoL in terms of physical rather than social or emotional domains. At least two reservations can be identified. First, there were three studies in which there was no evidence of greater concordance for physical compared with social functioning. These involved three different measures of QoL and so findings cannot be attributed to some problem with a specific scale. Examination of the properties of these three measures (see *Table 3*) suggests that two do tend to be quite long and involved. The third (QWB) was not developed for work with children. Second, even where there is greater concordance for physical compared with social functioning, it is clear that there is much greater heterogeneity in how social or clinical functioning is measured compared with physical functioning. This can be seen from the number of different terms used to describe the social/emotional functioning domains. Thus, the apparent relationship may be an artefact of the way in which domains of physical and social functioning are measured in QoL scales. The general implication is that there is reason to examine more carefully how much overlap there is between scales purporting to measure the same domain in different measures. This is a problem for all domains, but may apply particularly to social and emotional QoL.

Do parents over- or under-estimate their child's QoL?

This question was addressed in three papers. Children reported lower QoL than their parents for five of the seven scales assessed in the Dutch QoL Scale (TACQOL; physical complaints, motor functioning, autonomy, cognitive functioning and positive emotions).[129] Children's scores suggested lower health status for domains measuring physical complaints, motor functioning and positive emotions. Further analyses suggested that for both health status and QoL scales, there was poor agreement for physical complaints. For motor functioning and autonomy, correlations were lower for health status than QoL but there were no differences for physical complaints, social and cognitive functioning.

In developing the How Are You? instrument, Bruil[119] distinguished between 'prevalence' and 'quality of performance' as separate contributors to overall QoL. Prevalence items were phrased in terms of 'Have you remembered what you learned in school?' 'Have you done maths assignment?' For a group of children with a chronic illness, parents reported significantly lower QoL than their child, to the extent that they reported lower 'prevalence' of physical activities and cognitive tasks (parents said the child took part in these activities less often than the child reported or less often completed school-based assignments). In addition, parents reported a lower 'quality of performance' on cognitive tasks and social activities than their children ('How well are you able to keep your attention on schoolwork?'; 'How well are you able to play with other children after school?'). In contrast, parents reported fewer physical complaints than their child did. Similar results were obtained for a healthy group, in that these parents also reported fewer physical complaints than their children. However, parents of healthy children reported a better 'quality of performance' on cognitive tasks than their children.

Graham and co-workers[113] also explored concordance between parents and their children as part of the development of the Children's QoL scale. Although QoL was investigated for four groups of children (physical illness, psychiatric disorder, mental retardation and healthy) it was only possible to collect concordance data for the physically ill and psychiatric disorder groups. Concordance was assessed separately for the three indicators of QoL defined by the Children's QoL Scale: function, upset and satisfaction. Correlations between mother and child for 'function' varied between –0.01 and +0.80 for the physically ill group and between 0.15 and 0.79 for the psychiatric group. Similar correlations were found for the 'upset' and 'satisfaction' scores. In subsequent analyses repeated multivariate analyses of variance were conducted between mother and child ratings. No significant differences were reported for function, or satisfaction. (Differences were found for 'upset' scores but are not surprising given that mothers rate their own rather than the child's upset.)

How is degree of concordance affected by child age, gender and illness status?

Age

Only two of the studies examined the effect of age on parent–child agreement. However, the relationship was not simple and the effects of age were moderated by the child's emotional state. Theunissen and co-workers[129] reported that the child's age was related to agreement on the autonomy and positive emotions scales. Older children (10–11 years) with low positive emotions scores agreed less with their parents than the younger children (8–9 years). Conversely, older children with high positive emotions scores agreed more with their parents.

Differences in concordance between children and their parents compared with adolescents and their parents were studied by Varni and co-workers[143] For children and parents, mean concordance was greater for cognitive functioning, followed by psychological functioning, physical functioning, disease and treatment symptoms and social functioning. For adolescents and parents the agreement was greatest for physical functioning, disease and treatment symptoms, psychological functioning, social functioning and cognitive functioning.

Other studies focus on adolescents but do not include comparisons with parent–child data. However findings are contradictory. For example, Eiser and co-workers[126] reported relatively good correlations between parents and adolescents with cancer; Czyzewski and co-workers[165] reported poor correlations between parents and adolescents with CF. Given differences in disease, treatment and methods of data collection in these studies, it is not possible to draw any firm conclusions about the extent of concordance between parents and their adolescents.

Gender

Only one study[129] included an assessment of the role of gender in moderating parent–child concordance. The results are difficult to interpret. Boys

with low autonomy scores and their parents showed poorer concordance than girls with low autonomy and their parents. However, boys with high autonomy scores had higher concordance with their parents than girls with high autonomy scores. No theoretical explanation for these findings is offered.

Illness

Although parents' ratings of the child's QWB scores correlated weakly with clinical indicators of physical health, there were no similar correlations between adolescent's QWB ratings and clinical indicators. Thus, parents' perceptions of the child's QWB may reflect clinical status, but this is not necessarily so for the child.[165]

Many of the children (54%) in the survey reported by Theunissen and co-workers[129] were suffering from minor temporary illness (e.g. influenza). Parent–child agreement for health status (physical complaints or social functioning scores) was greater for children with a temporary illness compared with those who were well. However, agreement between child and parents was not affected by chronic illness (19% of the sample), life events, the position of the child in the family or educational level of parents.

Walker and Heflinger[263] reported concordance between parent and child for somatic symptoms and disability to be much lower for psychiatric and well groups (somatic symptoms: –0.09 and +0.15, respectively; functional disability: –0.08 and +0.12, respectively) than for two groups with abdominal pain (somatic symptoms: 0.47 and 0.48; functional disability: 0.42 and 0.56, respectively). They suggest that this lack of concordance may be due to the fact that these were areas of less concern in the former groups and thus there may have been less communication between parents and children regarding the presence or absence of any symptoms or disability.

This same theme, emphasising the importance of communication was reported by Bruil.[119] For healthy children, intraclass correlations (ICCs) were weaker for domains that were less accessible to parent observation. As this was not found for children with a chronic condition, it was suggested that parents of sick children are more attentive than parents of healthy children. For children with a chronic illness, parents were less positive about their child's functioning than their children. The exception was for physical symptoms, where parents reported fewer symptoms than their children.

The results of the studies in this section provide little consistent evidence regarding the relationship between child and parent ratings as a function of variables such as age, gender or illness status. There is some evidence that parent–child concordance is greater for dimensions of greater importance to the parent–child dyad, particularly for sick children and their parents.

Medical staff as proxy raters

Billson and Walker[266] used the HUI (Feeney *et al.*;[78] see chapter 4) to assess health status from the perspective of doctors, parents and survivors of childhood cancer (aged 2–17 years). Consultations with the doctor were informally structured to enable the doctor to assess the child's health status for each of six attributes (senses, mobility, emotion, cognition, self-care and pain). Some modifications were made to the original measure to ease comprehension for parents and children. Parents completed the measure for children under 8 years of age, while parents and children aged 8 years and above completed the measure together. Children over 14 years completed the measure themselves, with help from their parents if needed.

A total of 48 assessment pairs (patient/parent and doctor) were collected. No deficits on any of the six attributes were identified for 16 patients (33%) on the basis of their own assessments or those of their parents and for 19 patients (40%) on the basis of doctors' assessments. Assessments by doctors identified fewer deficits than those made by parents or patients and this difference was most pronounced for pain. In 17 cases, there was no difference between assessment by the doctor and patient or parent. Disagreements were found on only one attribute for 15 patients and on two attributes for ten patients. The degree of correlation between patients/parents (range, 0.29–1.00) and patients/doctors (range, 0.31–1.00) was not significantly different.

Phipps and co-workers[142] used BASES to determine relationships between parent, nurse and child ratings during initial hospitalisation for bone marrow transplant. There are three versions of the BASES.[141,142] The versions for completion by parent and nurse consist of 38 items; the child version consists of 14 items. All versions examine QoL in terms of five domains (somatic distress, compliance, mood disturbance, quality of interactions and activity). Correlations between nurse and parent ratings on Day 0 were good (27 of 28 correlations were significant at the 0.05 level). On day 14, 22 out of 27 correlations were significant. Further correlations were conducted

between child, nurse and parent ratings. Most correlations were in the low-to-moderate range, but parent–child correlations tended to be more significant than nurse–child ratings. In line with previous findings of greater concordance for observable than non-observable behaviours, agreement was higher for scales measuring somatic distress and activity compared with scales measuring mood disturbance, compliance and quality of interactions.

Teachers as proxy raters

Glaser and co-workers[265] asked parents, teachers and children (previously treated for a CNS tumour) to complete a modified version of the HUI (Billson and Walker, 1994).[266] There was good agreement between teacher and child for domains measuring cognition, hearing, vision and pain. Children reported their QoL in terms of domains of ambulation, speech and emotion to be lower than comparable ratings made by their teachers.

Discussion

As with all work in this area, the extent to which findings can be explained by differences in study design needs to be considered.[253] A number of problems recur in assessing the quality of these studies.

Definitions of QoL domains

Answers to the questions identified are complicated because of variability in the content of domains. This is a general criticism of QoL scales for children, but contributes to difficulties in answering specific questions about the value of proxy raters. Most developers of scales adopt the WHO-derived definition[57] of QoL as a multi-dimensional concept, and attempt to assess domains including physical, social and emotional QoL. However, the precise content of items included in these domains varies considerably in emphasis and generality. Thus, the physical domain of QoL is variously measured in terms of physical activities,[119] physical symptoms,[126] somatic distress[119] or physical complaints.[119,129] The emphasis may be on physical symptoms, participation in physical activities or distress caused by limitations in physical activities. There is even greater variability in nomenclature for social domains (see *Tables 2* and *3*). These reservations need to be borne in mind when evaluating this literature.

Statistical issues

The question of concordance between parent and child ratings has most frequently been considered by determining the Pearson's product–moment correlations between child and proxy ratings. These analyses lack the sophistication necessary to answer questions about parent–child concordance.

Occasionally authors make *a priori* predictions about the expected value of correlations. Thus, Varni and co-workers[143] predicted effect sizes to be larger for externalising compared with internalising problems, and Theunissen and co-workers[129] predicted greater agreement for ratings of health status compared with QoL.

The practice of determining child–proxy agreement on the basis of significance of Pearson's product–moment correlations may be insufficient. Although strong correlations between parent and proxy data demonstrate some validity, they do not assure that patient and proxy ratings are interchangeable in terms of mean values. (Parent and child ratings may correlate but this can mask differences in degree; i.e. whether parents rate the child's QoL to be much worse than children). ICCs may therefore be preferable. The intra-class correlation is a measure of the proportion of overall variability accounted for by the variability among individuals, and is a combined score from the *t*-test and product–moment correlation and ranges from 0 to 1. High ICC suggests little variability in measurement by either parent or child and therefore that parent and child ratings are interchangeable. Given that an ICC of 0.80 or more indicates that a scale is highly reliable between raters, the majority of domains fall short of this, with most correlations in the small-to-medium effect range. Relationships between the magnitude of scores and amount of proxy agreement could give scatter bias or random fluctuation. To account for this, some authors have advocated regression analyses where child reports are regressed against parent reports.

The multitrait-multimethod approach can be used to examine the inter-relationships between assessed items to determine the extent to which data obtained from various sources (e.g. patient, parent, doctor) reflect the same underlying constructs. However, this approach was only employed in one study.[129]

Sneeuw and co-workers[259] argue that to be adequate a methodology requires the separate determination of the reliability and validity of patient and proxy ratings. In some instances, limited reliability of the measurement scales may be the reason why there appears to be poor correlations at the individual level. If the reliability of the measures is low, agreement between respondents can never be high. There is therefore a need

for direct comparisons of reliability estimates based on both patient and proxy ratings.

Rather than focusing on concordance between parent and child, some authors prefer to focus on the degree of difference between raters. In these cases, authors report a series of *t*-tests to determine variability as a function of different domains, or occasionally multivariate analyses have been used.[113] Multivariate analyses have the advantage in that account is taken of the number of analyses being conducted.

Data collection methods

In many cases, researchers fail to specify the circumstances in which data were collected. Questionnaires may be completed while children and parents are waiting for clinic appointments[123,125] or sent by post.[119,129,154]

The practice of sending questionnaires by post for home completion may be justified in terms of simplicity. Clinic time can be saved if patients bring completed questionnaires with them. It might also be argued that families are stressed during clinic visits, and that more reliable data may be obtained when families complete the measures in their own homes. A disadvantage of home completion however is that it is not a controlled situation, and children and parents may 'help' each other. For this reason, it might be expected that higher concordance be found where measures are completed at home and returned by post.

In attempting to clarify this issue, Glaser and co-workers[265] used the modified HUI with survivors of CNS tumours (aged 5–16 years) and compared parent–child ratings in home and clinic settings. Nineteen parent–child pairs completed the measures at home, and 28 patient–parent pairs completed the measure in the clinic. In both situations, there was good agreement between children and parents on all attributes except for vision and pain. Parent–child agreement was higher where questionnaires were completed at home than in the clinic, suggesting that some 'helping' takes place at home.

Mothers or fathers as proxy raters

As shown in *Table 6*, many studies include both mothers and fathers in their samples, although the numbers of fathers tend to be small. There have been no attempts to date to determine differences between mothers and fathers in terms of their abilities to act as accurate proxy for their child. Choice of proxy rater needs to be considered carefully, and differences between raters acknowledged. Given that mothers tend to be more involved in childcare, systematic differences favouring mothers might be anticipated.[15] The potential value of other proxies has not been considered, and may be particularly an issue when working with ethnic groups. In some cultures, grandparents may be more involved in childcare and consequently make reliable proxy informants. In other circumstances, particularly where children come from very deprived backgrounds, it may be difficult to identify any single 'primary' caregiver.

Implications for using proxy raters

Our conclusions need to be considered in the light of two issues. First, how good are the data on which the findings are based? Second, what degree of concordance might we expect to find between parent and child?

With regard to quality, much that has been written regarding child–proxy agreement has resulted from developmental work involving new measures. Thus, the question has been investigated from the standpoint of establishing concurrent validity for a new measure. As test developers tend to predict moderate agreement between parent and child, moderate agreement is often reported as evidence for the validity of a newly developed measure.[113,123,129,153]

With regard to the question of concordance, there are many reasons why parent and child may not agree. Children may wish to protect their parents and therefore hide their awareness of disease and its implications. They may want to present themselves in a positive light and therefore minimise the extent of their difficulties or they may simply be unaware of the potential restrictions associated with their disease. Proxy raters and patients do not share an identical 'data pool'. They differ in their interpretations of the questions, in their interpretations of the same event and do not share the same expectations.[253]

In addition, parents may have different perceptions of the consequences of disease; different from children and medical staff. They are likely to be more in tune with the child's emotional or behavioural functioning and more aware that the family will only be eligible for professional help if they report the child's limitations. Parents' views may be informed by the burden of care-giving, and their own mental health, well-being and concerns. Differences in agreement that might be attributable to demographic factors such as age or gender, and parent–child relationships have rarely been

addressed. To the extent that parents are more closely involved in the day-to-day lives of younger children, it might be expected that concordance would be greater than for older children or adolescents. Against this, the greater ability of the older child and adolescent to verbalise their experiences and feelings may mean that higher concordance is found for older groups. The current literature does not adequately address this question. Indeed, a different type of basic research may be needed.

The quality of the parent–child relationship may well be further modified by the presence of physical illness.[113,119,263] Where illness changes parent–child relationships, it may be necessary to conduct separate analyses for dyads involving healthy and chronically ill samples. In all currently available reports greater concordance for sick children has been attributed to the greater need for communication between parents and sick child as a consequence of need for treatment. However, such an explanation has never been seriously investigated. There is a need to determine more satisfactorily how far concordance is affected by the child's health. Parent–child concordance is likely to be affected by the extent to which parents are responsible for giving treatment, the child's willingness or compliance with treatment and whether or not a condition is perceived to be life-threatening.

It remains difficult to establish parent–child concordance in those situations where the need for a proxy is most urgent.[259] Children who are very young, and those with learning disabilities or special educational needs may all be excluded from studies unless special provision is made so that they are able to provide information about their own QoL. In these circumstances, proxy ratings may provide the only means of assessment of the child's QoL. From adult work, questions regarding the accuracy of proxy raters in different circumstances, particularly when patients are terminally ill, have been raised. For children too, the need for proxy raters in these circumstances is critical.

Discrepancies between child and parent ratings are generally regarded as a nuisance factor, challenging the validity of measures and questioning the value of assessing QoL at all. In fact, discrepancies may be useful clinically, in suggesting areas of family conflict. Discrepancies may also reflect different time perspectives. For example, although parents and children may agree that short stature is not a problem at the moment, differences between child and parent ratings may reflect the parents' greater concern about possible future problems. Such concerns may reflect parents' wider experience. Unfortunately there are currently no guidelines on how to resolve discrepant information. Where children report a QoL very different from that reported by their proxy, there is no way of knowing whether the views of child or proxy reflect the 'truer' picture.

As a consequence, there are consistent recommendations that until more conclusive evidence is obtained indicating that one informant is more reliable, information should be collected from multiple informants.[123,129] The parent report may provide a substitute for children's QoL at a group level, but large differences can exist in proxy agreement at the individual child–parent level.[129] There is a need to re-frame the question about proxy reporting away from 'Who is right?' to 'What does each rater contribute to the understanding of the child's QOL?'.[267]

Summary

- The question of adult–child concordance was considered in relation to three questions:
  - Is concordance greater for some domains than others?
  - Do parents over- or underestimate their child's QoL?
  - How is degree of concordance affected by child age, gender and illness status?
- We found limited evidence for the assumption of greater concordance between child and parent ratings for observable behaviours such as physical functioning, and less for non-observable functioning such as emotion or social QoL. Contrary data were reported in three studies.[129,154,165] There is also much greater heterogeneity in measures of social and emotional, compared with physical functioning and this may contribute to inconsistent results.
- The question of whether parents over- or underestimate their child's QoL was explored in three studies. These studies were based on different measures of QoL and involved different domains of QoL. No clear conclusions can be drawn given differences in measures and critically domains of QoL.
- There was also little literature considering how far concordance is affected by age, gender or illness status. Where it has been studied, concordance is greater for parent–child agreement for groups of chronically sick compared with healthy children.
- All of these results may be influenced by the

specific measure of QoL employed. Differences in quality of the individual measures may contribute to the findings.

- Given the status of work in this area, it is important to obtain information from parents and children whenever possible.

TABLE 6 *Concordance between child and parent ratings on generic measures*

Study and questionnaire used	N	Proxy	Age of child	Disease group(s)	Domains	Inter-rater reliability	Comments
How Are You? (HAY) Bruil, 1999[119]	995	Chronic illness: Mothers 88% Fathers 9% Both 5% Healthy: Mothers 84% Fathers 15% Both 1%	7–13	Chronic illness *n* = 577 Healthy children *n* = 418	Quality of performance: Physical activities Cognitive tasks Social activities Prevalence: Physical complaints Physical activities Social activities Cognitive tasks Happiness	 0.65;[a] 0.30[b] 0.51; 0.35 0.44; 0.27 0.61; 0.45 0.60; 0.59 0.49; 0.52 0.37; 0.37 0.36; 0.32	ICCs also computed which were comparable with Pearson's correlations
Quality of Well-Being (QWB) Czyzewski *et al.* [165]	199	Mothers 92% Fathers 8%	0.2–17.9 years	CF	Symptom complex Functionality Overall QWB	0.23 0.55 0.39	Pearson product–moment correlations
Perceived Illness Experience (PIE) Eiser *et al.*, 1995[126]	35	Parents (no. of mothers, fathers or other not specified)	8–20	Cancer	Physical symptoms Functional disability Restrictions Parent behaviour Peer rejection Psychological symptoms Physical appearance Interference-activity Manipulation Preoccupation-illness School Treatment Disclosure Total	0.84 0.81 0.62 0.60 0.59 0.56 0.56 0.54 0.54 0.52 0.51 0.39 0.13 (n.s) 0.35	*t*-tests showed that there were no differences in mean ratings between adolescents and their parents for any of the domains

continued

TABLE 6 contd *Concordance between child and parent ratings on generic measures*

Study and questionnaire used	N	Proxy	Age of child	Disease group(s)	Domains	Inter-rater reliability	Comments
Perceived Illness Experience (PIE) Eiser *et al.*, 1999[125]	31	Mothers only	8–25	Treated by limb salvage procedures for a primary bone tumour	Food Parental behaviour Disclosure Appearance School/work Peer rejection Manipulation Preoccupation Interference with activity Total	0.69 0.66 0.61 0.57 0.54 0.53 0.48 0.43 0.43 0.59	*t*-tests revealed that there were no significant differences between the means for child and parent report
Glaser *et al.*, 1997[265]	27	Parents (*n* = 21) Teachers (*n* = 27)	6–17 years	CNS tumours = 27 Siblings = 21 Age- and sex-matched controls	Cognition Hearing Vision Pain Dexterity Ambulation Speech Emotion Confidence for the future Self-esteem	1.0 (kappa) 1.0 0.59 0.66 0.68 0.65 0.24 0.26 1.0 1.0	Good agreement between parent and child for all domains Patients perceived ambulation and speech worse than teachers
HUI Mark III Glaser *et al.*, 1997[264]	19 (at home) 28 (in clinic)	Parents (no. of mothers, fathers, or other not specified) Teachers	5–16 years	CNS tumours	Speech, hearing, dexterity, ambulation, emotion, vision, pain, cognition, self-esteem, confidence for future		Good agreement for all attributes in both settings
Children's Quality of Life Graham *et al.*, 1997[113]	25 26	Mothers	9–15 9–15	Psychiatric disorder (PD) Chronic physical disorder (CPD)	Perception of level of function (CPD: 0.57; PD 0.57)		ICCs

continued

TABLE 6 contd *Concordance between child and parent ratings on generic measures*

Study and questionnaire used	N	Proxy	Age of child	Disease group(s)	Domains	Inter-rater reliability	Comments
TACQOL Theunissen *et al.*, 1998[129]	1105	Mothers 84% Fathers 11% Both 5%	8–11	School-based sample	QoL: Cognitive functioning Physical complaints Social functioning Motor functioning Autonomy Health status: Cognitive functioning Physical complaints Autonomy Motor functioning Social functioning QoL/health status: Negative emotions Positive emotions	 0.61 0.54 0.51 0.48 0.47 0.62 0.56 0.53 0.52 0.48 0.55 0.39	ICCs also computed which were comparable with Pearson's *r* Forward stepwise, linear regression was conducted for each scale to reveal relationships between absolute proxy agreement and background variables
PedsQL Varni *et al.*, (submitted)[124]	281	Mothers 87%	8–18	Cancer	Physical Pain Nausea Procedural anxiety Treatment anxiety Cognitive Psychological Worry Social Appearance Communication with physician/nurse Core (physical, social, psychological)	0.57 0.55 0.51 0.48 0.45 0.42 0.38 0.35 0.30 0.29 0.13 0.53	ICCs reported

continued

TABLE 6 contd *Concordance between child and parent ratings on generic measures*

Study and questionnaire used	N	Proxy	Age of child	Disease group(s)	Domains	Inter-rater reliability	Comments
Battery of measures Walker & Heflinger, 1998[263]	257	Mothers Teachers	6–18	Recurrent abdominal pain *n* = 90 Organic abdominal pain *n* = 63 Well *n* = 56 Psychiatric *n* = 48	Children's somatisation Abdominal pain Functional disability Academic competence Children's depression	0.47; 0.15 0.46; 0.22 0.42; 0.12 0.38; 0.44 0.30; 0.22	Correlations shown for the recurrent abdominal pain and well groups only

[a]*Children with chronic illness*
[b]*Healthy children*
n.s. not significant

TABLE 7 *Concordance between child and parent ratings on disease-specific measures*

Study	N (child–parent pairs)	Proxy	Age of child	Disease group(s)	Domains	Inter-rater reliability	Comments
Quality of Life in Headaches –Youth report (QLHY) Langeveld *et al.*, 1997[154]	64	Parent (no. of mothers or fathers not specified)	12–18	Headache and non-headache	Stress Impact on activities Social interaction (sibling) Social interaction (peers) Depression VAS health satisfaction Somatic symptoms (other) Home and school VAS life satisfaction Leisure activities Harmony Fatigue Strength/vitality Cheerful mood Optimism about future	0.48[a] 0.47 0.47 0.44 0.43 0.40 0.39 0.36 0.33 0.31 0.31 0.30 0.28 0.27 –0.13 (n.s)	ICCs reported

continued

TABLE 7 contd *Concordance between child and parent ratings on disease-specific measures*

Study	N (child–parent pairs)	Proxy	Age of child	Disease group(s)	Domains	Inter-rater reliability	Comments
Behavioural Affective Somatic Experiences Scale (BASES) Phipps *et al.*, 1999[142]	42	Parent (no. of mothers or fathers not specified)	5–20	Cancer patients undergoing bone marrow transplantation	Somatic distress Quality of interactions Activity Mood disturbance Compliance	0.57 0.49 0.43 0.35 0.29	Correlations were also computed between parent–nurse and nurse–child
Pediatric Cancer Quality of Life Questionnaire (PCQL) Varni *et al.*, 1998[143]	157	Parent (no. of mothers and fathers not specified)	8–18	Cancer	Physical functioning Symptoms Psychological Social Cognitive	0.43 0.42 0.34 0.27 0.26	Adolescent–parent concordance using data standardised by Simpson's paradox yielded results essentially identical to those based on the original data
Paediatric Cancer Quality of Life-32 items Varni *et al.*, 1998[144]	274	Parent (no. of mothers or fathers not specified)	8–18	Cancer	Physical Cognitive Symptoms Psychological Social	0.59 0.46 0.45 0.41 0.36	

[a] *Only ICC at first administration reported here. ICC also conducted at week 1, 2 and 3*
VAS, visual analogue scale

Chapter 7

How feasible and reliable are proxy measures of QoL in different disease contexts?

Aims and scope

In this chapter we consider the reasons for adopting a battery approach to measurement. The battery approach refers to the use of measures related to, but not specifically developed for the measurement of QoL (e.g. physical functioning or depression). Based on our review, we identified five studies where this approach was adopted. In evaluating these papers, we conclude that there are considerable methodological problems with this approach, including the length and consequent response burden for the child.

Background

As described in previous chapters, there are a limited number of psychometrically valid measures currently available for evaluation of QoL in children. As a consequence, measures of concepts related to different domains of QoL have sometimes been used as proxies. Indeed, it is possible to identify measures that could be used to assess physical function (e.g. physical symptoms and pain), psychological function (e.g. anxiety, depression or body image), and social function (e.g. employment, social relations or marital status). Thus, a combination of different, but previously validated measures may be considered preferable to a single measure of QoL, particularly where few psycho-metric data are available. This **battery** or **modular** approach refers to collections of specific measures that are scored independently and reported as individual scores.[82]

Arguments in favour of a battery of measures to assess QoL focus on three issues. First, as has been shown in previous chapters, there were relatively few reliable and valid measures available to assess QoL, at least until very recently. Second, generic QoL measures may lack sensitivity in the domain of greatest interest. For example, if the task is to assess the impact of limb amputation on the child's QoL, then the limited number of questions included to assess physical functioning in generic measures is clearly inadequate. A battery approach would allow for the addition of other questions or measures that would provide a more comprehensive assessment. Third, with few exceptions[117,268] measures are not suitable for self-report by children under 11 years of age. Researchers may therefore decide it is preferable to use measures of proven value for self-report in young children (e.g. self-esteem: Harter and Pike[247]).

Results

Fifteen studies were identified in which a battery approach to the measurement of QoL was adopted. Although these studies had passed the inclusion criteria identified in chapter 3, further examination suggested that additional criteria were needed. In particular, these studies were characterised by small sample sizes and unacceptable amounts of missing data. In order to maintain some degree of quality control over the contents of the review, we therefore decided to define additional exclusion criteria for these studies. As a result, studies were excluded on the basis of limited sample sizes (< 30), which in most cases was combined with missing data, and a restricted number of proxy measures. In this latter case, we excluded studies where the range of proxy measures included was not such as to provide a multidimensional measure of QoL (focusing perhaps on physical to the exclusion of social QoL).

Of the fifteen studies identified, ten failed to fulfil these inclusion criteria. The five studies included are summarised in *Table 8* (see page 74), and the excluded studies are summarised in *Table 9* (see page 77). It is important to note that three of the papers[164,269,263] were book chapters resulting from papers presented at a conference concerned with assessment of QoL in children and adolescents (October 1996, Case Western Reserve University in Cleveland, Ohio).

Aims

The main purpose of four papers was limited to describing QoL in a group of children with chronic illness.[263,269,270,275] The remaining paper reported an evaluation of a health promotion intervention and compared the efficacy of a battery approach with a single QoL measure.[164.]

Study design

Four studies were described as longitudinal. However, only the baseline assessments were described by Austin and co-workers,[270] and the necessary comparisons between baseline and follow-up data were not described in their later report.[275] Longitudinal data were reported by Walker and Heflinger[263] who made follow-up phone calls 5 years after initial entry into the study. Czyzewski and co-workers[164] evaluated patients and families at baseline and follow-up at 18 and 30 months later.

Selection of tests

All researchers broadly justified their choice of measures as reflecting the multidimensional nature of QoL. Thus, they included measures of physical, social and emotional well-being. Beyond this, there was little attempt to provide information about choice of specific proxy measures (i.e. why one measure of depression rather than another).

In the five papers included, 16 different standardised measures were employed in the proxy assessment of QoL (*Table 10*). Gortmaker and co-workers[269] combined items from previously reported measures to represent six domains (health perceptions, physical resilience, physical functioning, psychological functioning, social and role functioning and symptoms). Two papers reported using interviews or questionnaires that were developed specifically for the study to obtain general information relevant to QoL. These concerned issues such as number of school days missed or days in bed.[269,270]

TABLE 10 *Proxy measures employed in five studies that adopted a battery approch to measurement of QoL*

Measure	Study
CBCL	Austin *et al.*, 1994,[270] 1996[275] Czyzewski *et al.*, 1998[164] Gortmaker *et al.*, 1998[269]
Self Concept	Austin *et al.*, 1994,[270] 1996[275]
Self Perception Profile	Walker & Heflinger, 1998[263]
Health Resources Inventory	Walker & Heflinger, 1998[263]
Depression	Walker & Heflinger, 1998[263]
Functional Disability Index	Walker & Heflinger, 1998[263]
Abdominal Pain Index	Walker & Heflinger, 1998[263]
Children's Somatisation Inventory	Walker & Heflinger, 1998[263]
Aids Controlled Trials Group QoL form	Gortmaker *et al.*, 1998[269]
National Health Interview Survey	Gortmaker *et al.*, 1998[269]
Behaviour Problems Index	Gortmaker *et al.*, 1998[269]
Parenting Stress Index	Czyzewski *et al.*, 1998[164]
Impact on Family	Czyzewski *et al.*, 1998[164]
Vineland Adaptive Behaviour Scales	Czyzewski *et al.*, 1998[164]
Family AGPAR	Austin *et al.*, 1994,[270] 1996[275]
Children's attitude to illness scale	Austin *et al.*, 1994,[270] 1996[275]
NIH	Czyzewski *et al.*, 1998[164]

Length of the battery

Information concerning the number of items included in the battery, or the time taken for administration and scoring was rarely reported. An exception was Gortmaker and co-workers,[269] who reported that the battery of measures in their study took approximately 20 minutes for parents to complete.

Appropriate use of tests

There was a tendency for researchers to select subscales of measures rather than use complete versions, thus potentially compromising established psychometric criteria. Choices were justified on the grounds of relevance to the particular domains under consideration. Thus, Walker and Heflinger[263] assessed perceived academic and social competence with subscales of the Self Perception Profile for Children[247] or Adolescents[299] Austin and co-workers[270, 275] used subscales from the Child Behaviour Checklist (CBCL),[300] which they considered most clearly reflected the psychological, social and school domains of QoL.

Response rates

In their baseline study, Austin and co-workers[270] reported that only one family declined to participate because of the amount of time involved. In their follow-up, Austin and co-workers[275] reported that seven families (out of 270) chose not to participate, 14 had incomplete data, 18 were lost to follow-up and three were dropped for other reasons. The remaining studies did not report response rates.

Age of children

Given the number of questionnaires that make up a battery, and consequently the response burden imposed, there may be limits on how far children can assess their own QoL. Gortmaker and co-workers[269] reported that only children above 12 years of age were able to provide data directly. In the studies by Austin and co-workers[270,275] children were aged between 8 and 12 years. As it was perceived that children in this age range may not be able to complete all measures in the battery, some domains were assessed by mothers only, and for other domains different tests were used for child or mother completion. Walker and Heflinger[263] included children aged between 6 and 18 years, presumably using interviewers to read the questions to the younger children. In

the study by Czyzewski and co-workers[164] data were collected only from parents and adolescents, but not younger children (the sample included patients aged between 1 and 18 years).

Psychometric properties of battery measures

Gortmaker and co-workers[269] reported the reliability and validity of the QoL battery employed. Reliability (internal consistency) for the six scales was based on coefficient alpha. Construct validity was determined by comparing scale differences between children (all of whom were infected with HIV) who were either diagnosed as having AIDS or not.

Czyzewski and co-workers[164] examined the inter-relationships between measures both within domains, in order to understand the contributions of multiple measures of the same domain, and between the QWB and specific domain measures. Small-to-moderate correlations were reported within domains (e.g. 'impact related to illness' with 'parental stress'), and the relationship between the QWB and other QoL measures depended on the informant.

Austin and co-workers[270] attempted to determine the contributions of the different measures to domain specific QoL by examining intercorrelations among the 15 QoL variables, which were grouped into four domains (physical, psychological/behavioural, social and school). A similar procedure was adopted in their follow-up study. Austin and co-workers[275] grouped 19 QoL variables into three domains (psychological, social and school). Clinical validity was determined by comparing children with asthma and epilepsy. In addition Austin and co-workers (1996)[275] examined validity on the basis of disease severity (active or inactive). As predicted, adolescents with active epilepsy reported the poorest QoL.

Walker and Heflinger (1998)[263] made limited attempts to examine the reliability of their battery by examining the degree of concordance between child and parent report (see chapter 6). They also examined discriminant validity by comparing differences between the diagnostic groups with respect to various domains of QoL and predictive validity by identifying measures that significantly predicted recovery at follow-up.

Determination of an overall QoL score

Only one study attempted to integrate the separate findings to produce a general index of QoL. Gortmaker and co-workers[269] conducted a factor analysis on the separate domains of QoL to determine how far the separate measures in the battery could be combined to yield an overall score. For those diagnosed with AIDS, two factors were identified. The first (50% of the variance) included all domains except behaviour problems, which loaded on the second factor (17%). For those who were infected but not diagnosed with AIDS, the first factor included physical functioning, physical resilience and symptoms (39%), and the second included psychosocial functioning and health perceptions. This raises questions about the appropriateness of assuming that QoL can be measured for children with different health problems using the same items. In the remaining studies, results were reported for the separate domains and no attempt was made to combine them to provide an overall QoL index.

How are results used to determine QoL?

In two studies, conclusions about QoL were made on the basis of comparisons with normative data.[164,269] In the remaining studies, QoL of the target group was assessed by comparison with other sick groups.[263,270,275]

Discussion

Appraisal of the reliability of proxy measures of QoL is made difficult by:

- the limited number of studies that adopt this approach and conform with basic psychometric criteria ($n = 5$)
- the lack of information regarding reasons for employing proxy measures of QoL
- failure to justify the selection of measures
- lack of information concerning the practicality of the battery (e.g. time taken to administer, total number of items)
- failure to report psychometric properties of the battery.

There are a number of potential problems with a battery approach.

- In the absence of any agreement concerning suitable domains of child QoL, each investigator is likely to select those of most interest. Thus, any battery does not necessarily include all the domains typically identified in definitions of QoL. This makes comparisons of QoL in different diseases particularly difficult.[198] The inclusion of measures in a battery appears to be *ad hoc*, with the rationale for the inclusion of particular measures rarely made explicit. It is

impossible to be clear simply from the titles of instruments or their constituent scales and dimensions what precisely is being measured.[69]
- Analysis of changes over time is particularly difficult when multiple measures are used.
- Considerations of statistical power have not been well addressed in any of the studies. In addition, the precision with which key parameters (e.g. reliability) have been assessed varies significantly. Authors have generally failed to consider psychometric properties of the batteries they employ.
- Different measures are used for different age groups. This may not be conducive to long-term follow-up and can result in unacceptable levels of missing data.
- A potential danger with this approach is that the multiple comparisons that can be made lead to erroneous conclusions, to the extent that simply by chance one or two of the comparisons will yield significant results. Multivariate analyses are necessary to guard against such errors.
- Even if the individual measures employed have established construct validity, the combination of proxy measures within a domain may not be appropriate. The scales used may not reflect QoL, but may reflect other constructs. This is demonstrated in the study by Gortmaker and co-workers[269] where the results of a factor analysis suggest that the battery reflects at least two domains, rather than a single construct of QoL.

Choice of measures is discussed more fully by Gortmaker and co-workers[269] who argue that it may be legitimate to drop specific domains for some studies where there is no expectation that QoL might be effected. Thus, it may not be necessary to include a measure reflecting the domain of physical QoL if there is no reason to expect that physical functioning is affected. This would allow QoL measurement to be tailored for individual studies and keep the battery within manageable time limits. (In practice, situations where it is possible to be so confident are difficult to identify).

Thus, decisions have to be made concerning whether assessment of QoL should be comprehensive, with inclusion of measures that cover a broad range of QoL domains, versus inclusion of measures that tap into only those domains that are relevant to the treatments or disease under study. This would mean restricting assessment of children with recurrent pain to a measure such as the Abdominal Pain Index,[284] rather than including additional measures that might tap a wider range of issues affecting the child. However, it is important to balance the need for comprehensiveness with the burden imposed on the child when asked to complete a lengthy battery. The use of a general measurement strategy may be particularly important when the effects of the treatment or disease on QoL are largely unknown. However, when QoL components that are not linked to a specific treatment or disease are included in the battery, they may simply add unnecessary complexity to the design and analysis. It may therefore be preferable to use measures with clearly stated hypotheses concerning the relationships between the measures selected and the treatments or disease processes under study.[301]

Comparison of the efficacy of the battery approach over single measures of QoL was attempted in one study. Czyzewski and co-workers[164] discuss the utility of a multidimensional battery versus a single measure to assess QoL in children and adolescents with CF in the context of a family health education intervention. They used a battery of measures, which they considered reflected the key domains of QoL (health, psychological, social, and functional status). Results were compared with the QWB Scale [279] as a single measure.

Apart from the advantages of cost, the use of established measures in a battery may enable comparisons to be made with population norms. As such, this approach may be seen to have merits over any currently available QoL measure, as there are few such measures with population norms available. In using norms in this way, it must be recognised that information about subscales of measures are unlikely to be reliable (see chapter 2). Any changes to the measures used may effect the comparability of the data with norms and comparisons across studies. Proxy measures that tap into various domains of QoL may provide more detailed information concerning the impact of a disease on the patient's QoL, and may provide information that may not be obtained from a single QoL measure.

Any battery of tests is only as good as the measures included. There is some irony (and indeed concern) in the fact that measurement of QoL grew out of a general dissatisfaction with the alternatives frequently used for assessment of sick children, and then these same measures may be used in batteries. The CBCL[300] is a case in point. This measure has been criticised as inappropriate for work with sick children on a number of grounds, including the fact that use of items concerning somatic symptoms may result in overestimation of behaviour problems for sick

children.[302] Consequently, proxy measures used to assess children with chronic illness should be carefully reviewed to determine if any items are confounded by illness factors. As a consequence, standardised measures may need to be modified (either by omitting certain items or adjusting scores) to account for the inclusion of any items that would artificially inflate the scores of sick children.

Summary

- The battery approach has not been frequently used with children. We found that studies using this approach fell short of our original inclusion criteria. In particular, the inevitably lengthy test batteries result in unacceptably high levels of missing data.
- In the five studies that satisfied the inclusion criteria, there was nevertheless considerable variability in selection of measures.
- Evaluation is difficult because authors fail to describe how they chose measures for their batteries, and do not routinely report critical information such as completion rates or missing data.
- Decisions about use of a battery to assess QoL are complex. On the one hand, use of existing measures can potentially eliminate the time and expense required to develop a single measure of QoL. On the other hand, a full battery of standardised tests may be expensive in terms of the time needed to administer and score. In addition, battery measures tend to be lengthy to complete and therefore demanding for sick children.
- Decisions have to be made concerning the inclusion of measures that cover a broad range of domains that are relevant to the treatments or disease under study. The use of a general measurement strategy may be more desirable when the effects of the treatment or disease on QoL are largely unknown. However, when QoL components that are not linked to a specific treatment or disease are included in the battery, they may simply add unnecessary complexity to the design and analysis.
- It may therefore be preferable to use measures with clearly stated hypotheses concerning the relationships between the measures selected and the treatments or disease processes under study.[301] Gortmaker and co-workers[269] argue that a modular approach allows QoL measurement to be tailored for individual studies so that domains may be dropped for some studies where there is no expectation of impact. At the same time, this approach allows for the addition of more sensitive measures concerned with particular aspects of QoL that are expected to be most affected by the condition or treatment of interest.
- The battery approach may be particularly inappropriate for self-completion by younger children. Use of measures developed by different authors for different purposes may impose some burden on children given that measures may differ in response format. The battery approach is also likely to result in longer measures and greater response burden for children.

TABLE 8 *Summary of studies using a battery approach to the assessment of QoL included in the review*

Study	Aim(s)	Sample	Domains measured (as specified by authors)	Respondent and instrument (subscales in parentheses)	Comments
Austin *et al.*, 1994[270]	Longitudinal evaluation of the natural history of adaptation to asthma and epilepsy	270 children aged 8–12 years with epilepsy or asthma	Physical:	Mother: Interview (school absences etc.)	Results reported from the first data collection of a longitudinal study
			Psychological/behaviour:	Child: Self Concept[271] (anxiety, happiness/satisfaction)	Different respondents for different domains
				Attitude to illness[272]	Intercorrelations reported among each of the QoL variables
				Mother: CBCL (internalising, externalising, social, school, peers) [273]	Reliability and validity of the measures not reported for the samples included
			Social:	Child: Family AGPAR[274]	Comparisons were made with norms for the CBCL and CBCL-TRF
				Mother: CBCL (peer relations, social activity, social problems)[273]	Total of 15 QoL variables
			School:	Mother: CBCL (school progress)[273]	No attempt to create overall score
				Teacher: CBCL-TRF (school achievement, internalising, externalising behaviour)[273]	Authors note that physical domain was not measured as comprehensively as the other domains
Austin *et al.*, 1996[275]	To compare QoL in inactive or active epilepsy with those with asthma Describe differences in QoL related to sex and illness severity	Longitudinal follow-up (see Austin *et al.*, 1994) 228 children 12 to 16 years with epilepsy or asthma	Psychological:	Child: Self concept (self anxiety, self happiness)[271]	Results reported from the second data collection of a longitudinal study
				Attitude to illness[272]	No comparisons were made with the data obtained at Time 1 (Austin *et al.*, 1994)
				Mother: CBCL (internalising, externalising problems, thought problems, attention problems)[273]	Different respondents for different domains
			Social:	Child: Family AGPAR[274]	Total of 19 QoL variables assessed
				Mother: CBCL (peer relations, activities, social problems)[273]	No attempt to create overall score

continued

TABLE 8 contd *Summary of studies using a battery approach to the assessment of QoL included in the review*

Study	Aim(s)	Sample	Domains measured (as specified by authors)	Respondent and instrument (subscales in parentheses)	Comments
				Teachers: CBCL (social problems)[273]	
			School:	Child: Self concept (intellectual)[271] Mother: CBCL (school progress)[273] Teacher: CBCL-TRF (school achievement, internalising, externalising, thought problems, attention problems)[273]	
Czyzewski *et al.*, 1998[164]	To compare a battery and a single metric approach; assess the differential impact of CF on a range of domains; evaluate intervention programme	Children with CF < 1 to 18 years Caregiver dyads (n = 199) Dyads evaluated at baseline and 18 to 30 months after intervention	Health status:	Clinician: NIH[276]	QoL inferred from comparison with population norms
			Psychosocial:	Parent: Vineland Adaptive Behaviour[277] Parent/Youth self-report: CBCL Parent: Parenting Stress Index[278] Impact on Family[241]	No attempt to create overall QoL index Concluded that the QWB is less sensitive than the battery approach in detecting current QoL or expected age-related difference in QoL
			Family functioning:	Parent and child: QWB[279]	
			Single metric:		
Gortmaker *et al.*, 1998[269]	To test the General Health Assessment for Children (GHAC); a modular measure of QoL to be used in AIDS clinical trials	444 children aged 5 to 11 with HIV-1	Health perceptions:	Health ratings adapted from ACTG QoL form for adults[280]	The GHAC demonstrated very good internal consistency
			Physical resilience:	7 questions from National Health Interview Survey[281]	Significant difference between children with and without AIDS

continued

TABLE 8 contd *Summary of studies using a battery approach to the assessment of QoL included in the review*

Study	Aim(s)	Sample	Domains measured (as specified by authors)	Respondent and Instrument (subscales in parentheses)	Comments
	To compare children with and without AIDS GHAC forms were designed for parents of children aged 6 months to 4 years, and 5 to 11 years. An existing adult report QoL questionnaire was utilised for 12 to 20 year-olds and parents also filled out this form		Physical function: Psychological function: Social and role function: HIV related symptoms:	6 items adapted from ACTG QoL[280] Behaviour Problems Index[282] 5 items from National Health Interview Survey[281] ACTG QoL[280]	Original 6 scales subject to factor analysis to create 2 factors. Factor 1 included physical functioning, physical resiliance, HIV-related symptoms) and factor 2 included psychosocial functioning and health perceptions Employed different assessment instruments for different age groups QoL inferred from comparison with normative data
Walker & Heflinger, 1998[263]	To compare different groups To predict QoL outcomes 5 years after clinic visit To compare child and parent ratings	Longitudinal 257 children aged 6–18 years Abdominal pain: 90 with recurrent abdominal pain; 63 with an organic diagnosis; 48 psychiatric conditions; 56 healthy	Physical: Functional: Emotional: Social:	Children's Somatisation Inventory[283] Abdominal Pain Index[284] Functional Disability Index[285] Children's Depression Inventory (CDI)[286,287] Self Perception[288] Health Resources Inventory (academic competence & peer relations)[289]	Child report (not parent report) was a significant predictor of recovery No attempt to create overall index Follow-up interview was conducted with abdominal pain patients over the telephone 5 years after the initial interview

CBCL-TRF. Child Behaviour Checklist – Teacher Report Form
ÑIH, National Institutes of Health
ACTG, AIDS Controlled Trial Group

TABLE 9 *Studies using a battery approach to the assessment of QoL that failed to reach the inclusion critera*

Study	Design	Domains	Instruments	Results	Comments
Chin *et al.*, 1991[290]	Sample: 26 children (age 0.58–14.2) undergoing orthotopic (OLT) for end-stage liver disease Aim: to investigate outcome and to evaluate areas of potential ongoing concern after OLT. QoL was assessed in those with a minimal follow-up period of 12 months	Socialisation, daily living skills, communication, motor skills	Vineland Adaptive Behaviour Scales liver transplantation	Comparisons were made with the normal range In all but 3 children QoL as perceived by parents was good or excellent Socialisation scored relatively highly while gross motor and community daily living scales performed less well	Small sample Single instrument used to assess QoL
Cleary *et al.*, 1994[291]	Sample: 19 patients aged 4.5–18 years who were participants in a Phase II trial of the value of rG-CSF in children with congenital agranulocytosis Aim: questionnaires were administered on 8 occasions during a clinical trial, to evaluate patient QoL	Functional status: General health perceptions: Activity limitation: Disease-related symptoms: Discomfort related to therapy:	FSQ Rand HIS Rand HIS Author developed	Treatment with rG-CSF results in significant improvement in general health perceptions, limitations of daily activities, and symptoms of the disease Only minor discomfort due to therapy was noted	Small sample Missing data (not all patients completed all the questionnaires) Psychometric properties not examined
Felder-Puig *et al.*, 1998[292]	Sample: 60 patients aged 15 to 30 years. One year since treatment for bone cancer Aim: Evaluation of psycho-social QoL	Psychosocial	State Trait Anxiety Inventory Frankfurt Self-Concept Scales Questionnaire on Life Goals and Satisfaction with Life	Approximately 80% of patients revealed minor psychosocial problems Neither clinical data nor physical or functional sequelae affected psychosocial adjustment	Assessment of psychosocial QoL 54% participation rate

continued

TABLE 9 contd *Studies using a battery approach to the assessment of QoL that failed to reach the inclusion critera*

Study	Design	Domains	Instruments	Results	Comments
Herndon *et al.*, 1986[9]	Sample: 12 survivors with > 80% total body burns. Age range 8 months to 12.4 years at time of burn Aim: examination of QoL after major thermal injury in children	Physical, psychological	Standard mental status evaluation Louisville Behaviour Checklist The Burn Injury Questionnaire The Activity Questionnaire	Physical impairment was approximately 60%. 50% of children were completely independent in activities of daily living One-third of the children had excessive fear, regression, neurotic and somatic complaints	Small sample
Johnson *et al.*, 1994[293]	Sample: 13 survivors of medulloblastoma age 1.7–15.9 years Aim: to examine the quality of long-term survival	Intellectual functioning/ school: Behavioural/ emotional	WISCC-R WISC-R Wide range Achievement test-R Controlled Oral Word Association Test Hooper Visual Organisation Test Purdue Pegboard Wisconsin card sorting test Benton Visual Retention Test Wechsler Memory Scale Rey Auditory Verbal Learning Test CBCL Symptom Checklist: self-report Vineland Adaptive Behavior Scales Harter Self Perception Profiles	Perceptual motor task performance was below average in more than half of the participants, but motor dexterity was more severely affected than perception Problems in learning and a delay in physical growth and development were seen in the majority of patients	Small sample 57% of the patients completed all components of the study

continued

TABLE 9 contd *Studies using a battery approach to the assessment of QoL that failed to reach the inclusion critera*

Study	Design	Domains	Instruments	Results	Comments
Kun *et al.*, 1983[294]	Sample: 30 children with primary brain tumours aged 1 year 9 months to 15 years 8 months at diagnosis Aim: to examine intellectual, emotional and academic function	Intellectual: Emotional: Academic:	Bloom's functional status McCarthy Scales of Children's Abilities WISC-R Peabody Individual Achievement Test Personality Inventory for Children Mathematics Reading recognition Reading comprehension Spelling	Subnormal IQ levels were found in 8 children who had cranial irradiation Social and emotional problems were encountered in a large proportionof patients Statistically, only the scale of psychotic tendencies was abnormally represented	Small sample Not all the children complete the assessment tests
MacPhee *et al.*, 1998[295]	Sample: 30 children were aged 12–18 years with Crohn's disease or ulcerative colitis Aim: examination of psycho-social factors that influence adolescent and family adjustment and QoL in newly diagnosed (< 5 years) inflammatory bowel disease patients	Psychosocial:	Social Support Network Questionnaire QOL for adolescents and parents (Garret & Drossman) Family Inventoy of Life Events and Changes for Adolescents (A-FILE) and Parents (FILE) Coping Health Inventory for Parents Adolescent Coping Orientation for Problem Experiences	Adolescents' QoL scores significantly correlated with satisfaction and degree of closeness with their social support network Severity of illness did not correlate with adolescents' QoL scores Mothers also completed the same measures to assess their own QoL etc.	Small sample Variable duration and severity of disease Selection bias (only dyads included)

continued

TABLE 9 contd *Studies using a battery approach to the assessment of QoL that failed to reach the inclusion critera*

Study	Design	Domains	Instruments	Results	Comments
Martinez-Climent *et al.*, 1994[296]	Sample: 39 children 4 to 23 years (mean 14 years) who were survivors of cranial posterior tumours Aim: to develop a scale for assessing QoL in survivors of posterior fossa tumours	Psychological: Endocrine/ growth status: Functional status:	WISC/WISC-R Terman-Merril test (children < 4) Growth and endocrine status Bloom's scale	QoL scores were high in 49% (good QoL) of patients, intermediate in 20%, and low (poor QoL) in 31%	Small sample
Nespoli *et al.*, 1995[297]	Sample: 36 children aged 3 to 19 years who had undergone bone marrow transplants (BMT) Aim: evaluation of QoL in children who underwent BMT in childhood	Anxiety: Depression: Self image:	Busnelli Anxiety Scale (BAS; children 8–15 years) CDI (children 9–16 years) OSIQ (adolescents 13–19years)	Average scores for the BAS did not fall outside the normal range reported for a sex and age-matched Italian population On the CDI shows a slight inadequacy on Fac 1 but was in the normal range for Fac 2 OSIQ scores were in the normal range except for lower scored on the superior adjustment scale for BMT girls	Small sample Not all children completed each of the measures
Packer *et al.*, 1987[298]	Sample: 24 long-term survivors of medullo-blastoma of the posterior fossa Aim: to characterise the QoL of children surviving medulloblastoma and the factors affecting outcome	Domains were not clear Performance level: Intelligence: School performance and other measures of outcome:	WISC-R Wide Range Achieve-ment Test Vineland Social Maturity Scale Finger Tapping and Grooved Pegboard Beery Developmental Test of Visual Motor Integration Target Test Benton Visual Retention Test Sentence Memory Rey Auditory Verbal Learning Test Selective Reminding Memory Procedure Verbal Fluency	12% of patients had an Full Scale Intelligence Quotient below the normal range Specific learning, memory and fine-motor disabilities were found in over 50% of patients	Not all children completed each of the measures

FSQ, Functional Status Questionnaire; Rand HIS, Rand Health Insurance Study; WISC. Wechsler Intelligence Scale for Children

Chapter 8

The use of QoL measures in different contexts

Aims and scope

So far our review has drawn heavily on issues involved in measuring QoL, and attempting to answer the specific questions identified by the HTA. The focus on measurement is important, but the resulting debates about quality threaten to undermine the use and integration of QoL measurement in both clinical practice and research.

In justifying the need for this review in chapter 1, it was argued that QoL measures were potentially useful in a number of different contexts. In this chapter, we attempt to review how far QoL measures have been used to measure outcomes. The review is based on the searches conducted to answer the four main questions of the review, and therefore does not necessarily include all the available and relevant review. The findings nevertheless point to the limited use of QoL measures in research. In trying to account for these findings, we consider a number of barriers or misconceptions about the measurement of QoL in research and clinical practice.

QoL measurement in practice

Comparing outcomes in clinical trials

QoL measures were originally considered in terms of their potential value for assessing outcomes in clinical trials involving patients with life-threatening conditions.[303,304] More recently, their value has also been recognised in a greater range of situations, including evaluations involving non-life-threatening conditions.

QoL measures might be useful to compare different chemotherapy regimens such as continuous or intermittent therapy, or alternative methods of delivering nebulised budesonide in the treatment of steroid-dependent children with asthma.[305] "Traditional end-points to measure effectiveness, such as disease-free survival and survival, need to be supplemented by outcomes reflecting the impact of treatment on the patient, and in some instances, families".[305]

Bradlyn and co-workers[303] conducted a national survey to determine how far QoL measures were employed in Phase III clinical trials in paediatric oncology. Research publications were identified in bibliographies from two major US Oncology Groups (Pediatric Oncology Group and Children's Cancer Group). They did not restrict their work to the use of comprehensive QoL measures but defined QoL very broadly to include any assessment of individual domains such as mobility, or relations with parents. Based on this broad definition, QoL measures were only included in 3% of paediatric oncology trials. (Comparable figures suggest that some 8.2% of cancer trials for adults include QoL measures.[306]) Any evaluation based strictly on the use of standardised QoL measures would suggest an even lower incidence. In adult work, despite the increasing availability of QoL measures, researchers continue to rely on *ad hoc* questions in evaluations of clinical trials.[307] In paediatric work, QoL is often discussed, but not yet routinely integrated in trial evaluation.

Some tangential evidence that QoL measures could be used to compare outcomes in chemotherapy trials was provided by Barr and co-workers.[308] They used the HUI (marks 2 and 3) to document health-related QoL in children with acute lymphoblastic leukaemia in remission during post-induction chemotherapy. Eighteen families took part, and children ranged from 11 months to 14 years in age. Ratings were made by nurses, parents and where appropriate, by children themselves. The data suggest that the burden of morbidity is cyclical in nature, mirroring the schedule of chemotherapy. The impact on QoL was least at the onset of the treatment cycle (following a week of no treatment) and greatest at the beginning of the second week (following use of steroids). Pain was the most frequently reported indicator of morbidity, followed by emotion and mobility. This study is exemplary in documenting the relationship between steroid therapy and QoL, and suggests also that the HUI (marks 2 and 3) is sufficiently sensitive in this context. In providing direct evidence of the morbidity associated with maintenance chemotherapy, and evidence for the sensitivity of the HUI measures, it paves the way for future work in which QoL outcomes can be assessed in clinical trials.

The growth in measures of QoL specifically for work with children with non-life-threatening

conditions such as asthma or eczema will provide the opportunity to evaluate treatment change in terms of QoL rather than, or in addition to clinical indicators. Such a shift involving the focus on outcomes of interest to the patient, is to be welcomed. An additional benefit may involve greater insight into factors such as patient adherence to treatment recommendations. (It might be expected that patients who perceive improvements in their QoL would be generally more adherent than those who perceive no changes or a reduction in QoL.[309] Future research concerned with the relationship between treatment regimens, QoL and adherence may offer useful insights into patient behaviour and decision-making.)

It is important that new treatments are assessed in terms of safety efficacy as well as QoL. If patients themselves do not notice an improvement they are unlikely to adhere to protocols. Thus, Harper and co-workers[310] compared multiple short courses of cyclosporin (12 weeks) with continuous therapy for a year with respect to efficacy, safety, tolerability and QoL. Although there were no differences in QoL at 12 weeks, it was better for both child and family in the continuous treatment group at 12 months.

QoL measures also have potential in evaluating the effects of community-based screening programmes. They may be used in two ways. First screening may result in considerable distress for some families. Where a child is identified as having a condition, QoL is necessarily compromised. However, it is also possible that even where the child is identified as not at risk, the degree of anxiety is accelerated. Thus, QoL measures may be useful in assessing the immediate emotional burden of screening. There is also the issue of what follows from screening. If a child is identified through screening to have a hearing problem, then justification for the value must be assessed in relation to the ensuing benefits to the child of early identification of the problem. Interventions to improve language or cognition are important but overall evaluation of benefits may be expected to have wider ramifications in terms of facilitating the child's development and normal family relationships. Again, there seems to be huge scope for QoL measures in community-based screening programmes and interventions.

Evaluating interventions

QoL measures have a potential role in evaluating the efficacy of psychosocial interventions. The extensive literature describing the social and psychological consequences of chronic disease for children suggests that much more attention needs to be paid to alleviating the impact of disease and treatment rather than focusing on identification and description.[311] In practice, there have been few reported evaluations of psychosocial interventions for children that include specific measurement of QoL. (This is not to say that many interventions are not undertaken, but they are often not evaluated.) In part, the lack of formal evaluation can be attributed to the non-availability of suitable measures. There may also be ethical objections where interventions are offered only to some children.

One exception was reported by Kazak and co-workers,[312] who devised an intervention to reduce distress among children undergoing bone marrow aspirations and lumbar punctures. Children were compared in two arms of a randomised, controlled prospective study. One group ($n = 45$) received a pharmacological intervention only, and the other group a combined pharmacological and psychological intervention ($n = 47$). Outcome measures included child distress, parent-rated assessment of child QoL[145] and parenting stress. Child distress was lower in the combined intervention group, but there were no differences in QoL scores. (Subsidiary analyses suggested that QoL for both groups improved over the 6-month course of the study. Thus, there is evidence that this QoL measure lacks sensitivity to identify differences as a function of the intervention, but may more adequately reflect changes over time.)

The importance of QoL data in evaluating interventions is emphasised in a study by Jelalian and co-workers.[313] They reported an intervention to increase food intake and weight gain in children with CF. Although the intervention was successful (in terms of weight gain), families rated the accompanying improvements in the child's energy levels and ability to participate in sports as equally important.

These studies support recommendations to include QoL measures in evaluating treatment interventions. Perceived improvements in QoL may be more important to child and family QoL than changes in clinical function noted by paediatricians.

Allocating scarce resources

QoL measures have often been justified in terms of their potential role in allocation of resources and public policy decision-making.[83] Parents too have expressed concern about the value of treatment for all infants.[314] In recognition of this, guidelines have been published regarding eligibility for intensive care.[315] Of relevance to this question is work by de Keizer and co-workers,[316] who attempted to predict health status and QoL in children following admission to intensive care,

based on HUI scores. The most important predictors of health status 1 year after admission were level of sensation, mobility and cognition on admission. Other contributing variables included ability for self-care, systolic blood pressure, oxygen, Glasgow Coma Scale scores, glucose and age.

While considerable work needs to be conducted in this area, the ability to predict future health status on the basis of simple measurement such as this is potentially useful in clinical decision-making. Currently available QoL measures, however, may not be perceived to be sufficiently well developed for use in this context. 'Ultimately, QoL assessment can provide data regarding the true cost of childhood cancer, taking into account not only life-years saved, but also reflecting a quality-adjusted analysis of those benefits'.[83] QoL data may have some potential for prioritising and allocating resources, but currently available measures may lack the sophistication that would be demanded.

Assessing the outcomes of treatment in the long term

Increasing long-term survival, albeit with greater morbidity, has accounted for much of the work involving QoL. The issue is particularly acute for survivors of potentially life-threatening conditions such as childhood cancer or renal disease, survivors of neonatal intensive care, and previously healthy children who experience major accidents or injuries. For many of these children, survival is associated with major stresses and compromises to everyday function. The financial cost involved in continuing medical care is a considerable burden for the health service. Assessment of QoL is important in order to provide an appropriate follow-up service in both the immediate and long term.

The most comprehensive series of follow-up studies involving QoL assessments have been reported for survivors of neonatal intensive care by Saigal and co-workers.[317] Based on assessments involving the HUI measures,[78] survivors of neonatal intensive care report a greater burden of morbidity and rate their QoL to be significantly lower than age-matched healthy controls.[317] Even so, most survivors were relatively satisfied with their QoL. Findings such as these perhaps justify use of QoL measures in assessing outcomes for these survivors. Based on simple outcome measures, it is clear that the group as a whole have considerable morbidity. However, children are able to enjoy a good QoL and find ways to compensate for any physical morbidity. However, in that it was not possible to include the most severely affected individuals in the study, some reservations remain about the QoL of survivors of neonatal intensive care. Considerable follow-up of survivors of childhood cancer has also taken place (for review see Eiser[318]), although much of this has not been based on formal assessment of QoL. A number of outcome measures are now available for this purpose; indeed some measures have been developed specifically to assess QoL in survivors of childhood cancer (e.g. Varni *et al.* [143]).

Palliative care

The focus in medical training on treatment and cure may make it difficult for some medical and nursing staff to accept when palliative care is more appropriate. For some, a decision to recommend palliative care can be perceived to be a failure.[319, 320] Continuing intensive treatment in the face of progressive disease can be a major obstacle to the attainment of good QoL. One difficulty in assessing QoL in palliative care lies in defining exactly when curative care ceases and terminal care begins. Many children may experience a period of normal life even after the point when cure is no longer perceived to be possible. "The transition between aggressive treatments to cure or prolong good quality life and palliative care may not be clear".[321] As has been recognised in adult work, it is important to evaluate QoL at an early stage and not wait till the patient is too ill to benefit from any assessment.

Guidelines for improving the QoL of children in palliative care have been defined by a Working Party convened by the Societe Internationale D'Oncologie Pediatrique and chaired by Jankovic.[322] In considering good practice in palliative care, they recommend:

- evaluation of the child's physical, psychological and practical needs
- research into the use of optimising analgesia, control of emesis and psychological symptoms
- consideration of the impact of the palliative care team on the child and family
- consideration of the role of cytotoxic agents and radiotherapy
- assessment of QoL
- formal training within medicine on management of the dying child.

Although she does not focus her discussion of palliative care on QoL, Goldman[321] makes a number of relevant points. Factors such as the child's understanding of illness and death need to be considered and appropriate means of communication (play, drama, art, school-work) adopted. Goldman discusses the use of faces scales to assess pain, but does not consider formal QoL

scales. Certainly no measure currently available would seem to be appropriate for these purposes. However, standardised measures of QoL in palliative care may be useful in making decisions about the efficacy of alternative treatments and assessing change in the individual child's evaluation of their own QoL. They may also have a place in determining the child's views about the value of treatment.

Reluctance to assess QoL during palliative care is invariably based on assumptions that patients are too ill to rate their own QoL. The assumption has rarely been tested. For children, a good QoL measure for work in palliative care probably needs to take into account the control of symptoms, the impact on the family of treatment both in the immediate and longer term, and the child's emotional and spiritual well-being. Assessments of the effectiveness of treatment must not only focus on the immediate consequences of the child's death, but also the longer-term implications for the family. Clearly none of the currently available generic measures have sufficient sensitivity for this purpose. There remains however, a need for some means of assessing how far treatments used in palliative care really improve QoL either for the child or for the family.

Barriers to adoption of QoL measures

Despite the amount of work invested in developing QoL measures, it is disappointing how little they have been adopted for use in either research or general clinical practice. The dissemination of information about QoL, and inclusion in everyday practice must be a goal, but for the moment interest remains at an academic level. In the absence of hard data, we may need to speculate as to why QoL measures have not been used in outcome studies as much as might be hoped. They may be perceived as 'soft' and time consuming.[323] Clinicians question particularly the significance or meaning of small changes in QoL and have difficulty understanding how far changes in scores have clinical, rather than statistical, significance.[68,324] Thus, there is potential conflict between scale developers with a concern for statistical significance and clinicians interested in clinical significance. The relationship between statistical and clinical significance needs to be clarified.

Additional barriers relate to the interpretation of QoL data. In terms of clinical need, the question of meaningfulness of QoL scores has not generally been addressed by those involved in developing measures. Thus, if change in QoL is measured on a seven-point Likert scale, we need to determine the practical meaning to the patient of a change from a rating of 4 to 6. Does this represent twice as good an improvement as a change from 4 to 5? There is scope here for considerable research work.

As so often in work with children, in the absence of hard data we need to draw on work with adults to answer this question. Although little is formally known about the attitudes of clinicians to use of QoL measures, a report by Taylor and co-workers[325] yields some insight into clinician attitudes. Sixty adult cancer clinicians were interviewed about their understanding of QoL. They were asked about their understanding of definitions and measurement, and their attitudes to using QoL information in their practice. Although most were aware that QoL could be defined and measured, a number of barriers to use were identified.

The majority questioned the reliability and validity of currently available measures and were of the view that collecting QoL data would impose additional time burdens on staff. In addition, rather than facilitate decision-making, there was a view that inclusion of QoL data made decisions more difficult. Particularly where QoL data is collected by other professionals, clinicians felt that their authority was challenged, and distance created between them and their patients.

Despite this, clinicians felt that responsibility for QoL data rested with nursing or healthcare workers. QoL data were perceived to be interesting but not necessarily the responsibility of medical staff. They questioned the face validity of patient responses to QoL scales and identified their lack of knowledge and inability to keep abreast of emerging issues as a reason for their reluctance to use QoL instruments. Questions about the generalisability or specificity of instruments was a further problem. In the context of clinical trials, many accepted the potential value of QoL data in assessing the overall value of the trial, but argued that the data lacked specificity for individual care.

Discussions with paediatricians suggest other reasons. There is of course, a concern that measurement of QoL takes too much time for integration in a busy paediatric clinic. Use of paper and pencil measures may be seen to compromise attempts to establish rapport with the child. This is a real concern and provides an additional reason why modification of adult measures may not be appropriate for work with children. However, we must work toward a

balance between the very informal approaches to collection of QoL data adopted by most paediatricians in everyday practice, and the rigidness and impersonal approaches that characterise most paper and pencil measures currently available.

Summary

- In justifying the need for this review in chapter 1, it was argued that QoL measures were potentially useful in a variety of contexts. These may include:
  - comparing outcomes in clinical trials
  - evaluating interventions
  - allocating scarce resources
  - assessing the outcomes of treatment in the long term, and
  - palliative care.
- Despite the increasing interest in developing measures of QoL, they have not been used routinely in research as evidenced by our computer-based searches. (It is not known how far they are used in routine clinical practice.)
- In accounting for the limited interest in using QoL measures, we speculate that there are a number of perceived barriers, which probably include the following:
  - concern with psychometric properties of measures such as validity and reliability and much less focus on the content and acceptability to families
  - perceptions that QoL measurement increases the time of a consultation, is the domain of nurses and psychological support staff rather than the paediatrician, and imposes an additional and unnecessary burden on families.

Chapter 9

Measuring QoL in children – the state of the art

Aims and scope

We first consider the implications of our review for the four key questions identified by the HTA Programme. We go on to make wider inferences, regarding the future development and value of QoL measures in paediatrics. Although adult work provides one frame of reference for paediatric work, disagreements among those involved in adult QoL work suggests that considerable caution needs to be taken. Given that interest in QoL is relatively recent, our review is necessarily limited both by our own approach to identification of papers and by the quality of work identified. For this reason, we consider in some detail, the limitations of the review.

Given our findings, we make some recommendations regarding the most suitable measures for assessing child QoL in different contexts (e.g. clinical trials, evaluating interventions). We conclude with an agenda for priorities for future development. First, there does remain a need to continue to refine and improve the quality of measures available. A strategy for future development is outlined.

QoL – the heart of paediatrics?

Priorities for the healthcare of children have been the subject of some review.[326] Based on the principles defined by recent legislation (The Children Act, 1999[327] and the Education Act, 1996[328]) guidance is given for purchasing authorities about the services that should be provided to enable as many children as possible to reach adult life with their potential uncompromised by illness, environmental hazard or unhealthy lifestyle.[329] Social Service, Education Departments and Child Health Services are required to work together in order to assess and provide for children in need and those with special education needs. How this works has been subject to scrutiny but recommendations have been made for the definition of joint planning mechanisms. However, decisions about resource allocation and service provisions are blighted by the absence of evidence-based methods to distinguish between different therapeutic interventions or modes of service delivery.

The extension of QoL research, from a small area of well-defined and serious health problems such as childhood cancer or very low birth-weight, to include measures of the well-being of children in general, and to include those children whose needs span the interdisciplinary and interagency professional boundaries, is an essential step towards measuring health gain and efficiency within an integrated children's service. QoL is no longer the sole province of those involved in high-technology disciplines, but is increasingly of concern to a wider audience concerned with the well-being of children in general. The QoL of children is fundamental to the aims of paediatricians, as exemplified by the mission statement of the Royal College of Paediatrics and Child Health (RCPCH), and implicit within the accepted 'duties' of a paediatrician.

> "The Royal College of Paediatrics and Child Health **advances** standards of medical care for children, **builds** programmes of research and training and **champions** the rights of children." [http://www.rcpch.ac.uk/rcpch/about/appeal/index.htm] (original emphasis)

Among the 'Duties of a Paediatrician' specified by the RCPCH are the following.

- Paediatricians should pay due regard to the domestic, sociological, environmental and genetic dimensions of the health of children.
- Paediatricians, whatever their speciality interest, should understand their particular responsibilities for the holistic and life-long health of children who come under their care: each contact is an opportunity for health promotion and disease prevention.
- Paediatricians should be aware of current medical and political affairs affecting the lives and health of children.
- Paediatricians should serve as advocates for the health needs of children locally, nationally and internationally.
- Paediatricians should see themselves as ambassadors for children and for the speciality of paediatrics.
- Above all, paediatricians should be courteous and compassionate in all their professional dealings with children, their parents and other carers, placing the child's best interests at the centre of all clinical considerations.

Although QoL is not mentioned specifically, it is clear that the attainment of the child's QoL is at the heart of paediatrics. Whether defined as the child's best interests, or the 'holistic and life-long health of children,' QoL is a central concern in practice.

In this final chapter, we first summarise our findings with respect to the four key questions defined at the outset. We then attempt to extend our conclusions, and consider the wider implications of the review for the future development of QoL measures and their integration into clinical practice and research.

Implications of the review for the four key questions

To what extent are adult measures used in the evaluation of healthcare interventions in children?

Given that work involving measurement of adult QoL is more established than comparable work for children, it follows that it should be possible to learn from experience gained in the adult field. From this perspective, asking how far, and how well, measurement of QoL for children can benefit from what has been learnt in adult work, is very reasonable. We need to remember that there remain many difficulties in measuring QoL in adults, and disagreements between key workers about theory, definition and measurement.

In reviewing adult measures, Aaronson[330] emphasises the biases in scale development, amounting in many instances to a 'reinvention of the wheel.' He observes that the preoccupation with development of measures "appears to reflect the system of rewards operating in the social sciences, rather than a legitimate gap in the available resource pool".

Any decision to develop child measures by changing adult measures needs to proceed in the knowledge of the limitations that have come to light. There are concerns among those involved in the development of adult measures about the limitations of current methods and theoretical perspectives employed. In addition, measures developed for adults are likely to be inappropriate for work with children. These difficulties need to be considered in relation to other areas where it is apparent that adult work is not an appropriate model for paediatrics. Children have specific needs and cannot be considered as small adults. This also needs to be recognised at a policy level. There is inadequate representation of children and adolescents in membership of modernisation action teams to draw up action plans for health.[331] Children are not routinely considered to be a special group with regard to funding for research or development; improving child health is not a high priority for local service delivery; specific services for children are not available in 40% of the UK.[191] Services for children can only be improved when we stop treating them as an extension of adult services and make more appropriate provision for, and evaluation of, their QoL.

Conclusions

- Given the more extensive work concerning adult QoL, it is reasonable to ask how far this can guide work in paediatrics.
- As a model for work in paediatrics, we need to understand the limitations of adult work, particularly in terms of disagreements about theory, definition and method.
- There needs to be greater recognition of the special needs of children in provision of services and development and evaluation of QoL measures.
- Considerable caution needs to be exercised in using adult work as a model for paediatrics.

How appropriate are adult measures for use with children?

We identified a number of reasons why care needs to be taken in applying adult models for work with children. Measures developed for adults are likely to be too long for children. Insufficient account is taken of children's different cognitive and language skills, or their abilities to understand and use rating scales. At the least, assessment of the reading age necessary to complete a measure should be provided. Critically, key domains used to measure adult QoL may lack appropriateness for assessment of children's QoL. This response burden and lack of suitability of domains are key arguments against general adaptation of adult measures for children. Our review leads us to conclude that modifying adult measures is not to be recommended.

New measures for children need to draw on theories of cognitive, social and emotional development. They need to adopt more child-friendly approaches, and we recommend the use of puppets, pictures and computers. These may be more attractive for children, and incidentally counter paediatricians' arguments that collection of QoL data jeopardises the establishment of rapport

with children. Crucially, development of measures needs to draw on work that describes children's abilities to use rating scales of different lengths, and their understanding of emotional terms of the type often used as anchors.

Conclusions

- Measures need to be developed which take into account the child's understanding of health and illness, language and reading skills, and ability to use rating scales.
- Publication of new QoL measures should include information about reliability, validity and sensitivity, as well as estimates of reading age necessary for completion.
- We need to move away from paper and pencil assessment of QoL, and develop measures that are more child friendly (e.g. include pictorial support, use of computers).

To what extent do child self-reports correspond with the assessment made by parents and carers?

Implicit in many definitions of QoL is the notion that it is unique to the individual. However, we have to accept that there are circumstances in which we have to rely on proxy raters. The most common include situations where children are too young or too ill to answer for themselves. In these situations, a critical question is how far ratings between child and proxy can be expected to match. There are a number of reasons why we might not expect concordance between children and their parents. They do not necessarily share the same perspective or experience of an event. Indeed, we must accept that it is a natural part of development that children do not want to share all their thoughts and feelings with their parents. Furthermore, parents have their own concerns and anxieties which may colour their perception of the child's QoL.

Obtaining QoL data directly from children may be the most desirable option, but we have to allow for situations where this is not possible. In order to anticipate the range of situations in which QoL ratings may be useful, it is important that measures of QoL provide for parallel ratings by child and proxy. Given the importance of this, basic research is needed to determine the relationship between child and proxy ratings, taking into account variables such as age, gender and health status. The use of proxies other than parents needs to be explored. Additional work needs to clarify how far child and proxy ratings differ depending on the domain of QoL assessed.

Conclusions

- For children who are too young, ill, or unable to answer for themselves, all measures should include some provision for proxy ratings.
- All measures should include versions for self-completion by children as well as parallel forms for proxies. Questions should be re-phrased away from the search for concordance and towards understanding the circumstances in which parents and children agree (or disagree) about the child's QoL.
- Parents' ratings of the child's QoL are influenced by their own health and well-being. It is important to develop a research programme to increase understanding of how parent mental health influences their perceptions of child QoL.
- Proxies other than parents may have some value and should be used more widely.
- Basic research to identify characteristics of parent and child that are associated with more 'accuracy' is needed, as well as to determine behaviours and emotions that are more amenable to accurate proxy report.

How feasible and reliable are proxy measures of QoL in different disease contexts?

In the absence of comprehensive and validated QoL measures, the battery approach, which involves assessment of established measures related to QoL, is potentially useful. However, the use of a battery of measures to assess QoL in children has not been a popular method. Many studies that were identified were excluded for lack of methodological rigour. Any battery is only as good as the individual measures involved. The practice of selecting some items or scales from established proxy measures is not recommended, as it results in measures of unknown psychometric quality. Batteries tend to impose considerable response burden on children. As more comprehensive measures of QoL become available, there would seem to be few arguments in favour of this method.

Conclusions

- The battery approach is limited by the quality of measures available.
- The quality of studies using this approach was very limited, and restricts our ability to evaluate the approach.
- There are considerable practical disadvantages in using a battery approach with children (e.g. length).
- The development of comprehensive QoL measures limits the value of this approach.

The future development and integration of QoL measures into research and practice

Theoretical and methodological issues

In reviewing adult work, Battista and Hodge[332] conclude that much of the work in the last two decades has focused on three issues. These include:

- the concern with 'tool-making' with the emphasis on developing measures and establishing their psychometric properties
- the emphasis on multidimensional constructs and concern to define the domains that should be sampled
- the question of whether multidimensionality and specificity can be captured in generic measures, and thus how far disease-specific instruments are needed.

These issues are as relevant in paediatric QoL assessment as in adults.

Tool-making

The need for high methodological standards in the development of measures for both children and adults is paramount. Many currently available measures fall short of the standards that could reasonably be expected. First, a common procedure in determining items in a measure involves preliminary interviews or focus groups with a small number of patients (or clinicians). The items in a measure are necessarily based on the views of a very small number of patients who have agreed to be involved and may not be representative of the total population.

To date, considerable attention has been given to establishing the psychometric properties of a measure, and much less to issues of content and face validity. There has been little recognition of how far psychometric properties such as reliability and validity are dependent on sample characteristics, or across different populations and cultures. A measure with validity established in one population will not necessarily perform as well in another.

Where there is a clear conceptual understanding about what is to be measured, the use of appropriate psychometric techniques helps to ensure that measures have acceptable psychometric properties. It is customary to report reliability (internal consistency) or validity, with an assumption that if a measure fulfils these requirements, little else needs to be considered. However, adherence to recommended psychometric techniques does not by itself guarantee that the resulting measure reflects the concept to be measured.

First, the practice of excluding certain items on the basis of statistical criteria can lead to some bias. Items that may be important to individuals are excluded while items that fulfil psychometric requirements may be included, even if they would not be endorsed by any single individual.

Second, it is possible for items to hang together to form a scale but not necessarily measure QoL. What psychometrics does is to establish the items to be included purely on statistical grounds, in particular by determining the inter-correlations between items. As such, psychometrics does not take into account the content or meaning of the items. There is therefore no guarantee that the scale really covers the full range of items of interest. The risk is that some scales may be highly internally consistent but reflect a very narrow range in terms of items of interest. Thus, we need to look beyond levels of Cronbach's alpha when evaluating a measure.

Third, however the initial item pool is determined, further issues relate to how these items are organised into domains. Some argue that factor analysis is the ideal method for establishing construct validity.[94] Multitrait-multimethod modelling, exploratory factor analysis and confirmatory factor analysis are sophisticated statistical techniques, which are yet to be exploited fully in the measurement of QoL. In many cases, given small sample sizes, factor analysis is not possible, and decisions about the organisation of items into domains are made on more *ad hoc* bases. Validity testing is often based on correlations with clinical indices such as severity of the disease, or time since diagnosis. Clinical status measures of this kind are not necessarily related to QoL in any simple fashion. Indeed, if they were, we might ask why we want to measure QoL at all.

Disillusionment with conventional approaches to scale development has resulted in calls for more individualised or qualitative approaches. For adults, these include approaches based on patient-generated measures. This approach draws on definitions of QoL involving perceived differences between an individual's hopes and expectations and their present experience.[63]

Ruta and co-workers[333] describe the development of the Patient Generated Index (PGI) which attempts to reflect patients' views about their health and its meaning in the context of their lives. The PGI is completed in three stages. Patients are first asked to list five areas of their lives affected by their condition. They then rate how badly they are

affected in each of these areas on a scale from 0 to 100, where 0 represents the worse they can imagine and 100 represents exactly how they would like to be. A sixth box allows them to rate all other areas of their life. In the third and final stage, patients are asked to imagine they can improve some or all of the chosen areas in their life. They are given 60 points to spend across one or more areas. Points allocated to each area represent the relative importance of potential improvements in each area. By multiplying each of the six ratings by the proportion of points allocated to the area and summing these scores, an index is generated between 0 and 100. The resulting score is assumed to represent the extent to which reality falls short of patients' hopes and expectations in areas of life that they value most. The authors provide evidence which suggests that the method is acceptable and comprehensible to adult patients.

The basic assumption of the PGI, that the domains of QoL need to reflect an individual's lifestyle, wants and aspirations, is shared by a number of other workers.[334] These assumptions are reflected in other measures for adults (The Schedule for the Evaluation of Individual QoL[335]), but have not yet been fully utilised for children.

Theoretical concepts and their interpretations

There is increasing disillusionment with theory, application and measurement among some of those working in adult QoL. The lack of theory is perceived to be a major limitation. Without theory 'there is no means of linking what is actually and what is supposedly being measured'.[336] Lack of theory has contributed to the diversity in number and definition of domains and lack of clarity about how the separate domains relate together.

The need for a more theoretically driven account of QoL has been suggested by a number of workers. Within adult QoL research, some attempts to use causal modelling approaches to specify the relationships between factors contributing to QoL have been made.[48,337] Fayers and co-workers[95] stress the need to distinguish between factors that are **causal** and those that are **consequences** or **effects** of QoL. Causal factors include symptoms and side-effects experienced as part of treatment programmes which in themselves may cause poor QoL. Other items such as anxiety or depression may reflect underlying QoL. These effect indicators can serve to highlight poor QoL. The authors suggest that causal and effect indicators have fundamentally different relationships with QoL. Causal indicators may be a good indicator of poor, but not of good health-related QoL. This is because the presence of symptoms may reliably indicate poor QoL but the absence of symptoms does not necessarily mean that patients report good QoL. In contrast, effect indicators may successfully indicate both good and poor QoL.

To the extent that most definitions of QoL emphasise the patient's perspective, a theory which attempts to account for the processes underlying how patients make decisions about their QoL is needed. Thus, we need to specify how the individual makes a decision as to whether QoL is good or bad. These decisions are likely to be made on the basis of comparisons between perceptions of past QoL, or between current functioning and how far this differs from activities engaged in by salient comparison groups. Decisions about who children compare themselves with (e.g. a healthy peer or patient with similar disease), and whether the comparison is made along dimensions that result in positive or negative views about themselves, will have implications for their judgements of their own QoL.[338, 339]

To a large extent, the lack of theory is at the heart of the confusion underlying QoL. Without a theory "We have to guard against defining QoL in terms of what is measured by instruments named QoL."[340]

Multidimensionality and specificity in generic measures, and disease-specific instruments

Although developed to take into account the 'patient's perspective, many scales focus heavily on objective indices such as experience of symptoms, and relatively little on the patient's perspective. Functional status and symptoms are not synonymous with QoL'.[336] Many articles begin by defining QoL in terms of its multidimensional nature, and imply that there is agreement about the nature of these domains. In fact, as we have shown, there is considerable variability in both the number and definition of domains in both child and adult work.

Issues about the number and definition of domains are at the heart of arguments about the merits of generic and disease-specific measures. Generic and disease-specific measures tend to be developed for different purposes. Many generic measures were developed for use in population-based health surveys, and are considered appropriate for all children regardless of health status.

There are, however, serious limitations with disease-specific measures, in that they do not allow for comparisons across samples of children differing in health status. Disease-specific measures were developed to understand the range of responses

to be found within a population of children with a defined condition. They include items to assess the child's concern about specific aspects of treatment and symptoms that reflect the disease.

There is a tendency to see generic and disease-specific approaches to be diametrically opposite. Assumptions that disease-specific measures are preferable in clinical trial work have not been formally established. Use of both generic and disease-specific measures may have some advantage but imposes a huge burden on patients. Yet neither alone will give a complete assessment. Solutions may include the development of measures involving a generic core with disease-specific modules added where appropriate.[123] These provide the opportunity to compare children with different conditions (based on generic sections), and to assess the impact of disease (based on disease-specific sections). Modules are currently available for conditions including cancer, diabetes, asthma and arthritis. However, we need to bear in mind that this approach is not suitable for children with more than one condition, and that many children will have a disease for which no module is available.

Limitations of the review

It is important to consider two main sources of limitations in our review. Our conclusions are necessarily limited by the processes involved in conducting the searches and the decisions we made regarding inclusion and exclusion criteria. In addition, they are limited by the quantity and quality of literature available at the time. Assessment of QoL in children is a newly emerging field, and our conclusions are therefore based on a relatively limited body of empirical work.

In conducting the literature searches, we took the decision to focus on specific databases that were available to us. It is possible that different studies might have been identified if we had searched different databases. In addition, we restricted this review to work in English language journals. Any other decision would have increased the costs, and required the employment of translators. The result may be some bias to the inclusion of studies conducted in the USA and UK.

We also took a decision to adopt very generous criteria regarding definition of QoL so that we included any measure supposedly measuring QoL, health status, or well-being. As a consequence, we probably included a larger number of studies than if we had employed stricter criteria focusing specifically on QoL. Our view was that such a decision would result in so few articles to be included that few meaningful conclusions could be drawn. Adopting more strict criteria and including only measures that reflect QoL as opposed to health status would have resulted in a very brief review. Given the current interest in QoL measure-ment, we felt it was important to adopt this broad-based approach to inclusion in order to contribute to greater clarification of exactly what is being measured in the future.

Comparison of our work with previous reviews suggests limited overlap in terms of measures of QoL identified. For example, we have included fewer measures than Bullinger and Ravens-Sieberer[18] and more than Levi and Drotar.[341] However, examination of the specific measures included in the different reviews suggests that discrepancies arise through lack of definition of QoL. Previous reviews differ in how far they include measures of function, well-being, or health status in addition to QoL. The discrepancies do not arise because we have systematically excluded important articles in languages other than English. The inclusion in our review of a number of measures, which were originally developed in European languages but subsequently translated into English, goes some way in off-setting any criticisms of bias.

Of greater fundamental importance, our conclusions, particularly those connected with the four questions we were asked to answer, are necessarily limited by the quality of published work in this area. This applies particularly to the questions involving use of proxy respondents for children who are unable to provide QoL data themselves. We feel that quite different conclusions might be found depending on the specific measures of QoL used. This criticism also adds strength to the argument that measurement of QoL needs to be theory driven, rather than subject to the vagaries of idiosyncratic measures.

Choices in measuring child QoL

Choice in relation to purpose

> "If the past two decades have seen the triumph of psychometrics, the field appears poised for a post-psychometric phase, for advancing QoL's role in policy requires advocacy of the very concept of QoL as a socially desirable and demanded outcome."[332]

All currently available measures have some limitations. The measurement of any psychological concept such as QoL is inherently different from measuring a physical concept such as height, and it may therefore be inevitable that we must live with some limitation in any measure. However, this is not to say that we should give up on measuring QoL. For children, QoL is too important to be disregarded. Further development of measures depends crucially on experience gained in using the measures that are now available. This is relevant not only for refinement of currently available measures, but also to enable the development of more sophisticated measures in the future. For these reasons, it is important to recognise the limitations of currently available measures, while also acknowledging that improve-ments can only be developed when we understand how current measures perform in practice.

Given the current state of the art, we draw on information about the performance characteristics of available measures summarised in *Tables 2* and *3*. Based on these data, we conclude that only three generic measures and two disease-specific measures fulfil the very basic criteria identified. Our own recommendations regarding choices for different purposes might therefore involve the following.

For work evaluating clinical trials, whether in the context of high technology medicine such as childhood cancer, or in a community setting, there is a need for a brief measure of QoL that can be completed during a regular clinic visit. In order to recruit a large sample of patients, a measure is needed that is simple to administer with minimal training or expertise. The measure needs to include those aspects of functioning that are most likely to be compromised by the treatment protocol. Thus there is a need for measures that focus on physical symptoms and emotional well-being. Particularly for children (compared with adults) and if there is any concern about cognitive side-effects of the protocol, then assessment of school or learning needs to be included. Given the concern with physical symptoms, it is likely that disease-specific measures might be more useful than generic. The PedsQL and its associated modules for work in oncology, asthma or diabetes[143] is one of the more thoroughly developed measures currently available. In asthma, the measure by Juniper and co-workers[140] also has much to recommend it.

The inclusion of QoL data in clinical trials creates new questions about statistical analyses which have not been resolved. The analysis of multivariate QoL data (and the inevitable missing data) poses a very different problem compared with analyses based on univariate outcomes such as survival. Strategies to manage missing data are important, as is the need for hypothesis-driven trials

The choice of measures for evaluation of psychosocial interventions is relatively similar. If the need is for a brief assessment then generic measures such as the PedsQL[123] or HUI2 or HUI3[78] have some merit. However, it is unlikely that either of these will address the full range of functioning that might need to be assessed (and indeed they were not designed to do so). Additional measures will therefore need to be included depending on the specific purpose of the intervention. Where the goal is to achieve greater school integration or improve family functioning, then the CHQ[19] may be more appropriate. However, advantages of the CHQ need to be set against the length of the currently available measure (though shorter forms are in the process of development).

There are also measures developed for specific purposes, such as the BASES[141] for work involving children undergoing bone marrow transplantation. This does fulfil the basic criteria we identified, and has potential use in evaluating interventions involving children undergoing bone marrow transplantation. It is clear that there are many other specific contexts in paediatrics where Qol measures may be desirable (e.g. palliative care), but no measure is currently available.

Cross-cultural work

In evaluating new treatments involving rare conditions, international measures of QoL have been advocated. Sharing data in drug trials may lead to more rapid evaluations of the efficacy of new treatment than relying on single-nation studies. International measures may have other uses in monitoring the health of populations and in programme evaluation. For cross-cultural work, it is perhaps attractive to consider translation of measures already available.

However, potential pitfalls in the translation of measures need to be recognised. First, measures developed for a specific purpose in one culture may be inappropriate elsewhere.

Second, specific items may differ in meaning between cultures. 'Feeling sick' means feeling nauseous to British children, but is used as a more general indicator of illness in other cultures.

Third, norms developed in one culture may not translate well to others. For example, cut-off points

for normal and abnormal behaviour on the CBCL[300] differ for USA and Australian children.[342,343] Differences between cultures may be attributable to genuine differences between the prevalence of attributes in cultures, or may be an artefact of ambiguous language use.[112] Translating a QoL instrument for use in different countries may appear a cheap and satisfactory option, but in fact requires considerable work to establish true comparability. Norms established in one country cannot be assumed to have the same meaning elsewhere.

In recognition of these problems, recommendations about the process of translation of QoL measures have been made, largely with respect to adult work,[344] and more specifically for work with children.[345] A process of forward and backward translation, comparison of the back translation with the original source, review by lay-panels, committee review and psychometric testing of the translated version is normally recommended. At least four levels of cross-cultural equivalence have been defined: conceptual equivalence; construct or item equivalence; operational equivalence and metric equivalence.[346, 347]

Conceptual equivalence refers to the extent to which the items in the target language are similar to the source, and must include both semantic meaning and formulation of the items as well as the underlying concept. This is achieved through the translation process and qualitative testing. **Construct** or **item equivalence** is the extent to which individuals in different cultural groups respond to the same item in similar ways and is evaluated using classical test theory (test–retest reliability). **Operational equivalence** refers to the relative performance of the instrument using different modes of administration (self-report, interview). **Metric equivalence** is the extent to which individuals in different cultures are ranked similarly along a continuum of QoL; (patients with CF in different countries who have similar pulmonary scores will have similar scores on a QoL measure).

There is neither agreement among researchers about defining these different aspects of equivalence, nor about how different types of equivalence should be assessed. On the basis of a literature search, Herdman and co-workers[348] for example, found that the four-way classification described above was too simple. He was able to identify 19 different types of equivalence. Conceptual (30%), semantic (12%), functional (8%) and scale equivalence (8%) were reported most frequently. Although there was almost universal agreement about the definition of semantic equivalence, there was much less agreement regarding conceptual and functional equivalence. Herdman and co-workers[348] concluded that there are real short-comings in current approaches involving translation of QoL instruments, which can be attributed to lack of agreement in definitions of equivalence, as well as the assumptions underlying cross-cultural research.

Notwithstanding rigorous translation processes, there is a fundamental question concerning how well any measure developed in one culture can satisfactorily address QoL in another. The approach of the WHOQOL Group[349] in which measures are developed simultaneously in different cultures represents one solution to this problem. Although this approach has resulted in a measure of QoL for adults,[349] work with children is at a more preliminary level.[350]

To the extent that there has been greater investment from the USA, many measures recommended above may in fact prove less acceptable, given cultural differences in the meaning of illness, organisation of healthcare services and relations between parents and children. Consideration needs to be given to the language (e.g. questions about 'difficulties walking one block' don't mean very much to children outside the USA).

Other issues may be even more critical. In the cancer-specific QoL measure described by Varni and co-workers[143] for example, a number of questions, directed at children ask about their concern about relapse, or their cancer coming back. Inclusion of such direct questions (or even use of the term 'cancer') may be unacceptable to some paediatricians and families in the UK. Despite the practical difficulties, a number of efforts are currently being made to translate QoL measures for children into different languages.

- The HUI is undergoing development in Austria, France, Japan, The Netherlands, Singapore and the UK.[351]
- A measure to assess QoL in CF developed initially in French by Henry and co-workers has been translated into German and Spanish.[352] In addition, Quittner and her group have reported initial steps in translation for use in the USA.[353] (The original measure was not included in the review as only abstracts were available).
- The CHQ[15] is being adopted by a number of different groups. Australian norms have been reported by Waters and co-workers.[354] Efforts to develop an Anglicised version of this measure are being undertaken by Eiser and colleagues in Swansea.[355]

- The Childhood Asthma questionnaires,[248] originally developed in the UK have been modified for work with Australian samples. Although mean QoL scores did not differ between British and Australian children with asthma, scores for the subset of terms relevant for normal children showed that QoL scores were higher for Australian children. These differences were interpreted in the context of cultural expectations about QoL. French and co-workers[356] note a number of problems in developing internationally equivalent measures. Australian children (aged 11–12 years) were noted to be more socially independent and knowledgeable about their asthma compared with British children. They were also more likely than British children to raise issues about being different because of their asthma. Only British teenagers but not younger children expressed concerns about being different. As a consequence, the authors decided to use the form of the questionnaire originally developed for British teenagers even though they were working with younger children in Australia.
- The PedsQoL[143,144] is undergoing translation into a number of languages. Details are available on the website (http://www.pedsql.org).

To a large extent, workers in the QoL field assume an 'absolutist' approach to cross-cultural work (i.e. that culture has only a minimal impact on QoL, which will consequently be essentially invariant across cultures). This 'imposed ethics' results in QoL being defined and operationalised in one culture and then imposed directly on to another. These assumptions are reflected in most guidelines currently available for translation of QoL instruments, which do not involve an initial assessment of the relevance of the concepts measured in the questionnaire to the target culture. The logical extension of the assumption that QoL measures are culture free would be the failure to assess conceptual equivalence at all. Similarly, the concept of scalar equivalence can be ignored on the grounds that the pattern of relationships between scales would also be equivalent. The emphasis in the absolutist approach is on establishing semantic equivalence. The absolutist approach needs to be contrasted with the universalist and relativist approaches as described by Berry and co-workers.[206] There needs to be a more open orientation to the idea that culture has a significant impact on the way concepts are expressed and acknowledges the substantial role of culture in variations in behaviour.

Conclusions and final recommendations

Despite the rapid growth in development of measures of child QoL, the results of this review suggest that a number of fundamental issues remain to be addressed. Many of the recommendations previously made concerning the development and selection of QoL measures for use with adults apply also to work with children (see Fitzpatrick *et al.*[69]). However, as a result of this review, we would add a number of additional recommendations. These include the need for empirical research concerning:

- the implementation of QoL measures in paediatric research
- the link between theory and measurement development
- adoption of child-centred approaches to measurement
- clarification of the relationship between child and proxy ratings.

The implementation of QoL measures in paediatric research

While a more rigorous theoretical and methodological approach to the development of measures is a clear precurser to their use in practice, experience with currently available measures is also important to establish a precedent for the evaluation of the child's QoL in paediatric practice.

In this context, the rationale for including QoL measures in evaluations of clinical trials for both life-threatening and less-serious disease has been well argued. However, to date, arguments over the merits of different measures have restricted the adoption of QoL measures. We need to identify a series of randomised trials where there are clear hypotheses regarding the likely consequences of treatment for QoL. This will involve national collaboration to ensure size of samples, and critical discussion to ensure that the trial involves genuine implications for QoL. The value of QoL also needs to be considered in some screening programmes. Neonatal screening for hearing impairment is a case in point. Although currently available measures have their limitations, the development of more satisfactory assessment tools is partially dependent on information about how current measures perform in practice.

The link between theory and measurement development

We need to develop more rigorous approaches to understanding how judgements of QoL are made

and change during the course of child development. Adoption of a theoretical model that enables us to make predictions about how groups will differ, or how QoL will change during the course of a disease are essential to move us away from the current descriptive approach to work. A theoretical approach will also clarify the domains to be measured, rather than relying on *ad hoc* selection of domains and items depending on specific circumstances.

In practice, a number of theoretical approaches may potentially be of value. These include:

- **theories of child and adolescent development**, which can contribute to our understanding of how QoL changes across the life-span, and the relationship between child and proxy ratings
- **theories from Social Psychology and Sociology**, which can contribute to understanding the processes underlying judgements of QoL, particularly in so far as our perceptions of ourselves are integral with our relationship to others.

Adoption of child-centred approaches to measurement

Paper and pencil measures are unlikely to be as useful in paediatrics as they have proved in adult work. They are less attractive for children, and also for paediatricians, who may feel that they compromise the establishment of rapport. Research is needed to determine the most reliable and effective methods for obtaining reliable information from young children.

- Direct comparisons between different response scales (e.g. Likert, visual analogue, smiley faces) are required, as well as work delineating the appropriateness of these methods for work with children of different ages.
- There is a need for the development of alternatives to paper and pencil measures, including the use of puppets and interactive computer presentation.
- Research involving children's understanding of time needs to be integrated in order to determine the most appropriate time frame for which children can accurately recall information.
- Developers and users of measures need to consider the readability of the measure.

Clarification of the relationship between child and proxy ratings

We need to accept that both child and proxy ratings have value. The question is to clarify how differences in perceptions of QoL arise between child and proxy and the implications for the child's QoL. This applies as much to clinicians as parents, teachers and other proxies. For clinicians, judgements will be influenced by their medical training, their views about this child's progress in comparison with children undergoing similar treatment, and in relation to the progress made over the course of treatment by the individual child. Judgements will also be limited to the extent that they reflect knowledge in very limited settings, usually a hospital or clinic. In contrast, parents will be influenced by the development of other children they know (their own or those of friends), their expectations and hopes for their child, additional life stresses, and their own mental health. Compared with clinicians, they will be aware of the child's QoL in a greater range of situations, and this may increase their sensitivity. Research should focus on the following.

- Determination of how proxy mental health influences ratings of the child's QoL. Beyond this it is important to clarify how parent mental health and perceptions of the child's disease influence QoL over time. This is relevant to issues concerning how parenting practices and family organisation can subsequently effect the child's QoL.
- Determination of how discrepancies arise between clinicians and parents in their perceptions of the child's QoL. Research of this kind may also be of relevance to issues of communication (and failed communication).

Acknowledgements

This research was commissioned by the NHS R&D HTA Programme. We are grateful for the comments received from the HTA referees.

We also wish to thank the following who commented on an earlier draft of this review: Professor Ronald Barr, Dr Jim Crossley, Professor Adrian Davis, Dr Adam Glaser, Dr Chris McCabe, Dr Ellen Perrin, Dr Ron Smith and Yvonne Vance. We are extremely grateful to them all for the time and effort they put into this task, and for their very helpful and insightful comments.

We would also like to thank all those who responded to our requests for information (see below), particularly to those who sent pre-publication copies of papers.

We are also grateful to Sarah Cox for her secretarial support in preparing the final manuscript.

The following individuals responded to requests for information

Professor Cliff Bailey, Leeds, St James's Hospital, UK

Professor Clare Bradley, Royal Holloway, UK

Dr Jacqueline Collier, Nottingham University, UK

Professor Joyce Cramer, Harvard University, USA

Dr Allan Colver, University of Newcastle-upon-Tyne, UK

Professor Adrian Davies, MRC Institute for Hearing Research, UK

Dr AY Finlay, University of Wales College of Medicine, UK

Dr Davina French, University of Western Australia, Australia

Dr David Goodwin, Jefferson Health System, Malvern PA, USA

Dr Peter Hoare, Edinburgh Sick Children's NHS Trust, UK

Professor Elizabeth Juniper, McMaster University, Canada

Dr Jeanne Landgraff, HealthAct, Boston, USA

Professor Alexander Quittner, University of Florida, USA

Dr Denise Reid, University of Toronto, Canada

Dr Anne Riley, Johns Hopkins University, USA

Professor Saroj Saigal, McMaster University, Canada

Dr Ulrike Ravens-Sieberer, Hamburg, Germany

Dr Sarah Smith, MRC Institute of Hearing Research, Nottingham, UK

Professor Leslie Spieth, Children's Hospital, Harvard Medical School, USA

Professor Barbara Starfield, Johns Hopkins University, USA

Dr Julie Trudel, Hospital Ste-Justine, Montreal

Dr AG Thomas, Manchester Children's Hospitals, UK

Professor James Varni, Children's Hospital and Health Center, San Diego, USA

Dr Ton Vogels, TNO Prevention and Health, The Netherlands

Dr Elizabeth Waters, University of Melbourne, Australia

References

1. Bergner M. Quality of life, health status and clinical research. *Med Care* 1989;**27**:S148–S156.
2. It's official: university makes you happier and healthier. *The Times* 2 2000 Apr 11;Pages 8–9.
3. Quality of life gets a higher profile. *The Guardian* 1998 Nov 24;Page 2.
4. Lindstrom B, Koehler L. Youth, disability and quality of life. *Pediatrician* 1991;**18**(2):121–8.
5. Nussbaum MC. The fragility of goodness. Cambridge: Cambridge University Press; 1986.
6. Jordan TE. Developing an international index of quality of life for children: The NICQL Index. *J R Soc Health* 1983;**4**:127–30.
7. UNICEF. State of the world's children. 1990.
8. Ditesheim JA, Templeton JMJ. Short-term v long-term quality of life in children following repair of high imperforate anus. *J Pediatr Surg* 1987;**22**(7):581–7.
9. Herndon DN, LeMaster J, Beard S, Bernstein N, Lewis SR, Rutan TC, *et al.* The quality of life after major thermal injury in children: an analysis of 12 survivors with greater than or equal to 80% total body, 70% third-degree burns. *J Trauma* 1986; **26**(7):609–19.
10. Henning P, Tomlinson L, Rigden SPA, Haycock GB, Chantler C. Long-term outcome of treatment of end stage renal failure. *Arch Dis Child* 1988;**63**(1):35–40.
11. Lansky LL, List MA, Lansky SB, Cohen ME, Sinks LF. Toward the development of a play performance scale for children (PPSC). *Cancer* 1985;**56**(S7):1837–40.
12. Lansky SB, List MA, Lansky LL, Ritter-Sterr C, Miller D. The measurement of performance in childhood cancer patients. *Cancer* 1987;**60**:1651–6.
13. Eisen M, Ware JE, Donald C, Brook RH. Measuring components of children's health status. *Med Care* 1979;**17**:902–21.
14. Stein REK, Jessop DJ. Functional Status II(R): a measure of child health status. *Med Care* 1990;**28**(11):1041–55.
15. Landgraf JM, Abetz L, Ware JE. The CHQ user's manual. 1st ed. Boston, MA: The Health Institute, New England Medical Center; 1996.
16. Mok JY, Laing IA, Ferguhar JW. Young diabetics: memories, current lifestyles and attitudes. *Diabetic Med* 1984;**1**(3):227–30.
17. Ahlfield JE, Soler NG, Marcus SD. The young adult with diabetes – impact of the disease on marriage and having children. *Diabetes Care* 1985;**8**(1):52–6.
18. Bullinger M, Ravens-Sieberer U. Health related QOL assessment in children: a review of the literature. *Revue Europeenne de Psychologie Appliquee* 1995;**45**(4):245–54.
19. Landgraf JM, Abetz LN. Functional status and well-being of children representing three cultural groups: initial self-reports using the CHQ-CF87. *Psychol Health* 1997;**12**(6):839–54.
20. Spieth LE, Harris C.V. Assessment of health-related quality of life in children and adolescents: an integrative review. *J Pediatr Psychol* 1996;**21**(2):175–93.
21. Eiser C. Choices in measuring quality of life in children with cancer: A comment. *Psycho-Oncology* 1995;**4**:121–31.
22. Eiser C, Jenney MM. Measuring symptomatic benefit and quality of life in paediatric oncology. *Br J Cancer* 1996;**73**:1313–16.
23. Koot HM. The study of quality of life in children and adolescents with chronic disabling conditions: concepts and methods. In: Koot HM, Wallander JL, editors. Quality of life in children and adolescents: concepts, methods and findings. UK: Harwood Academic Press. In press.
24. Vannucci RC. Current and potentially new management strategies for perinatal hypoxic-ischemic encephalopathy. *Pediatrics* 1990;**85**(6):961–8.
25. Ausejo M, Saenz A, Pham B, Kellner JD, Johnson DW, Moher D, *et al.* The effectiveness of glucocorticoids in treating croup: meta-analysis. *BMJ* 1999;**319**(7210):595–600.
26. Gortmaker S, Sappenfield W. Chronic childhood disorders: prevalence and impact. *Pediatr Clin North Am* 1984;**31**:3–18.
27. Newacheck PW, Taylor WR. Childhood chronic illness: prevalence, severity, and impact. *Am J Public Health* 1992;**82**:364–74.
28. Stiller CA. Population based survival rates for childhood cancer in Britain, 1980–1991. *BMJ* 1994;**309**:1612–16.
29. Newacheck PW, Halfon N. Prevalence and impact of disability and chronic conditions in children. *Am J Public Health* 1998;**88**:610–7.

30. Vollmer WM, Osborne ML, Buist AS. 20-year trends in the prevalence of asthma and chronic airflow obstruction in an HMO. *Am J Respir Crit Care Med* 1998;**157**:1079–84.

31. Gortmaker SL, Dietz WH, Sobol AM, Wehler CA. Increasing pediatric obesity in the United States. *Am J Dis Child* 1987;**141**:535–40.

32. Swanson JM, Lerner M, Williams L. More frequent diagnosis of attention deficit-hyperactivity disorder. *N Engl J Med* 1995;**333**:944.

33. Metcalfe MA, Baum JD. Incidents of insulin dependent diabetes in children under 15 year in the British Isles during 1988. *BMJ* 1991;**302**:443–7.

34. Engel BT. Psychosomatic medicine, behavioural medicine, just plain medicine. *Psychosom Med* 1986;**48**:466–79.

35. Hawkins MM, Stevens MCG. The long term survivors. *BMJ* 1996;**52**:898–923.

36. Brown RT. Cognitive aspects of chronic illness in children. New York: The Guilford Press; 1999.

37. Larsson B, Carlsson J. A school based, nurse administered relaxation training for children with chronic tension type head-ache. *J Child Psychol Psychiatry* 1996;**32**:821–32.

38. House of Commons Health Committee. Second, third, fourth, and fifth reports of session 1996–7. London. House of Commons Health Committee; 1997.

39. Royal College of Pediatrics and Child Health: Ethics Advisory Committee. Guidelines for the ethical conduct of medical research involving children. *Arch Dis Child* 2000;**82**:177–82.

40. Ellis R, Leventhal B. Information needs and decision making preferences of children with cancer. *Psycho-Oncology* 1993;**2**:277–84.

41. Alderson P. Will you help us with our research? *Arch Dis Child* 1995;**72**:541–2.

42. Postlethwaite RJ, Reynolds JM, Wood AJ, Evans JH, Lewis MA, Eminson DM. Recruiting patients to clinical trials: lessons from studies of growth hormone treatment in renal failure. *Arch Dis Child* 1995;**73**:30–5.

43. Weithorn LA, Campbell SB. The competency of children and adolescents to make informed treatment decisions. *Child Dev* 1982;**53**:1589–98.

44. Spilker B, Spilker B, editors. Quality of life and pharmacoeconomics in clinical trials. Philadelphia: Lippincott-Raven Publishers; 1996.

45. Feeny D, Furlong W, Mulhern RK, Barr RD, Hudson M. A framework for assessing health-related quality of life among children with cancer. *Int J Cancer* 1999;**S12**:2–9.

46. Patrick DL, Bergner M. Measurement of health status in the 1990s. *Annu Rev Public Health* 1990;**11**:165–83.

47. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality-of-life. *Ann Intern Med* 1993;**118**:622–9.

48. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. *JAMA* 1995;59–65.

49. Leidy NK. Functional status and the forward progress of merry-go-rounds: toward a coherent analytical framework. *Nurs Res* 1994;**43**:196–202.

50. Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health-status classification systems: health utilities index. *Pharmacoeconomics* 1995;**7**:490–502.

51. Starfield B. Measure of outcome: a proposed scheme. *Milbank Q* 1974;**52**:39–50.

52. Starfield B, Riley A. Profiling health and illness in children and adolescents. In: Drotar D, editor. Measuring health-related quality of life in children and adolescents: implications for research and practice. Mahwah, USA: Lawrence Erlbaum Associates, Inc; 1998.

53. Bowling A. Measuring disease. Buckingham: Open University Press; 1995.

54. Aaronson NK, Meyeravitz BE, Bard M, Bloom JR, Fauzy FI, Feldstein M, *et al.* Quality of life research in oncology: past achievements and future priorities. *Cancer* 1991;**67**(3):839–43.

55. Gotay CC, Korn EL, McCabe MS, Moore TD, Cheson BD. Quality of life assessment in cancer treatment protocols: research issues in protocol development. *J Natl Cancer Inst* 1992;**84**(575):579.

56. Ware JE. Conceptualising disease impact and treatment outcomes. *Cancer* 1984;**53**:2316–23.

57. World Health Organization. World Health Organization Constitution. Geneva: World Health Organization; 1947.

58. WHOQOL Group. World Health Organization Quality of Life Assessment (WHOQOL): position paper from the World Health Organization. *Soc Sci Med* 1995;**41**:1403.

59. Testa MA, Simonson DC. Assessment of quality-of-life outcomes. *N Engl J Med* 1996;**334**(13):835–40.

60. Schipper H, Clinch JJ, Olweny CLM. Quality of life studies: definitions and conceptual frameworks. In: Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials. 2nd ed. Philadelphia: Lippincott-Raven Publishers; 1996.

61. Cummins RA. Self-rated quality of life scales for people with an intellectual disability: a review. *Appl Res Intellect Disabil* 1995;**10**(3):199–216.

62. Leplege A, Hunt S. The problem of quality of life in medicine. *JAMA* 1997;**278**(1):47–50.

63. Calman KC. Definitions and dimensions of quality of life. In: Aaronson NK, Beckmann J, editors. The quality of life of cancer patients. New York: Raven Press; 1987.

64. Hayry M. Measuring quality of life: Why, how and what? In: Joyce CRB, O'Boyle CA, McGee H, editors. Individual quality of life: approaches to conceptualisation and assessment. Australia: Harwood Academic Publishers; 1999.

65. Gill TM, Feinstein AR. A critical appraisal of the quality of quality-of-life measurements. *JAMA* 1994;**272**:619–26.

66. Smith KW, Avis NE, Assmann SF. Distinguishing between quality of life and health status in quality of life research: a meta-anlysis. *Qual Life Res* 1999;**8**(5):447–59.

67. Guyatt GH, Deyo RA, Charlson M, Levine MN, Mitchell A. Responsiveness and validity in health-status measurement: a clarification. *J Clin Epidemiol* 1989;**42**(5):403–8.

68. Deyo RA, Patrick DL. Barriers to the use of health status measures in clinical investigation, patient care and policy research. *Med Care* 1989;**27**(S3):254–68.

69. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998;**2**(14).

70. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989;**27**(3):S127–S232

71. Fallowfield L. An overview of quality of life measurements. In: Trimble MR, Dodson WE, editors. Epilepsy and quality of life. New York: Raven Press; 1994.

72. O'Connor R. Issues in the measurement of health-related quality of life. Working Paper 30, NHMRC National Centre for Health Program Evaluation. Melbourne, Australia; 1993.

73. Bergner M, Bobbit RA, Carter SR, Gilson BS. The Sickness Impact Profile: development and final revision of a health status measure. *Med Care* 1981;**19**(8):787–805.

74. Plant P, McEwen J, Prescott K. Use of the Nottingham Health Profile to test the validity of census variables to proxy the need for health care. *J Public Health Med* 1996;**18**(3):313–20.

75. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): conceptual framework and items selection. *Med Care* 1992;**30**:473–83.

76. von Neumann J, Morgenstern O. Theory of games and economic behaviour. Princetown, New Jersey: Princetown University Press; 1944.

77. Feeny DH, Torrance GW. Incorporating utility-based quality-of-life assessment measures in clinical-trials – 2 examples. *Med Care* 1989;**27**:S190–S204.

78. Feeny D, Furlong W, Barr RD, Torrance GW, Rosenbaum P, Weitzman S. A comprehensive multiattribute system for classifying the health status of survivors of childhood cancer. *J Clin Oncol* 1992;**10**(6):923–8.

79. Boyle MH, Offord DR, Hoffman HG, Catlin GP, Byles JA, Cadman DT, *et al.* Ontario Child Health Study: methodology. *Arch Gen Psychiatry* 1987;**44**:826–31.

80. Juniper EF, Guyatt GH, Feeny DH, Griffith LE, Ferrie PJ. Minimum skills required by children to complete health- related quality of life instruments for asthma: comparison of measurement properties. *Eur Respir J* 1997;**10**(10):2285–94.

81. Patrick DL, Erickson P. Health status and health policy. Oxford: Oxford University Press; 1993.

82. Patrick DL, Erickson P. Assessing health-related quality of life for clinical decision making. In: Walker SR, Lancaster RM, editors. Quality of life: assessment and application. England: MTP Press; 1988.

83. Bradlyn AS, Ritchey AK, Harris CV, Moore IM, O'Brien RT, Parsons SK, *et al.* Quality of life research in pediatric oncology. *Cancer* 1996;**78**(6):1333–9.

84. Eiser C. Children's quality of life measures. *Arch Dis Child* 1997;**77**:350–4.

85. Mulhern RK, Horowitz ME, Ochs J, Friedman AG. Assessment of quality of life among pediatric patients with cancer. *Psychol Assess* 1989;**1**(2):130–48.

86. Jenney ME, Kane RL, Lurie N. Developing a measure of health outcomes in survivors of childhood cancer: a review of the issues. *Med Pediatr Oncol* 1995;**24**(3):145–53.

87. WHO/IACAPAP. Report of a WHO/IACAPAP Working Party on the Measurement of Quality of Life in Children; 1993 Jun 3–4; London, UK.

88. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;**16**:297–334.

89. Nunnally J. Psychometric theory. 2nd ed. New York: McGraw-Hill; 1978.

90. Bowling A. Measuring health: a review of quality of life measurement scales. 2nd ed. Buckingham: Open University Press; 1997.

91. Bradley C. Measuring quality of life in diabetes. In: Marshall SM, Home PD, Rizza RA. editors. The Diabetes Annual/10. Elsevier Science; 1996.

92. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull* 1968;**70**:213–20.

93. Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychol Bull* 1955;**52**:281–302.

94. Armstrong FD, Toledano SR, Miloslavich K, Lackman-Zeman L, Levy JD, Gay CL, *et al.* The Miami Pediatric Quality of Life Questionnaire: Parent Scale. *Int J Cancer* 1999;**S12**:11–17

95. Fayers PM, Hand DJ, Bjordal K, Groenvold M. Causal indicators in quality of life research. *Qual Life Res* 1997;**6**:393–406.

96. Hays RD, Anderson R, Revicki D. Psychometric considerations in evaluating health-related quality of life measures. *Qual Life Res* 1993;**2**:441–9.

97. Wechsler D. Wechsler Intelligence Scale for Children – Revised. New York: Psychological Corporation; 1974.

98. Kline P. The handbook of psychological testing. London: Routledge; 1992.

99. Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol Bull* 1959;**47**:81–105.

100. Hays RD, Hadorn DC. Responsiveness to change: an aspect of validity, not a separate dimension. *Qual Life Res* 1992;**1**:73–5.

101. Oppenheim AN. Questionnaire design, interviewing and attitude measurement. London: Pinter Publishers; 1992.

102. Revicki DA, Kaplan RM. Relationship between psychometric and utility-based approaches to the measurement of health-related quality of life. *Qual Life Res* 1993;**2**:477–87.

103. Rosenbaum PL, Saigal S. Measuring health-related quality of life in pediatric populations: conceptual issues. In: Spilker B, editor. Quality of life and phamacoeconomics in clinical trials. Philadelphia: Lippincott-Raven Publishers; 1996.

104. Binger C, Ablin A, Feverstein R, Kushner J, Zoger S, Mikkelson C. Childhood leukemia: emotional impact on patient and family. *N Engl J Med* 1969;**280**:414–18.

105. Pless IB, Pinkerton P. Chronic childhood disorder: promoting patterns of adjustment. London: Henry Kimpton; 1975.

106. Editorial. Age limits of pediatrics. *Pediatrics* 1972;**49**:463.

107. Riley AW, Forrest CB, Starfield B, Green B, Kang M, Ensminger M. Reliability and validity of the adolescent health profile-types. *Med Care* 1998;**36**(8):1237–48.

108. Riley AW, Green BF, Forrest CB, Kang M, Ensminger ME. A taxonomy of adolescent health: development of the adolescent health profile-types. *Med Care* 1998;**26**(8):1228–36.

109. Starfield B, Bergner M, Ensminger M, Riley A, Ryan S, Green B, *et al.* Adolescent health status measurement: development of the Child Health and Illness Profile. *Pediatrics* 1993;**91**(2):430–5.

110. Starfield B, Riley AW, Green BF, Ensminger ME, Ryan SA, Kelleher K, *et al.* The adolescent child health and illness profile. a population-based measure of health. *Med Care* 1995;**18**(4):465–72.

111. Starfield B, Forrest CB, Ryan SA, Riley AW, Ensminger ME, Green BF. Health Statistics of well vs ill adolescents. *Arch Pediatr Adolesc Med* 1996;**150**(12):1249–56.

112. Landgraf JM, Maunsell E, Speechley KN, Bullinger M, Campbell S, Abetz L, *et al.* Canadian-French, German and UK versions of the Child Health Questionnaire: methodology and preliminary item scaling results. *Qual Life Res* 1998;**7**(5):433–45.

113. Graham P, Stevenson J, Flynn D. A new measure of health-related quality of life for children: preliminary findings. *Psychol Health* 1997;**12**(5):655–65.

114. Wasson JH, Kairys SW, Nelson EC, Kalishman N, Baribeau P. A short survey for assessing health and social problems of adolescents. *J Fam Pract* 1994; **38**(5):489–94.

115. Wasson JH, Kairys SW, Nelson EC. Adolescent health and social problems: a method for detection and early management. *Arch Fam Med* 1995;**4**:51–6.

116. Eiser C, Vance YH, Seamark D. The development of a theoretically driven measure of quality of life for children aged 6–11 years. *Child Care Health Dev.* 2000;**26**:445–56.

117. Eiser C, Cotter I, Oades P, Seamark D, Smith R. Health-related quality-of-life measures for children. *Int J Cancer* 1999;**S12**: 87–90.

118. Collier J. Developing a generic child quality of life questionnaire. *The British Psychological Society, Health Psychology Update* 1997;**28**:12–16.

119. Bruil J. Development of a quality of life instrument for children with chronic illness. Ledien, The Netherlands: Health Psychology, Leiden University; 1999.

120. Ravens-Siberer U, Theiling S, Bullinger M. Assessing quality of life in chronically ill children – the parents and the parents' view. 11th International Congress of the European Society for Child and Adolescent Psychiatry; 1999 Sep 15–19; Congress Centrum Hamburg; **8**(S2).

121. Lindstrom B, Eriksson B. Quality of life among children in Nordic countries. *Qual Life Res* 1993;**2**(1):23–32.

122. Lindstrom B. Quality of life for children and disabled children based on health as a resource concept. *J Epidemiol Community Health* 1994;**48**(6):529–30.

123. Varni JW, Seid M, Rode CA. The PedsQL: measurement model for the pediatric quality of life inventory. *Med Care* 1999;**37**(2):126–39.

124. Varni JW, Seid M, Jacobs JR, Rode CA. The PedsQL: II. Development of the asthma and arthritis disease-specific modules for the Pediatric Quality of Life Inventory. Submitted.

125. Eiser C, Kopel S, Cool P, Grimer R. The perceived illness experience scale (PIE): reliability and validity revisited. *Child Care Health Dev* 1999;**25**(3):179–90.

126. Eiser C, Havermans T, Craft A, Kernahan J. Development of a measure to assess the perceived illness experience after treatment for cancer. *Arch Dis Child* 1995;**72**(4):302–7.

127. Raphael D, Rukholm E, Brown I, Hill-Bailey P. The quality of life profile – adolescent version: background, description and initial validation. *J Adolesc Health* 1996;**19**(5):366–75.

128. Iorio R, Pensati P, Botta S, Moschella S, Impagliazzo N, Vajro P, *et al.* Side effects of alpha-interferon therapy and impact on health-related quality of life in children with chronic viral hepatitis. *Pediatr J Infect Dis* 1997;**16**(10):984–90.

129. Theunissen NC, Vogels T, Koopman HM, Verrips GH, Zwinderman K, Verloove-Vanhorick SP, *et al.* The proxy problem: child report versus parent report in health-related quality of life research. *Qual Life Res* 1998;**7**(5):387–97.

130. Vogels T, Verrips GH, Verloove-Vanhorick SP, Fekkes M, Kamphuis RP, Koopman HM, *et al.* Measuring health-related quality of life in children: the development of the TACQOL parent form. *Qual Life Res* 1998;**7**(5):457–65.

131. Spencer NJ, Coe C. The development and validation of a measure of parent-reported child health and morbidity: The Warwick Child Health and Morbidity Profile. *Child Care Health Dev* 1996;**22**(6):367–79.

132. Boyle MH, Furlong W, Feeny D, Torrance GW, Hatcher J. Reliability of the Health Utilities Index Mark III used in the 1991 cycle 6 Canadian General Social Survey Health Questionnaire. *Qual Life Res* 1995;**4**:249–57.

133. Apajasalo M, Sintonen C, Holmberg C, Sinkkonen J. Quality of life in early adolescence: a sixteen-dimensional health-related measure (16D). *Qual Life Res* 1996;**5**(2):205–11.

134. Apajasalo M, Rautonen J, Holmberg C, Sinkkonen J. Quality of life in pre-adolescence: a 17-dimensional health-related measure (17D). *Qual Life Res* 1996;**5**(6):532–8.

135. Bradlyn AS, Harris CV, Warner JE, Ritchey AK, Zaboy K. An investigation of the validity of the quality of Well-Being Scale with pediatric oncology patients. *Health Psychol* 1993;**12**(3):246–50.

136. Mishoe SC, Baker RR, Poole S, Harrell LM, Avant CB, Rupp NT. Development of an instrument to assess stress levels and quality of life in children with asthma. *J Asthma* 1998;**35**(7):553–63.

137. Gibson PG, Henry RL, Vimpani GV, Halliday J. Asthma knowledge, attitudes, and quality of life in adolescents. *Arch Dis Child* 1995;**73**(4):321–6.

138. Christie MJ, French D, Sowden A, West A. Development of child-centered disease-specific questionnaires for living with asthma. *Psychosom Med* 1993;**55**(6):541–8.

139. French DJ, Christie MJ, Sowden AJ. The reproducibility of the Childhood Asthma Questionnaires. *Qual Life Res* 1994;**3**(3):215–24.

140. Juniper EF, Guyatt GH, Feeny DH, Ferrie PJ, Griffith LE, Townsend M. Measuring quality of life in children with asthma. *Qual Life Res* 1996;**5**(1):35–46.

141. Phipps S, Hinds PS, Channell S, Bell GL. Measurement of behavioural, affective, and somatic responses to pediatric bone marrow transplantation: development of the BASES scale. *J Pediatr Oncol Nurs* 1994;**11**(3):109–17.

142. Phipps S, Dunavant M, Jayawardene D, Srivastiva DK. Assessment of health related quality of life in acute in-patient settings: use of the BASES Scales in children undergoing bone marrow transplantation. *Int J Cancer* 1999;**S12**:18–24.

143. Varni JW, Katz ER, Quiggins DJL, Friedman-Bender A. The Pediatric Cancer Quality of Life Inventory-32 (PCQL-32): I. Reliability and validity. *Cancer* 1998;**82**(6):1184–96.

144. Varni JW, Katz ER, Seid M, Quiggins DJ, Friedman-Bender A, Castro CM. The Pediatric Cancer Quality of Life Inventory (PCQL). I. Instrument development, descriptive statistics, and cross-informant variance. *J Behav Med* 1998;**21**(2):179–204.

145. Goodwin DA, Boggs SR, Graham-Pole J. Development and validation of the Pediatric Oncology Quality of Life Scale. *Psychol Assess* 1994;**6**(4):321–8.

146. Hoare P, Russell M. The quality of life of children with chronic epilepsy and their families: preliminary findings with a new assessment measure. *Dev Med Child Neurol* 1995;**37**(8):689–96.

147. Wildrick D, Parker-Fisher S, Morales A. Quality of life in children with well-controlled epilepsy. *J Neurosci Nurs* 1996;**28**:192–8.

148. Keene DL, Higgins MJ, Ventureyra EC. Outcome and life prospects after surgical management of medically intractable epilepsy in patients under 18 years of age. *Childs Nerv Syst* 1997;**13**(10):530–5.

149. Cramer JA, Westbrook L, Devinsky O, Perrine K, Glassman M, Camfield C. Development of a quality of life inventory for adolescents: The QOLIE-AD-48. *Epilepsia* 1999;**40**(8):1114–21.

150. Duffy CM, Duffy KN. Health assessment in the rheumatic diseases of childhood. *Curr Opin Rheumatol* 1997;**9**(5):440–7.

151. Rabbett H, Elbadri A, Thwaites R, Northover H, Dady I, Firth D, *et al.* Quality of life in children with Crohn's disease. *J Pediatr Gastroenterol Nutr* 1996;**23**(5):528–33.

152. Lewis-Jones MS, Finlay AY. The Children's Dermatology Life Quality Index (CDLQI): initial validation and practical use. *Br J Dermatol* 1995;**132**(6):942–9.

153. Ingersoll GM, Marrero DG. A modified quality-of-life measure for youths: psychometric properties. *Diabetes Educucation* 1991;**17**(2):114–18.

154. Langeveld JH, Koot HM, Passchier J. Headache intensity and quality of life in adolescents: how are changes in headache intensity in adolescents related to changes in experienced quality of life? *Headache* 1997;**37**(1):37–42.

155. Langeveld JH, Koot HM, Passchier J. Do experienced stress and trait negative affectivity moderate the relationship between headache and quality of life in adolescents? *J Pediatr Psychol* 1999;**24**(1):1–11.

156. Reid DT, Renwick RM. Preliminary validation of a new instrument to measure life satisfaction in adolescents with neuromuscular disorders. *Int J Rehabil Res* 1994;**17**(2):184–8.

157. Rosenfeld RM, Goldsmith AJ, Tetlus L, Balzano A. Quality of life for children with otitis media. *Arch Otolaryngol Head Neck Surg* 1997;**123**(10):1049–54.

158. Juniper EF, Guyatt GH, Dolovich J. Assessment of quality-of-life in adolescents with allergic rhinoconjunctivitis – development and testing of a questionnaire for clinical-trials. *J Allergy Clin Immunol* 1994;**93**:413–23.

159. Pilpel D, Leiberman E, Zadik Z, Carel CA. Effect of growth hormone treatment on quality of life of short-stature children. *Horm Res* 1995;**44**:1–5.

160. Parkin PC, Kirpalani HM, Rosenbaum PL, Fehlings DL, Van Nie A, Willan AR, *et al.* Development of a health-related quality of life instrument for use in children with spina bifida. *Qual Life Res* 1997;**6**(2):123–32.

161. Climent JM, Reig A, Sanchez J, Roda C. Construction and validation of a specific quality of life instrument for adolescents with spine deformities. *Spine* 1995;**20**(18):2006–11.

162. Kaplan RM. Implications of quality of life assessment in public policy for adolescent health. In: Drotar D, editor. Measuring health-related quality of life in children and adolescents. New Jersey: Lawrence Erlbaum Associates; 1998.

163. Munzenberger PJ, Van Wagnen CA, Abdulhamid I, Walker PC. Quality of life as a treatment outcome in patients with cystic fibrosis. *Pharmacotherapy* 1999;**19**(4):393–8.

164. Czyzewski DI, Bartholomew L, Kay. Quality of life outcomes in children and adolescents with cystic fibrosis. In: Drotar D, editor. Measuring health-related quality of life in children and adolescents: implications for research and practice. Mahwah, USA: Lawrence Erlbaum Associates; 1998.

165. Czyzewski DI, Mariotto MJ, Bartholomew LK, LeCompte SH, Sockrider MM. Measurement of quality of well being in a child and adolescent cystic fibrosis population. *Med Care* 1994;**32**(9):965–72.

166. Kaplan RM, Anderson JP. The general health policy model: Update and applications. *Health Serv Res* 1988;**23**:203–35.

167. Deyo RA, Inui TS, Leininger J, Overman SS. Psychosocial function in rheumatoid-arthritis-clinical use of a self administered health-status instrument. *Arch Intern Med* 1982;**142**(5):879–82.

168. Deyo RA, Inui TS, Leininger JD, Overman SS. Measuring functional outcomes in chronic disease: a comparison of traditional scales and a self-administered health status questionnaire in patients with rheumatoid arthritis. *Med Care* 1983;**21**:180–92.

169. Nelson EC, Wasson JH, Johnson DJ, Hays R. Dartmouth COOP Functional Health Assessment Charts: brief measures for clinical practice. In: Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials. Philadelphia: Lippincott-Raven Publishers; 1996. p. 161–8

170. Sintonen H, Pekurinen M. A fifteen-dimensional measure of health-related quality of life (15D) and its applications. In: Walker SR, Rosser RM, editors. Quality of life assessment: key issues in the 1990s. Dordrecht: Kluver Academic Publishers; 1993.

171. Juniper EF, Guyatt GH, Ferrie PJ, Griffith LE. Measuring quality of life in asthma. *Am Rev Respir Dis* 1993;**147**:832–8.

172. Karnofsy DA, Burchenal JH. The clinical evaluation of chemotherapeutics in cancer. In: MacLeod CM, editor. Evaluation of chemotherapeutic agents. New York: Columbia University Press; 1949.

173. Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI) – a simple practical measure for routine clinical use. *Clin Exp Dermatol* 1994;**19**:210–16.

174. The DCCT Research Group. Reliability and validity of a diabetes quality-of-life measure for the diabetes control and complications trial (DCCT). *Diabetes Care* 1988;**11**:725–32.

175. Devinsky O, Vickrey BG, Cramer JA, Hermann B, Meador K, Hays RD. Development of an instrument of health-related quality of life for people with epilepsy. *Epilepsia* 1995;**36**:1089–104.

176. Juniper EF, Guyatt GH. Development and testing of a new measure of health status for clinical trials in rhinoconjunctivitis. *Clin Exp Allergy* 1991;**21**:77–83.

177. Juniper EF, Howland WC, Roberts NB, Thompson AK, King DR. Measuring quality of life in children with rhinoconjunctivitis. *J Allergy Clin Immunol* 1998;**101**:163–70.

178. Kaplan RM, Bush JW. Health-related quality of life measurement for evaluation research and policy analysis. *Health Psychol* 1982;**1**:61–80.

179. Fanshel S, Bush JW. A health status index and its application to health services outcomes. *Operational Res* 1970;**18**:1021–65.

180. Kaplan RM, Bush JW, Berry CC. Health status: Types of validity for an index of well-being. *Health Serv Res* 1976;**11**:478–507.

181. Bergner M. Development, testing and use of the Sickness Impact Profile. In: Walker SR, Rosser RM, editors. Quality of life assessment: key issues in the 1990s. Lancaster: MIT Press; 1988.

182. Deyo RA, Diehl AK. Measuring physical and psychosocial function in patients with lower back pain. *Spine* 1983;**8**(6):635–42.

183. Sullivan M, Ahlmen M, Bjelle A. Health status assessment in rheumatiod arthritus. I. Further work on the validity of the Sickness Impact Profile. *J Rheumatol* 1990;**17**(4):439–47.

184. Longstreth WT Jr, Nelson L, Linde M, Munoz D. Utility of the Sickness Impact Profile in Parkinson's disease. *J Geriatr Psychiatr Neurol* 1992;**5**(3):142–8.

185. Weinberger M, Samsa GA, Tierney WM, Belyea MJ, Hiner SL. Generic versus disease specific health status measures: comparing the Sickness Impact Profile and the Arthritis Impact Measurement Scales. *J Rheumatol* 1992;**19**(4):543–6.

186. Smith HE. Head injury follow-up: Is the Sickness Impact Profile a useful clinical tool? *Clin Rehabil* 1992;**6**(1):31–9.

187. Vickrey BG, Perrine KR, Hays RD. QOLIE-89 (version 1.0) Scoring Manual. Santa Monica, CA: RAND; 1993.

188. Long AF, Dixon P. Monitoring outcomes in routine practice: defining appropriate measurement criteria. *J Eval Clin Pract* 1996;**2**(1):71–8.

189. Guttmann-Bauman I, Flaherty BP, Strugger M, McEvoy RC. Metabolic control and quality-of-life self-assessment in adolescents with IDDM. *Diabetes Care* 1998;**21**(6):915–18.

190. Jacobson AM, De Groot M, Samson J. The effects of psychiatric disorders and symptoms on quality of life in patients with type 1 and type 2 diabetes. *Diabetes Care* 1994;**17**:268–74.

191. Aynsley-Green A, Barker M, Burr S, Macfarlane A, Morgan J, Silbert J, *et al.* Education and debate: who is speaking for children and adolescents and for their health at the policy level? *BMJ* 2000;**321**:229–32.

192. Graham P. Depression in pre-pubertal children. *Dev Med Child Neurol* 1974;**16**(3):340–9.

193. Kazdin AE. Childhood depression. *J Child Psychol Psychiatry* 1990;**31**(1):121–60.

194. Carpay HA, Arts WFM. Outcome assessment in epilepsy: available rating scales for adults and methodological issues pertaining to the development of scales for childhood epilepsy. *Epilepsy Res* 1996;**24:**127–36.

195. Beckung E, Uvebrant P. Motor and sensory impairments in children with intractable epilepsy. *Epilepsia* 1993;**34**:924–9.

196. Huttenlocher PR, Hapke RJ. A follow-up study of intractable seizures in childhood. *Ann Neurol* 1990;**28**:699–705.

197. Wallander J, Varni JW. Effects of pediatric chronic physical disorders on child and family adjustment. *J Child Psychol Psychiatry* 1998;**39**(1):29–46.

198. Rosenbaum P, Cadman D, Kirpalani H. Pediatrics: assessing quality of life. In: Spilker B, editor. Quality of life assessments in clinical trials. New York: Raven Press; 1990.

199. Garrison WT, McQuiston S. Chronic illness during childhood and adolescence: psychological aspects. Developmental Clinical Psychology and Psychiatry Series. Vol. 19. Newbury park, CA: Sage publications; 1989.

200. Kazak AE. Families of chronically ill children: a systems and social-ecological model of adaption and challenge. *J Consult Clin Psychol* 1989;**57**:25–30.

201. Shalet SM, Gibson B, Swindell R. Effect of spinal irradiation on growth. *Arch Dis Child* 1987;**62**(5):461–4.

202. McGee HM. Quality of life: assessment issues for children with chronic illness and their families. In: Christie MJ, French DJ, editors. Assessment of quality of life in children with asthma. Langhorne, PA: Harwood Academic Publishers; 1994.

203. Eiser C, Kopel SJ. Children's perceptions of health and illness. In: Petrie KJ, Weinman JA, editors. Perceptions of health and illness: current research and applications. Singapore: Harwood Academic Publishers; 1997.

204. Bibace R, Walsh ME. R.Bibace, Children's conceptions of illness. In: Walsh ME, editor. New directions for child development: No 14 Children's conceptions of health, illness and bodily functions. San Fransisco: Jossey-Bass; 1981.

205. Harbeck C, Peterson L. Elephants dancing in my head – a developmental approach to children's concepts of specific pains. *Child Dev* 1992;**63**(1):138–49.

206. Berry SL, Hayford JR, Ross CK, Pachman LM, Lavigne JV. Conceptions of illness by children with juvenile rheumatiod arthritis: a cognitive developmental approach. *J Pediatr Psychol* 1993;**18**:83–97.

207. Sayer AG, Willett JB, Perrin EC. Measuring understanding of illness causality in healthy children and in children with chronic illness: a construct validation. *J Appl Dev Psychol* 1993;**14**:11–36.

208. Meadows AT, McKee L, Kazak AE. Psychosocial status of young-adult survivors of childhood cancer – a survey. *Med Pediatr Oncol* 1989;**17**(6):466–70.

209. Cummins RA. On the trail of the gold standard for subjective well-being. *Social Indicators Research* 1995;**35**(2):179–200.

210. Borthwick-Duffy SA. Quality of life of persons with severe or profound mental retardation. In: Schalock RL, Begab MJ, editors. Quality of life: perspectives and issues. Washington DC: American Association on Mental Retardation; 1990.

211. Barr RD, Pai MR, Weitzman S, Feeny D, Furlong W, Rosenbaum P, *et al.* A multiattributional approach to health-status measurement and clinical management illustrated by an application to brain-tumors in childhood. *Int J Oncol* 1994;**4**:639–48.

212. Spirito A, deLawyer DD, Stark LJ. Peer relations and social adjustment of chronically ill children and adolescents. *Clin Psychol Rev* 1991;**11**:539–64.

213. Bernt TJ, Hoyle SG. Stability and change in child and adolescent friendships. *Dev Psychol* 1985;**21**(6):1007–15.

214. Rubin KH, Chen X, McDougall P, Bowker A, McKinnon J. The Waterloo Longitudinal Project: predicting adolescent internalizing and externalizing problems from early and mid-childhood. *Dev Psychopathol* 1995;**7**:751–64.

215. Noeke-Hoeksema S, Girgus JS. The emergence of gender differences in depression during adolescence. *Psychol Bull* 1994;**115**: 424–43.

216. Mulhern RK. Neuropsychological late effects. In: Bearison DJ, Mulhern RK, editors. Pediatric psycho-oncology: psychological perspectives on children with cancer. New York: Oxford Press; 1994.

217. Rovet JF, Ehrlich RM, Hoppe M. Intellectual deficits associated with early onset of insulin dependent diabetes mellitus in children. *Diabetes Care* 1987;**10**(4): 510–15.

218. Lemanek KL, Hood C. Asthma. In: Brown RT, editor. Cognitive aspects of chronic illness in children. New York: The Guilford Press; 1999.

219. Bender BG. Are asthmatic children educationally handicapped? *School Psychol Q* 1995;**10**:274–91.

220. Wolters PL, Brouwers P, Perez LA. Pediatric HIV infection. In: Brown RT, editor. Cognitive aspects of chronic illness in children. New York: The Guilford Press; 1999.

221. Frank NC, Allison SM, Cant MEC. Sickle cell disease In: Brown RT, editor. Cognitive aspects of chronic disease in children. New York: The Guilford Press; 1999.

222. Delamater AM, Brady EU, Blumberg MJ. Cardiac conditions. In: Brown RT, editor. Cognitive aspects of chronic illnesses in children. New York: The Guilford Press; 1999.

223. Heimentz JR, Hynd GW, Jimenz M. Seizure disorders. In: Brown RT, editor. Cognitive aspects of chronic illnesses in children. New York: The Guilford Press; 1999.

224. Erickson EH. Identity and the life-cycle. *Psychol Issues* 1959;**1**(18):164.

225. Hopwood P. The assessment of body image in cancer patients. *Eur J Cancer* 1992;**29A**:276–81.

226. Kopel SJ, Eiser C, Cool P, Grimer RJ, Carter SR. Brief report: assessment of body image in survivors of childhood cancer. *J Pediatr Psychol* 1998;**23**:141–7.

227. Elkind D. Egocentricism in adolescence. *Child Dev* 1967;**38**:1025–34.

228. Smith C. Smoking among young people: some recent developments in Wales. *Health Education Journal* 1991;**50**:8–11.

229. Walker LS, Garber J, Greene JW. Somatic complaints in pediatric patients: a prospective study of the role of negative life events, child, social, and academic competence, and parental somatic symptoms. *J Consult Clin Psychol* 1994;**62**:1213–21.

230. Blount RL, Landolf-Fritshe B, Powers SW, Sturges JW. Differences between high and low coping children and between parent and staff behaviors during painful medical procedures. *J Pediatr Psychol* 1991;**16**(6):795–809.

231. Manne SL, Jacobsen PB, Gorfinkle K, Gerstein F, Redd WH. Treatment adherence difficulties among children with cancer: the role of parenting style. *J Pediatr Psychol* 1993;**18**:47–62.

232. Simopoulos AP. Diet and health: scientific concepts and principles. *Am J Clin Nutr* 1987;**45**:1015–410.

233. Armstrong N. Children's physical activity patterns and coronary heart disease. In: Coronary Prevention Group. Should the prevention of heart disease begin in childhood? London: Coronary Prevention Group; 1989.

234. Stone WL, Lemanek KL. Developmental issues in children's self-reports. In: La Greca A, editor. Through the eyes of the child: obtaining self-reports from children and adolescents. Boston: Allyn and Bacon; 1990.

235. Titman P, Smith M, Graham P. Assessment of the quality of life of children. *Clin Child Psychol Psychiatry* 1997;**2**(4):597–606.

236. Flesch R. A new readability yardstick. *J Appl Psychol* 1948;**32**:221–33.

237. Siegal M, Patty J, Eiser C. A re-examination of children's conceptions of contagion. *Psychol Health* 1990;**4**:159–65.

238. Pantell RH, Lewis CC. Measuring the impact of medical care on children. *J Chronic Dis* 1987;**40**(S1):99–108.

239. West A. Methodological issues in the assessment of quality of life in childhood asthma: what educational research has to offer. In: Christie MJ, French DJ, editors. Assessment of quality of life in children with asthma. Langhorne, PA: Harwood Academic Publishers; 1994.

240. Torrance GW. Measurement of health state utilities for economic appraisal: a review. *J Health Econ* 1986;**5**:1–30.

241. Stein REK, Reissman CK. Development of an impact-on-family scale: preliminary findings. *Med Care* 1980;**18**(4):465–72.

242. Juniper EF, Guyatt G, Willan AR, Griffith LE. Determining a minimal important change in a disease-specific quality of life instrument. *J Clin Epidemiol* 1994;**47**:81–7.

243. Ridgeway D, Waters E, Kuczaj SA. Acquisition of emotion-descriptive language: receptive and productive vocabulary norms for ages 18 months to 6 years. *Dev Psychol* 1985;**21**:901–8.

244. Harris PL. Children's understanding of the link between situation and emotion. *J Exp Child Psychol* 1983;**36**:490–509.

245. Harris PL, Lipian MS. Saarni C, Understanding emotion and experiencing emotion. In: Harris PL, editor. Children's understanding of emotion. Cambridge: Cambridge University Press; 1989.

246. McGrath P, McAlpine LM. Psychological perspectives on pediatric pain. *J Pediatr* 1993;**122**:52–8.

247. Harter S, Pike K. The Pictorial Scale of Perceived Competence and Social Acceptance for young children. *Child Dev* 1984;**48**(5):80–7.

248. French D, Christie M. Developing outcome measures for children: "Quality of life" assessment for pediatric asthma. In: Huchinson A, McColl E, Christie M, Riccalton C, editors. Health outcome measures in primary and out-patient care. Australia: Harwood Academic Publishers; 1996.

249. Friedman WJ. Children's representation of the pattern of daily activities. *Child Dev* 1990;**61**:1399–412.

250. Ginsberg HP. Entering the child's mind: the clinical interview in psychological research and practice. New York: Cambridge University Press; 1997.

251. Mohay H. Interviewing children and adolescents. In: Sanders M, Mitchell C, Byrne G, editors. Medical consultation skills: behavioral and interpersonal dimensions of health care. Melbourne: Addison Wesley; 1997.

252. Mize J, Ladd GW. Predicting preschoolers peer behaviour and status from their interpersonal strategies – a comparison of verbal and enactive responses to hypothetical social dilemmas. *Dev Psychol* 1988;**24**(6):782–8.

253. Sprangers MAG, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: a review. *J Clin Epidemiol* 1992;**45**:743–60.

254. Camp LD, O'Sullivan PS. Comparison of medical, surgical, and oncology patients' descriptions of pain and nurses' documentation of pain assessments. *J Adv Nurs* 1987;**12**:593–8.

255. Teske K, Daut RL, Cleeland CS. Relationships between nurses' observations and patients' self-reports of pain. *Pain* 1983;**16**(3):289–96.

256. Derogatis LR, Abeloff MD, McBeth CD. Cancer patients and their physicians in the perception of psychological symptoms. *Psychosomatics* 1976;**17**:197–201.

257. Jachuck SJ, Brierley H, Jachuck S, Willcox PM. The effects of hypotensive drugs on the quality of life. *J R Coll Gen Pract* 1982;**32**:103–5.

258. Epstein AM, Hall JA, Tognetti J, Son LH, Conant L. Using proxies to evaluate quality of life. Can they provide valid information about patients' health status and satisfaction with medical care? *Med Care* 1989;**27**:S91–S98.

259. Sneeuw KCA, Aaronson NK, Sprangers MAG, Detmar SB, Wever LD, Schornagel JH. Comparison of patient and proxy EORTC QLQ-C30 ratings in assessing the quality of life of cancer patients. *J Clin Epidemiol* 1998;**51**(7):617–31.

260. Achenbach TM, McConaughy SH, Howell CT. Child/adolescent behavioural and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychol Bull* 1987;**101**(2):213–32.

261. Ennett ST, DeVellis BM, Earp JA, Kredich D, Warren RW, Wilhelm CL. Disease experience and psychosocial adjustment in children with juvenile rheumatoid arthritis: children's versus mothers' reports. *J Pediatr Psychol* 1991;**16**(5):557–68.

262. Fivush R, Hamond NR, Harsch N, Singer A, Wolf A. Content and consistency of young children's autobiographical recall. *Discourse Processes* 1991;**14**:372–88.

263. Walker LS, Heflinger CA. Quality of life predictors of outcome in pediatric abdominal pain patients: findings at initial assessment and 5-year follow-up. In: Drotar D, editor. Measuring health-related quality of life in children and adolescents: implications for research and practice. Mahwah, USA: Lawrence Erlbaum; 1998.

264. Glaser AW, Davies K, Walker D, Brazier D. Influence of proxy respondents and mode of administration on health status assessment following central nervous system tumours in childhood. *Qual Life Res* 1997;**6**(1):43–53.

265. Glaser AW, Abdul RN, CL U, Walker DA. School behaviour and health status after central nervous system tumours in childhood. *Br J Cancer* 1997;**76**(5):643–50.

266. Billson AL, Walker DA. Assessment of health status in survivors of cancer. *Arch Dis Child* 1994;**70**(3):200–4.

267. Parsons SK, Barlow SE, Levy SL, Supran SE, Kaplan SH. Health-related quality of life in pediatric bone marrow transplant survivors: according to whom? *Int J Cancer* 1999;**S12**:46–51.

268. French DJ, Christie MJ, West A. Quality of life in childhood asthma: development of the childhood asthma questionnaires. In: Christie MJ, French DJ, editors. The assessment of quality of life in children with asthma. Chur: Harwood; 1994.

269. Gortmaker SL, Lenderking WR, Clark C, Lee S, Fowler MG, Oleske JM and the ACTG 219 Team. Development and use of a pediatric quality of life questionnaire in AIDS clinical trials: reliability and validity of the General Health Assessment for Children. In: Drotar D, editor. Measuring health-related quality of life in children and adolescents: implications for research and practice. Mahwah, USA: Lawrence Erlbaum; 1998.

270. Austin JK, Smith MS, Risinger MW, McNelis AM. Childhood epilepsy and asthma: comparison of quality of life. *Epilepsia* 1994;**35**(3):608–15.

271. Piers EV. Piers-Harris. Children's Self-Concept Scale. Revised manual. Los Angeles, CA: Western Psychological Services; 1984.

272. Austin JK, Huberty TJ. Development of the Child Attitude to Illness Scale (CATIS). *J Pediatr Psychol* 1993;**18**:467–80.

273. Achenbach TM, Edelbrock C. Manual for the child behaviour checklist and revised child behaviour profile. Burlington, VT: Queen City Printers; 1983.

274. Austin JK, Huberty TJ. Revision of the Family APGAR for use by 8-year-olds. *Fam Syst Med* 1989;**7**(3):323–7.

275. Austin JK, Huster GA, Dunn DW, Risinger MW. Adolescents with active or inactive epilepsy or asthma: a comparison of quality of life. *Epilepsia* 1996;**37**(12):1228–38.

276. Sockrider MM, Swank PR, Seilheimer DK, Schidlow DV. Internal validity and reliability of a modified NIH score. *Pediatr Pulmonol* 1994;**17**:86–96.

277. Sparrow SS, Balla DA, Cicchetti DV. Vineland Adapative Behaviour Scales. Circle Pines, MN: American Guidance Service; 1994.

278. Abidin RR. Parenting Stress Index Manual. Charlottesville, VA: Pediatric Psychology Press; 1983.

279. Kaplan RM. Health outcome models for policy analysis. *Health Psychol* 1989;**8**:723–35.

280. Testa MA, Lenderking WR. Quality of life considerations in AIDS clinical trials. In: Finkelstein DM, Schownfeld DA, editors. AIDS clinical trials. New Yory: Wiley-Liss, Inc; 1995.

281. Adams PF, Marano MA. Current estimates from the National Health Interview Survey, 1994. National Centre for Health Statistics. *Vital Health Stat* 1995;**10**(193).

282. Peterson JL, Zill N. Marital disruption, parent-child relationships and behaviour problems in children. *J Marriage Family* 1986;**48**:295–307.

283. Garber J, Walker LS, Zeman JL. Somatization symptoms in a community sample of children and adolescents: further validation of the Children's Somatization Inventory. *Psychol Assess* 1991;**3**:588–95.

284. Walker LS, Smith CA, Garber J, Van Slyke DA. Development and validation of the Pain Response Inventory for children. *Psychol Assess* 1997;**9**:392–405.

285. Walker LS, Greene JW. The Functional Disability Inventory: measuring a neglected dimension of child heath status. *J Pediatr Psychol* 1991;**16**:39–58.

286. Kovacs P. Rating scales to assess depression in school-aged children. *Acta Paedopsychiatr* 1980;**46**:305–15.

287. Radloff LS. The CES-D scale: a self report depression scale for research in the general population. *Appl Psychol Measure* 1977;**1**:385–401.

288. Harter S. Manual for self-perception profile for adolescents. Denver: University of Denver Press; 1988.

289. Gesten EL. A health resources inventory: the development of a measure of the personal and social competence of primary grade children. *J Consult Clin Psychol* 1976;**44**:775–86.

290. Chin SE, Shepherd RW, Cleghorn GJ, Patrick MK, Javorsky G, Frangoulis E, *et al.* Survival, growth and quality of life in children after orthotopic liver transplantation: a 5 year experience. *J Pediatr Child Health* 1991;**27**(6):380–5.

291. Cleary PD, Morrissey G, Yver A, Oster G. The effects of rG-CSF on health-related quality of life in children with congenital agranulocytosis. *Qual Life Res* 1994;**3**(5):307–15.

292. Felder-Puig R, Formann AK, Mildner A, Bretschneider W, Bucher B, Windhager R, *et al.* Quality of life and psychosocial adjustment of young patients after treatment of bone cancer. *Cancer* 1998;**83**:69–75.

293. Johnson DL, McCabe MA, Nicholson HS, Joseph AL, Getson PR, Byrne J, *et al.* Quality of long-term survival in young children with medulloblastoma. *J Neurosurg* 1994;**80**(6):1004–10.

294. Kun LE, Mulhern RK, Crisco JJ. Quality of life in children treated for brain tumors. Intellectual, emotional, and academic function. *J Neurosurg* 1983;**58**(1):1–6.

295. MacPhee M, Hoffenberg EJ, Feranchak A. Quality-of-life factors in adolescent inflammatory bowel disease. *Inflammatory Bowel Disease* 1998;**4**(1):6–11.

296. Martinez-Climent J, Castel SV, Esquembre MC, Verdeguer MA, Ferris TJA. Scale for assessing quality of life of children survivors of cranial posterior fossa tumors. *J Neurooncol* 1994;**22**(1):67–76.

297. Nespoli L, Verri AP, Locatelli F, Bertuggia L. The impact of pediatric bone marrow transplantation on quality of life. *Qual Life Res* 1995;**4**(3):233–40.

298. Packer RJ, Sposto R, Atkins TE, Sutton LN, Bruce DA, Siegel KR, *et al.* Quality of life in children with primitive neuroectodermal tumors (medulloblastoma) of the posterior fossa. *Pediatr Neurosci* 1987;**13**(4):169–75.

299. Harter S. The perceived competence scale for children. *Child Dev* 1982;**53**:87–97.

300. Achenbach TM. Manual for the Child Behaviour Checklist/4-18 and 1991 Profile. Burlington: University of Vermount Press; 1991.

301. Shumaker SA, Anderson RT, Czajkowski SM. Psychological tests and scales In: Spilker B, editor. Quality of life assessments in clinical trials. New York: Raven Press; 1990.

302. Perrin EC, Stein REK, Drotar D. Cautions in using the Child Behaviour Checklist: observations based on research about children with a chronic illness. *J Pediatr Psychol* 1991;**16**(4):411–21.

303. Bradlyn AS, Harris CV, Spieth LE. Quality of life assessment in pediatric oncology: a retrospective review of phase III reports. *Soc Sci Med* 1995;**41**(10):1463–5.

304. Eigen H, Clark NM, Wolle JM. Clinical-behavioural aspects of cystic fibrosis – directions for future research. *Am Rev Respir Dis* 1987;**136**(6):1509–13.

305. Ilangovan P, Pederson S, Godfrey S, Nikander K, Noviski N, Warner JO. Treatment of severe steroid dependent preschool asthma with nebulised budesonide suspension. *Arch Dis Child* 1993;**68**(3):356–9.

306. Sanders C, Egger M, Donovan J, Tallon D, Frankel S. Reporting on quality of life in randomised controlled trials: bibliographic study. *BMJ* 1998;**317**(7167):1191–4.

307. Chassany O, Bergmann JF, Caulin C. Authors are creating database of quality of life questionnaires [letter]. *BMJ* 1999;**318**(7191):1142–3.

308. Barr RD, Petrie C, Furlong W, Rothney M, Feeny D. Health-related quality of life during post-induction chemotherapy in children with acute lymphoblastic leukemia in remission. *Int J Oncol* 1997;**11**:333–9.

309. Abbott J, Dodd M, Bilton D, Webb AK. Treatment compliance in adults with cystic fibrosis. *Thorax* 1994;**49**:115–20.

310. Harper JI, Ahmed I, Barclay G, Lacour M, Hoeger P, Cork MJ, *et al.* Cyclosporin for severe childhood atopic dermatitis: short course versus continued therapy. *Br J Dermatol* 2000;**142**:52–4.

311. Thompson RJ, Gustafson KE. Adaption to chronic childhood illness. Washington: American Psychological Society; 1996.

312. Kazak AE, Penati B, Boyer BA, Himelstein B. A randomized controlled prospective outcome study of a psychological and pharmacological intervention protocol for procedural distress in pediatric leukemia. *J Pediatr Psychol* 1996;**21**(5):615–31.

313. Jelalian E, Stark LJ, Reynolds L, Seifer R. Nutrition intervention for weight gain in cystic fibrosis: a meta analysis. *J Pediatr* 1998;**132**:486–92.

314. Harrison H. The principles for family-centred neonatal care. Pediatrics 1993;**92**:643–50.

315. Allen MC, Donohue PK, Dusman AE. The limit of viability: neonatal outcome of infants born at 22-weeks to 25-weeks gestation. *N Engl J Med* 1993;**329**(22):1597–601.

316. de Keizer NF, Bonsel GJ, Gemke RJBJ. Health status prediction in critically ill children: a pilot study introducing standardised health ratios. *Qual Life Res* 1997;**6**:192–9.

317. Saigal S, Feeny D, Rosenbaum P, Furlong W, Burrows E, Stoskopf B. Self-perceived health status and health-related quality of life of extremely low-birth-weight infants at adolescence. *JAMA* 1996;**276**(6):453–9.

318. Eiser C. Practitioner review: long-term consequences of childhood cancer. *J Child Psychol Psychiatry* 1998;**39**(5):621–33.

319. Wiener JM. Response of medical personnel to the fatal illness of a child. In: Schoenberg B, Carr AC, Peretz D, Kutscher AH, editors. Loss and grief. New York: Columbia University Press; 1970.

320. Levine A, Artiss, KL, Susman EJ. The impact of childhood cancer on doctors and nurses. Proceedings of the American Cancer Society: National Conference on the Care of the Child with Cancer. American Cancer Society; 1979.

321. Goldman A. ABC of palliative care: special problems of children. *BMJ* 1998;**316**:49–52.

322. Jankovic M, Van Dongen-Melman JEWM, Vasiatou-Kosmidis H, Jenney MEM. Improving the quality of life for children with cancer. *Tumori* 1999;**85**:273–9.

323. Skeel RT. Do we need a new end point in clinical trials today – response. *J Natl Cancer Inst* 1989;**81**(14):1107.

324. Revicki DA. Relationship between health utility and psychometric health status measures. *Med Care* 1992;**27**:254–68.

325. Taylor KM, Macdonald KG, Bezjak A, Ng P, DePetrillo AD. Physicians' perspective on quality of life: an exploratory study of oncologists. *Qual Life Res* 1996;**5**:5–14.

326. Child Health in the Community: a guide to good practice. Department of Health; 1996.

327. The Children Act. The Stationary Office; 1989.

328. The Education Act. The Stationary Office; 1996.

329. Towards a combined child health service. British Paediatric Association. 1990.

330. Aaronson N. Assessing the quality of life of patients in cancer clinical trials: common problems and common sense solutions. *Eur J Cancer* 1990;**28A**:1304–6.

331. Department of Health. Health secretary announces membership of modernisation action teams to draw up national plan for health. London. Department of Health (Press release 2000/02081). 2000.

332. Battista RN, Hodge MJ. Quality of life research and health technology assessment: a time for synergy. *Qual Life Res* 1996;**5**:413–18.

333. Ruta DA, Garratt AM, Leng M, Russell IT, MacDonald LM. A new approach to the measurement of quality of life: the patient generated index. *Med Care* 1994;**32**(11):1109–26.

334. Joyce CRB, O'Boyle CA, McGee H. Individual quality of life: approaches to conceptualisation and assessment. Australia: Harwood Academic Publishers; 1999.

335. McGee HM, O'Boyle CA, Hickey AM, O'Malley K, Joyce CRB. Assessing the quality of life of the individual: the SEIQoL with a healthy and a gastroenterology unit population. *Psychol Med* 1991;**21**:749–59.

336. Hunt S. The concept of outcome. *Scand J Urol Nephrol* 1997;**183**:S71–S74

337. Hays RD, Stewart AL. The structure of self-reported health in chronic disease patients. *Psychol Assess* 1990;**2**:22–30.

338. Eiser C, Eiser JR. Social comparisons and quality of life among survivors of childhood cancer. *Psychol Health* 2000;**15**:435–50.

339. Skevington SM. Social comparisons in cross-cultural quality of life assessment. *Int J Mental Health* 1994;**23**:29–47.

340. Wallander JL. Theoretical and developmental issues in QL for children and adolescents. In: Wallander JL, Koot HM, editors. Quality of life in children and adolescents: concepts, methods and findings. UK: Harwood Academic Publishers. In press.

341. Levi R, Drotar D. Critical issues and needs in health-related quality of life assessment of children and adolescents with chronic health conditions. In: Drotar D, editor. Measuring health-related quality of life in children and adolescents. New Jersey: Lawrence Erlbaum; 1998.

342. Bond L, Nolan T, Adler R, Robertson C. The Child Behaviour Checklist in a Melbourne urban sample. *Australian Psychologist* 1994;**29**(2):103-9.

343. Hensley VR. Australian normative study of the Achenbach Child Behaviour Checklist. *Australian Psychologist* 1988;**23**(3):371–82.

344. Kuyken W, Orley J, Hudelson P, Sartorius N. Quality of life assessment across cultures. *Int J Mental Health* 1994;**23**(2):5–27.

345. Bullinger M, Anderson R, Cella D, Aaronson N. Developing and evaluating cross-cultural instruments from minimum requirements to optimal. *Qual Life Res* 1993;**2**(6):451–9.

346. Anderson RT, Aaronson NK, Alain PL, Wilkin D. International use and application of generic health-related measure. In: Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials. Philadelphia, PA: Lippincott-Raven; 1996.

347. Hui CCH, Traindis HC. Measurement in cross-cultural psychology. *J Cross Cultural Psychol* 1985;**16**(2):131–52.

348. Herdman M, Fox-Rusby J, Badia X. 'Equivalence' and the translation and adaptation of health-related quality of life questionnaires. *Qual Life Res* 1997;**6**(3):237–47.

349. WHOQOL Group. The World Health Organization Quality of Life Assessment (WHOQOL): development and general psychometric properties. *Soc Sci Med* 1998;**46**:1576–85.

350. Jirojankul P, Skevington S. Developing a quality of life measure for children aged 5–8 years. *Br J Health Psychol.* In press.

351. Feeny D, Furlong W, Torrence GW. The Health Utilities Index: an update. *Qual Life Newsletter* 1999;**22**:8–9.

352. Henry B, Grosskopf C, Aussage P, De Fontbrune S, Goehrs JM. Measuring quality of life in children with cystic fibrosis: The Cystic Fibrosis Questionnaires (CFQ). *Pediatr Pulmonol* 1997;**S14**:331.

353. Quittner AL, Sweeney S, Watrous M, Muzenberger P, Bearss K, Nitza AG, *et al.* Development of disease-specific quality of life measure for cystic fibrosis: linguistic validation of the French CFQ. *J Pediatr Psychol* 2000;**25**:403–14

354. Waters E, Salmon L, Wake M. The parent form Child Health Questionnaire in Austrailia: comparison of reliability, validity, structure and norms. *J Pediatr Psychol* 2000;**25**:381–91.

355. Eiser C, Kemp A, Jenney M, Maddocks A, Wandgraf J, Russell I, *et al.* Quality of life in UK children. Wales Office of Research and Development. Research in progress.

356. French DJ, Carroll A, Christie MJ. Health-related quality of life in Australian children with asthma: lessons for the cross-cultural use of life instruments. *Qual Life Res* 1998;**7**(5):409–19.

Appendix 1

Terms used in the search of electronic databases

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 or 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*)

Search 4 produces the main general papers on quality of life in children and adolescents. This result was supplemented by the results of search 5 which uses alternative terms for quality of life.

Searches 6 and 7 produced papers that report general measures of quality of life in children and adolescents with chronic illnesses.

Appendix 2

Data extraction form

Reference number:	Reviewer: Date of coding:

Name of instrument: Instrument in file: Y/N	
Disease specific ( ) specify	Generic

A. DETAILS OF PUBLICATION

1. Author(s):

2. Title:

3. Source and Ref:

4. Institution/contact address:

5. U.S.A. U.K. AUST/NZ. EURO. OTHER

B. RESEARCH QUESTION

Aim(s) *(Indicate here the aims of the study)*

(1) Development of QoL measure

(2) Determine statistical properties of QoL measure

(3) Compare proxy ratings

(4) Compare outcome in clinical trial

(5) Compare treatments

(6) Evaluate intervention

(7) Other *(specify)*

C. DESCRIPTION OF MEASURE *(indicate 'not reported' if any information is not given)*

1. Pre-existing measure *(specify name and changes made):*

2. (i) Did the development and testing of the instrument involve separate stages?

(ii) Is the pilot stage reported?

3. (i) Method of item generation *(e.g. semi-structured interview):*

(ii) Informant(s) *(e.g. children, parents, health professionals):*

4. Design of measure:

(i) (a) Adult (b) Child centred (c) Both

(ii) (a) Quality of Life (b) Health Status

5. Definition of QoL used:

(a) Multidimensional/WHO (b) Cost effectiveness (c) Goal-orientated

6. Age range(s) which measure is suitable for *(as stated by author):*

7. Number of domains:

8. Name of domains:

9. Total number of items:

10. Proxy ratings available? (Y/N)

11. Time taken to complete the measure

12. Response scale *(e.g. visual analogue, likert)*

13. Did questions involve assessment across time frames? *(e.g. last week, last month?)*

D. PROCEDURE *(indicate 'not reported' if any information is not given)*

STUDY DESIGN

1. **Design of study:** RCT
   Cohort study with matched concurrent controls
   Cohort study with unmatched concurrent controls
   Cohort study with historic controls
   Cohort study with no controls
   Other (specify)

2. **Method of randomisation:**

3. **Diagnostic classification (ICD-10):**

4. **Inclusion criteria:**

5. **Exclusion criteria:**

6. **Recruitment procedures:**

7. **SAMPLE CHARACTERISTICS**

   (i) Number of participants in each condition/target group/control:

   (ii) Mean age: Age range:

   (iii) Sex: — % female — % male

   (iv) Ethnicity:

   (v) Age at diagnosis:

   (vi) Time since diagnosis:

   (vii) Time since completion of treatment:

8. **(i) Representativeness of the sample:**

   (ii) Number of participants vs. non-participants:

9. **Informant** *(e.g. child, mother, father, clinician, teacher):*

10. **Method of administration and by whom** *(e.g. post/phone/interview):*

11. **Length of follow-up** *(complete if longitudinal design):*

E. EVALUATION *(indicate 'not applicable' to those sections which are not relevant)*

1. DEVELOPMENT OF AN INSTRUMENT
Validation measures used:
Child completed:

Parent completed:

Clinical:

RELIABILITY:

2. CROSS-SECTIONAL STUDIES
Base line measures *(specify whether child/parent completed)*

3. LONGITUDINAL STUDIES
Follow-up measures

F. ANALYSIS

1. Statistical techniques used

2. (i) Attrition rate

(ii) How was attrition managed?

3. Number followed up in each condition

G. RESULTS

Quantitative results *(free text)*

For example:-

- **Specify: validity (content, construct, concurrent, divergent & discriminant), and reliability (test-retest, inter-rater, internal consistency). Report correlation coefficients where necessary.**
- **Specify: differences between different populations, and differences between domains of QoL.**

H. QUALITY OF THE STUDY

1. **Effectiveness of measures** *(author's conclusions)*

2. **Effectiveness of measures** *(reviewer's conclusions, if different from above)*

3. **Limitations of measure and procedure** *(author's comments)*

4. **Limitations of measure and procedure** *(reviewer's comments, if different from above)*

Measure included with the article Y/N

I. REVIEWER'S DECISION

1. Does the study address the following (Y/N)? *include additional comments*

(i) The extent to which adult measures are used with children

(ii) Appropriateness of adult measures for use with children

(iii) Role of proxy raters and the feasibility

(iv) Reliability of proxy measures of QoL

2. Is the paper to be included? YES NO UNSURE

3. Reasons for decision

Appendix 3

Excluded studies

TABLE 11

Study	Excluded studies
Almond PS, Morell P, Matas AJ, Gillingham KJ, Chau KS, Brown A, *et al.* Transplanted children with long-term graft function have an excellent quality of life. *Transplant Proc* 1991;**23**(1 Pt 2):1380–1	QoL not measured
Alsarraf R, Jung CJ, Perkins J, Crowley C, Gates GA. Otitis media health status evaluation: a pilot study for the investigation of cost-effective outcomes of recurrent acute otitis media treatment. *Ann Otol Rhinol Laryngol* 1998;**107**(2):120–8	Clinical/functional health
Asikainen I, Kaste M, Sarna S. Predicting late outcome for patients with traumatic brain injury referred to a rehabilitation programme: a study of 508 Finnish patients 5 years or more after injury. *Brain Inj* 1998;**12**(2):95–107	Sample aged 0.8–71 Clinical/functional measures
Asonuma K, Inomata Y, Uemoto S, Egawa H, Kiuchi T, Okajima H, *et al.* Growth and quality of life after living-related liver transplantation in children. *Pediatr Transplant* 1998;**2**(1):64–9	Clinical (height and recovery weight)
Balestrini MR, Zanette M, Micheli R, Fornari M, Solero CL, Broggi G. Hemispheric cerebral tumors in children. Long-term prognosis concerning survival rate and quality of life – considerations on a series of 64 cases operated upon. *Childs Nerv Syst* 1990;**6**(3):143–7	Survival Neuropsychological assessment only
Baumgardner DJ, Burtea ED. Quality-of-life in technology-dependent children receiving home care, and their families – a qualitative study. *WMJ* 1998;**97**(8):51–5	Qualitative interviews
Bender BG. Measurement of quality of life in pediatric asthma clinical trials. *Ann Allergy Asthma Immunol* 1996;**77**(6):438–45	Review
Berardo DH, Berardo FM. Quality of life across age and family stage. *J Palliat Care* 1992;**8**(3):52–5	Comment
Berdeaux G, Hervie C, Smajda C, Marquis P. Parental quality of life and recurrent ENT infections in their children: development of a questionnaire. Rhinitis Survey Group. *Qual Life Res* 1998;**7**(6):501–12	Parental QoL
Boggs SR, Graham-Pole J, Miller EM. Life-threatening illness and invasive treatment: the future of quality of life assessment and research in pediatric oncology. In: Johnson JH, Johnson SB, editors. Advances in child health psychology. Gainseville, FL: University of Florida Press; 1991	Review article
Bomalaski MD, Teague JL, Brooks B. The long-term impact of urological management on the quality of life of children with spina bifida. *J Urol* 1995;**154**(2 Pt 2):778–81	Older than 18 years of age; education achievements
Bradford R. Children's psychological health status – the impact of liver transplantation: a review. *J R Soc Med* 1991;**84**(9):550–3	Review
Bradlyn AS, Harris CV, Spieth LE. Quality of life assessment in pediatric oncology: a retrospective review of phase III reports. *Soc Sci Med* 1995;**41**(10):1463–5	Review
Bradlyn AS, Pollock BH. Assessment of quality of life. *N Engl J Med* 1996;**335**(7):521	Comment
Brown SW. Quality of life – a view from the playground. *Seizure* 1994;**3** Suppl A:11–15	Comment
Bullinger M, Bullinger M, Ravens-Sieberer U. Health related QOL assessment in children: a review of the literature. *Eur Rev Appl Psychol/Revue Europeenne de Psychologie Appliquee* 1995;**45**(4):245–54	Review
Casey FA, Craig B G, Mulholland HC. Quality of life in surgically palliated complex congenital heart disease. *Arch Dis Child* 1994;**70**(5):382–6	Treadmill activity and school attendance

continued

TABLE 11 *contd*

Study	Excluded studies
Chang P. Psychosocial needs of long-term childhood cancer survivors: a review of literature. *Pediatrician* 1991;**18**(1):20–4	Review
Christie MJ, French D, Weatherstone L, West A. The patients' perceptions of chronic disease and its management: psychosomatics, holism and quality of life in contemporary management of childhood asthma. *Psychother Psychosom* 1991;**56**(4):197–203	Discussion
Clark DB, Kirisci L. Posttraumatic stress disorder, depression, alcohol use disorders and quality of life in adolescents. *Anxiety* 1996;**2**(5):226–33	Not a chronic illness
Cohen SR, Suzman K, Simms C, Burstein FD, Riski J, Montgomery GAC. Sleep apnea surgery versus tracheostomy in children: an exploratory study of the comparative effects on quality of life. *Plast Reconstr Surg* 1998;**102**(6):1855–64	Surgery
Crites LS, Siegel CA. Improving the quality of life for youngsters with HIV. *Child Today* 1990;**19**(3):24–7	QoL not measured
Crowley MJ, Kazdin AE. Child psychosocial functioning and parent quality of life among clinically referred children. *J Child Fam Stud* 1998;**7**(2):233–51	Parent QoL
Daniels H, Hogg B. An intercultural comparison of the quality of life of children and youth with handicaps in Denmark, Italy, the United Kingdom and Germany. *Educ Child Psychol* 1991;**8**(4):74–83	Disabilities
Ditesheim JA, Templeton JMJ. Short-term v long-term quality of life in children following repair of high imperforate anus. *J Pediatr Surg* 1987;**22**(7):581–7	No psychometric data
Dodson WE. Quality of life measurements in children with epilepsy. In: Trimble MR, Dodson WE, editors. Epilepsy and quality of life. New York: Raven Press; 1994	Review
Duffy CM, Duffy KN. Health assessment in the rheumatic diseases of childhood. *Curr Opin Rheumatol* 1997;**9**(5):440–7	Review
Eiser C, Jenney ME. Measuring symptomatic benefit and quality of life in paediatric oncology [editorial]. *Br J Cancer* 1996;**73**(11):1313–16	Editorial
Eiser C. Children's quality of life measures. *Arch Dis Child* 1997;**77**(4):350–4	Review
Eiser C, Havermans T. Knowledge and attitudes as determinants of 'quality of life' in children with asthma. In: Christie MJ, French DJ, editors. The assessment of quality of life in children with asthma. Chur: Harwood; 1994	Review
Enskar K, Carlsson M, von Essen L, Kreuger A, Hamrin E. Development of a tool to measure the life situation of parents of children with cancer. *Qual Life Res* 1997;**6**(3):248–56	Parent QoL
Erling A. Methodological considerations in the assessment of health-related quality of life in children. *Acta Paediatr Suppl* 1999;**88**(428):106–7	Review
Facione N. Quality of life issues in chronic otitis media with effusion: parameters for future study. *Int J Pediatr Otorhinolaryngol* 1991;**22**(2):167–79	Clinical, qualitative assessment
Faulkner MS, Clark FS. Quality of life for parents of children and adolescents with type I diabetes. *Diabetes Educ* 1998;**24**(6):721–7	Parent QoL
Feldman BM, Ayling-Campos A, Luy L, Stevens D, Silverman ED, Laxer RM. Measuring disability in juvenile dermatomyositis: validity of the childhood health assessment questionnaire. *J Rheumatol* 1995;**22**(2):326–31	Function
Ferry GD. Quality of life in inflammatory bowel disease: background and definitions. *J Pediatr Gastroenterol Nutr* 1999;**28**(4):S15–S18	Review
Finkelstein JW. Methods, models, and measures of health-related quality of life for children and adolescents. In: Drotar D, editor. Measuring health-related quality of life in children and adolescents. New Jersey: Lawrence Erlbaum; 1998	Review
Finlay AY, Lewis-Jones MS, Sharp JL, Dykes PJ. Children's Dermatology Life Quality Index: cartoon version validation [abstract]. *Br J Dermatol* 1998;**139**:15	Abstract

continued

TABLE 11 *contd*

Study *contd*	**Excluded studies**
Flat B, Srskaar D, Vinje O, Lien G, Aasland A, Moum T, Frre O. Measuring disability in early juvenile rheumatoid arthritis: evaluation of a Norwegian version of the childhood Health Assessment Questionnaire. *J Rheumatol* 1998;**25**(9):1851–8	Function
Fuggle P, Shand PA, Gill LJ, Davies SC. Pain, quality of life, and coping in sickle cell disease. *Arch Dis Child* 1996;**75**(3):199–203	Diary study
Gerard K. Determining the contribution of residential respite care to the quality of life of children with severe learning difficulties. *Child Care Health Dev* 1990;**16**(3):177–88	Severe learning disabilities
Glaser A, Walker D. Quality of life in surgically palliated complex congenital heart disease [letter; comment]. *Arch Dis Child* 1994;**71**(5):482	Letter
Gotay CC. Quality of life among survivors of childhood cancer: a critical review and implications for intervention. *J Psychosoc Oncol* 1987;**5**(4):5–23	Review
Goycochea-Robles MV, Garduno-Espinosa J, Vilchis-Guizar E, Ortiz-Alvarez O, Burgos-Vargas R. Validation of a Spanish version of the Childhood Health Assessment Questionnaire. *J Rheumatol* 1997;**24**(11):2242–5	Function
Grunberg SM, Groshen S, Steingass S, Zaretsky S. Comparison of conditional quality of life terminology and visual analogue scale measurements. *Qual Life Res* 1996;**5**(1):65–72	Sample aged 12–84 years
Hack MA. Consideration of the use of health status, functional outcome, and quality-of-life to monitor neonatal intensive care practice. *Pediatrics* 1999;**103**(1 Suppl E):319–28	Review article
Haverkamp F, Noeker M. 'Short stature in children – a questionnaire for parents': a new instrument for growth disorder-specific psychosocial adaptation in children. *Qual Life Res* 1998;**7**(5):447–55	Psychosocial
Hawkins WE, Hawkins MJ, Seeley J. Stress, health-related behavior and quality of life on depressive symptomatology in a sample of adolescents. *Psychol Rep* 1992;**71**(1):183–6	Not chronic illness
Hayes RP, Grinzaid KA, Duffey EB, Elsas LJ. The impact of Gaucher disease and its treatment on quality of life. *Qual Life Res* 1998;**7**(6):521–34	Sample aged 8–67 years
Henry B, Grosskopf C, Aussage P, de Fontbrune S, Goehrs JM. Measuring quality of life in children with cystic fibrosis: the cystic fibrosis questionnaires (CFQ). *Pediatr Pulmonol* 1997;**S14**:331	Abstract
Hinds PS. Quality of life in children and adolescents with cancer. *Semin Oncol Nurs* 1990;**6**(4):285–91	Review
Hinterberger W, Gadner H, Hocker P, Hajek-Rosenmayr A, Graninger W, Grabner G, *et al.* Survival and quality of life in 23 patients with severe aplastic anemia treated with bone marrow transplantation (BMT). *Blut* 1987;**54**(3):137–46	Clinical outcome
Hodges K, Wong MM, Latessa M. Use of the Child and Adolescent Functional Assessment Scale (CAFAS) as an outcome measure in clinical settings. *J Behav Health Serv Res* 1998;**25**(3):325–36	Function
Ishibashi A. Japanese nurse interested in the quality of life for children and adolescents with cancer [letter]. *J Pediatr Oncol Nurs* 1994;**11**(4):138	Letter
Ivan TM, Glazer JP. Quality of life in pediatric psychiatry: a new outcome measure. *Child Adolesc Psychiatr Clin N Am* 1994;**3**(3):599–611	Review
Jacobson AM, Fried K. Conceptual issues in developing quality of life assessments for children: illustrations from studies of insulin-dependent diabetes mellitus. In: Drotar D, editor. Measuring health-related quality of life in children and adolescents: implications for research and practice. New Jersey: Lawrence Erlbaum Associates; 1998	Review
Jenkin D, Danjoux C, Greenberg MAT. Subsequent quality of life for children irradiated for a brain tumor before age four years. *Med Pediatr Oncol* 1998;**31**(6):506–11	Survival and functional morbidity
Jenney ME, Kane RL, Lurie NA. Developing a measure of health outcomes in survivors of childhood cancer: a review of the issues. *Med Pediatr Oncol* 1995;**24**(3):145–53	Review

continued

TABLE 11 *contd*

Study	Excluded studies
Johnson SB, Kelly M, Henretta JC, Cunningham WR, Tomer A, Silverstein JH. A longitudinal analysis of adherence and health status in childhood diabetes. *J Pediatr Psychol* 1992;**17**(5):537–53	Clinical outcome
Jones B. The quality of life [editorial]. *Child Care Health Dev* 1981;**7**(2):65–6	Editorial
Jordan TE. Developing an international index of quality of life for children: the NICQL Index. *J R Soc Health* 1983;**103**(4):127–30	Not chronic illness
Joyce K, Singer M, Isralowitz R. Impact of respite care on parents' perceptions of quality of life. *Ment Retard* 1983;**21**(4):153–6	Parent QoL
Juniper EF. How important is quality of life in pediatric asthma? *Pediatr Pulmonol Suppl* 1997;**15**:17–21	Review
Juniper EF. Rhinitis management: the patient's perspective. *Clin Exp Allergy* 1998;**28**(S6):34–8	Review
Juniper EF, Guyatt GH, Feeny DH, Ferrie PJ, Griffith LE, Townsend M. Measuring quality of life in the parents of children with asthma. *Qual Life Res* 1996;**5**(1):27–34	Parent QoL
Kaplan RM. Implications of quality of life assessment in public policy for adolescent health. In: Drotar D, editor. Measuring health-related quality of life in children and adolescents. New Jersey: Lawrence Erlbaum; 1998	Review
Kaplan SH, Barlow S, Spetter D, Sullivan L, Khan A, Grand R. Assessing functional status and health related quality of life among school-aged children: reliability and validity of a new self-reported measure [abstract]. *Qual Life Res* 1995;**4**:444	Abstract
Keown PA, Shackleton CR, Ferguson BM. The influence of long-term morbidity on health status and rehabilitation following paediatric organ transplantation. *Eur J Pediatr* 1992;**151** Suppl 1:S70–5	Review
Kiltie AE, Gattamaneni HR. Survival and quality of life of paediatric intracranial germ cell tumour patients treated at the Christie Hospital, 1972–1993. *Med Pediatr Oncol* 1995;**25**(6):450–6	Evaluation of case notes; neurological and functional data
Kiltie AE, Lashford LS, Gattamaneni HR. Survival and late effects in medulloblastoma patients treated with craniospinal irradiation under three years old. *Med Pediatr Oncol* 1995;**28**(5):348–54	Survival, marital status, height
Kind P. Measuring quality of life in children. In: Christie MJ, French DJ, editors. The assessment of quality of life in children with asthma. Chur: Harwood; 1994	Review
Kohler JA, Radford M. Terminal care for children dying of cancer: quantity and quality of life. *BMJ* 1985;**291**(6488):115–16	Qualitative data
Koot HM, Bouman NH. Potential uses for quality-of-life measures in childhood inflammatory bowel disease. *J Pediatr Gastroenterol Nutr* 1999;**28**(4):S56–S61	Review
Korsch BM. Health-related quality of life after organ transplantation in childhood [editorial; comment]. *Pediatr Transplant* 1997;**1**(2):99–100	Editorial
Kraybill EN, Bose CL, Corbet AJ, Garcia-Prats J, Asbill D, Edwards K, *et al.* Double-blind evaluation of developmental and health status to age 2 years of infants weighing 700 to 1350 grams treated prophylactically at birth with a single dose of synthetic surfactant or air placebo. *J Pediatr* 1995;**126**(5 Pt 2):S33–S42	Morbidity
Krmar RT, Eymann A, Ramirez JA, Ferraris JR. Quality of life after kidney transplantation in children. *Transplantation* 1997;**64**(3):540–1	No psychometric properties
Lantos JD. Some moral and political pitfalls in measuring quality of life. In: Drotar D, editor. Measuring health-related quality of life in children and adolescents. New Jersey: Lawrence Erlbaum; 1998	Review
LeBidois J, Kachaner J, Vouhe P, Sidi D, Tamisier D. Heart transplantation in children: mid-term results and quality of life. *Eur J Pediatr* 1992;**151** Suppl 1:S59–64	Survival, clinical data

continued

TABLE 11 *contd*

Study	Excluded studies
Lenderking WR, Testa MA, Katzenstein D, Hammer S. Measuring quality of life in early HIV disease: the modular approach. *Qual Life Res* 1997;**6**(6):515–30	Sample aged 13–70 years
Leung WC. Use of health services by children. Objective measures of health status are essential [letter; comment]. *BMJ* 1999;**318**(7187):877	Letter
Levi R, Drotar D. Critical issues and needs in health-related quality of life assessment of children and adolescents with chronic health conditions. In: Drotar DM. editor. Measuring health-related quality of life in children and adolescents. New Jersey: Lawrence Erlbaum; 1998	Review
Lin ST, Yang CP, Liang DC. Health status of patients with childhood acute lymphoblastic leukemia in continuous complete remission for over five years. The Taiwan Children's Cancer Study Group. *J Formos Med Assoc* 1993;**92**(8):702–10	Clinical data
Lockey RF, DuBuske LM, Friedman B, Petrocella V, Cox F, Rickard K. Nocturnal asthma: effect of salmeterol on quality of life and clinical outcomes. *Chest* 1999;**115**(3):666–73	Predominantly adult sample
Mackie PC, Jessen EC, Jarvis SN. The lifestyle assessment questionnaire: an instrument to measure the impact of disability on the lives of children with cerebral palsy and their families. *Child Care Health Dev* 1998;**24**(6):473–86	Disability
Marra CA, Levine M, McKerrow R, Carleton BC. Overview of health-related quality-of-life measures for pediatric patients: application in the assessment of pharmacotherapeutic and pharmacoeconomic outcomes. *Pharmacotherapy* 1996;**16**(5):879–88	Review
McGee HM. Quality of life: assessment issues for children with chronic illness and their families. In: Christie MJ, French DJ, editors. The assessment of quality of life in children with asthma. Chur: Harwood; 1994	Review
Meijboom F, Szatmari A, Deckers JW, Utens EM, Roelandt JR, Bos E, *et al.* Cardiac status and health-related quality of life in the long term after surgical repair of tetralogy of Fallot in infancy and childhood. *J Thorac Cardiovasc Surg* 1995;**110**(4 Pt 1):883–91	Exercise capacity
Miller JJ. Long term outcomes and quality of life in juvenile RA [letter; comment]. *J Rheumatol* 1998;**25**(6):1244	Letter
Miller JR, Colbert AP, Osberg JS. Ventilator dependency: decision-making, daily functioning and quality of life for patients with Duchenne muscular dystrophy. *Dev Med Child Neurol* 1990;**32**(12):1078–86	No psychometric data
Morel P, Almond PS, Matas AJ, Gillingham KJ, Chau C, Brown A, *et al.* Long-term quality of life after kidney transplantation in childhood. *Transplantation* 1991;**52**(1):47–53	No psychometric data
Mostow EN, Byrne J, Connelly RR, Mulvihill JJ. Quality of life in long-term survivors of CNS tumors of childhood and adolescence. *J Clin Oncol* 1991;**9**(4):592–9	Adult survivors
Mulhern RK, Heideman RL, Khatib ZA, Kovnar EH, Sanford RA, Kun LE. Quality of survival among children treated for brain stem glioma. *Pediatr Neurosurg* 1994;**20**(4):226–32	Quality of survival; neurological and behavioural measures
Mulhern RK, Horowitz ME, Ochs J, Friedman AG. Assessment of quality of life among pediatric patients with cancer. *Psychol Assess* 1989;**1**(2):130–8	Review
Murray KJ, Passo MHA. Functional measures in children with rheumatic diseases. *Pediatr Clin N Am* 1995;**42**(5):1127–54	Review
Myatt HM, Myatt RA. The development of a paediatric quality of life questionnaire to measure post-operative pain following tonsillectomy. *Int J Pediatr Otorhinolaryngol* 1998;**44**(2):115–23	Surgery
Neff EJ, Dale JC. Assessment of quality of life in school-aged children: a method – Phase 1. *Matern Child Nurs J* 1990;**19**(4):313–20	No psychometric data
Ottenbacher KJ, Msall ME, Lyon NR, Duffy LC, Granger CV, Braun S. Interrater agreement and stability of the Functional Independence Measure for Children (WeeFIM): use in children with developmental disabilities. *Arch Phys Med Rehabil* 1997;**78**(12):1309–15	Disabilities

continued

TABLE 11 *contd*

Study	Excluded studies
Pal DK. Quality of life assessment in children: a review of conceptual and methodological issues in multidimensional health status measures. *J Epidemiol Community Health* 1996;**50**(4):391–6	Review
Parsons SK, Brown AP. Evaluation of quality of life of childhood cancer survivors: l a methodologica conundrum. *Med Pediatr Oncol* 1998;Suppl 1:46–53	Review
Price K. Quality of life for terminally ill children. *Soc Work* 1989;**34**(1):53–4	Review
Rose MA, Clark-Alexander B. Caregivers of children with HIV/AIDS: quality of life and coping styles. *J Assoc Nurses AIDS Care* 1998;**9**(1):58–65	Parent QoL
Rosenfeld RM, Goldsmith AJ, Madell JR. How accurate is parent rating of hearing for children with otitis media? *Arch Otolaryngol Head Neck Surg* 1998;**124**(9):989–92	QoL not assessed
Ruperto N, Ravelli A, Levinson JE, Shear ES, Murray K, Link TB, *et al.* Long-term health outcomes and quality of life in American and Italian inception cohorts of patients with juvenile rheumatoid arthritis. II. Early predictors of outcome. *J Rheumatol* 1997;**24**(5):952–8	Predominantly adult sample
Rutishauser C, Sawyer SM, Bowes G. Quality-of-life assessment in children and adolescents with asthma. *Eur Respir J* 1998;**12**(2):486–94	Review
Saigal S, Rosenbaum P, Hoult L, Furlong W, Feeny D, Burrows E, *et al.* Conceptual and methodological issues in assessing health-related quality of life in children and adolescents: illustration from studies of extremely low birthweight survivors. In: Drotar D, editor. Measuring health-related quality of life in children and adolescents. New Jersey: Lawrence Erlbaum; 1998	Review
Sanders J, Sullivan K, Witherspoon R, Doney K, Anasetti C, Beatty P, *et al.* Long term effects and quality of life in children and adults after marrow transplantation. *Bone Marrow Transplant* 1989;**4**(S4):27–9	Clinical and neuropsychological functioning
Schor EL. Children's health and the assessment of health-related quality of life. In: Drotar D, editor. Measuring health-related quality of life in children and adolescents. New Jersey: Lawrence Erlbaum; 1998	Review
Schulz RM, Dye J, Jolicoeur L, Cafferty T, Watson J. Quality-of-life factors for parents of children with asthma. *J Asthma* 1994;**31**(3):209–19	Parent QoL
Sejima H, Ito M, Shiraishi H, Kobayashi A, Kishi K, Mori CA. Quality of life in schoolchildren and students with epilepsy: a survey by questionnaire for nurse-teachers in Shimane prefecture. *Jpn J Psychiatry Neurol* 1993;**47**(2):359–60	No psychometric properties
Singh G, Athreya BH, Fries JF, Goldsmith DP. Measurement of health status in children with juvenile rheumatoid arthritis. *Arthritis Rheum* 1994;**37**(12):1761–9	Function
Sjbu LAB. Parents of children with cerebral palsy in Nordland (county in the north of Norway); factors connected to their quality of life and coping of the circumstances around the handicapped child. *Arctic Med Res* 1994;**53**(S1):30–1	Parent QoL
Skuse D, Elgar K, Morris E A B S. Quality of life in Turner syndrome is related to chromosomal constitution: implications for genetic counselling and management. *Acta Paediatr Suppl* 1999;**88**(428):110–13	Social and educational adjustment
Skuse DH, Cave S, O'Herlihy A, South R. Quality of life in children with Turner syndrome: parent, teacher, and individual perspectives. In: Drotar D, editor. Measuring health-related quality of life in children and adolescents. New Jersey: Lawrence Erlbaum; 1998	Adjustment
Slavc I, Salchegger C, Hauer C, Urban C, Oberbauer R, Pakisch B, *et al.* Follow-up and quality of survival of 67 consecutive children with CNS tumors. *Childs Nerv Syst* 1994;**10**(7):433–43	Survival and neurological sequelae
Sokal EM. Quality of life after orthotopic liver transplantation in children. An overview of physical, psychological and social outcome. *Eur J Pediatr* 1995;**154**(3):171–5	Review
Spieth LE, Harris CV. Assessment of health-related quality of life in children and adolescents: an integrative review. *J Pediatr Psychol* 1996;**21**(2):175–93	Review

continued

TABLE 11 *contd*

Study	Excluded studies
Staab D, Wenninger K, Gebert N, Rupprath K, Bisson S, Trettin M, *et al.* Quality of life in patients with cystic fibrosis and their parents: what is important besides disease severity? *Thorax* 1998;**53**(9):727–31	Parent QoL
Stabler B, Frank NC. Quality of life and the psychiatric status of individuals treated with GH in childhood. In: Drotar D, editor. Measuring health-related quality of life in children and adolescents. New Jersey: Lawrence Erlbaum; 1998	Review
Stone RD, Beasley PJ, Treacy SJ, Twente AW, Vacanti JP. Children and families can achieve normal psychological adjustment and a good quality of life following pediatric liver transplantation: a long-term study. *Transplant Proc* 1997;**29**(1-2):1571–2	Qualitative
Takei S, Hokonohara M. Quality of life and daily management of children with rheumatic disease. *Acta Paediatr Jpn* 1993;**35**(5):454–63	Review
Testa MA, Simonson DC. "Assessment of quality of life": Reply. *N Engl J Med* 1996;**335**(7):521–2	Comment
Townsend M, Feeny DH, Guyatt GH, Furlong WJ, Seip AE, Dolovich JA. Evaluation of the burden of illness for pediatric asthmatic patients and their parents. *Ann Allergy* 1991;**67**(4):403–8	Burden of illness
van der Sluis CK, Kingma J, Eisma WH, Duis HJ. Pediatric polytrauma: short-term and long-term outcomes. *J Trauma* 1997;**43**(3):501–6	Older than 18 years of age at follow-up
Vivier PM, Bernier JA, Starfield B. Current approaches to measuring health outcomes in pediatric research. *Curr Opin Pediatr* 1994;**6**(5):530–7	Review
Weber TR, Tracy TJ, Connors RH. Short-bowel syndrome in children. Quality of life in an era of improved survival. *Arch Surg* 1991;**126**(7):841–6	Survival and descriptive characteristics
West A. Methodological issues in the assessment of quality of life in childhood asthma: what educational research has to offer. In: Christie MJ, French DJ, editors. The assessment of quality of life in children with asthma. Chur: Harwood; 1994	Review
Wiklund I, Erling A, Albertsson-Wikland K. Critical review of measurement issues in quality of life assessment for children with growth problems. In: Drotar D, editor. Measuring health-related quality of life in children and adolescents. New Jersey: Lawrence Erlbaum; 1998	Review
Wray J, Radley-Smith R, Yacoub M. Effect of cardiac or heart-lung transplantation on the quality of life of the paediatric patient. *Qual Life Res* 1992;**1**(1):41–6	Development and cognitive function
Wright PS. Parents' perceptions of their quality of life. *J Pediatr Oncol Nurs* 1993;**10**(4):139–45	Parent QoL
Yang TF, Wong TT, Kwan SY, Chang KP, Lee YC, Hsu TC. Quality of life and life satisfaction in families after a child has undergone corpus callostomy. *Epilepsia* 1996;**37**(1):76–80	Family QoL
Yu VY, Manlapaz ML, Tobin J, Carse EA, Charlton MP, Gore JR. Improving health status in extremely low birthweight children between two and five years. *Early Hum Dev* 1992;**30**(3):229–39	Clinical indicators
Zamberlan KE. Quality of life in school-age children following liver transplantation. *Matern Child Nurs J* 1992;**20**(3-4):167–229	No psychometric data

Appendix 4

Summary of generic QoL measures identified for use with children

The Child Heath and Illness Profile – Adolescent Edition

(CHIP-AE: Riley *et al.*, 1998;[107,108] Starfield *et al.*, 1993,[109] 1995,[110] 1996[111])

Description	A taxonomy of health profile-types for describing adolescent health
Age range	11 to 17 years
Number of domains	6 domains (20 subdomains); conceptually derived and supported by factor analysis
Name of domains	Discomfort, satisfaction, disorders, achievement, resilience, and risks
Number of items	107 items plus an additional 46 items that are specific to disease or injury
Rating scale	3 to 5 point Likert scales
Report	Self-report only
Time to complete	30 minutes (approx)
Example item:	

SELF ESTEEM

	Completely agree	Mostly agree	Agree a little	Do not agree
I have a lot of good qualities	☐	☐	☐	☐

Reliability	*Test–retest:* good stability (1 week) Pearson correlation greater than 0.60 for 19 of the 20 subdomains (exception was 'home safety and health' subdomain in Resilience domain) *Internal consistency:* Cronbach's alphas exceeded 0.70 for all subdomains (except Academic Performance) in at least two out of four samples assessed
Validity	*Criterion:* parent–adolescent correspondence on the CHIP-AE was examined using chance-corrected measures of agreement on more observable measures of health *Convergent validity:* demonstrated for the subdomains of Emotional Discomfort and Family Involvement by comparing them with known measures of these constructs (correlations ranged between 0.59 and 0.68) Hypothesised age, gender, and socio-economic differences were found in virtually every case, taken as evidence of construct validity of the CHIP-AE

Child Health Questionnaire

(CHQ: Landgraf & Abetz, 1997;[19] Landgraf *et al.*, 1998[112])

Description	Developed to fulfil the same role as the SF-36, the CHQ is a generic health status measure
Age range	4 to 18 years
Number of domains	14
Name of domains	Physical functioning, role/social functioning (physical), role/social functioning (emotional), role/social (behavioural), general health perceptions, bodily pain/discomfort, general behaviour, mental health, self-esteem, parental impact (emotional), parental impact (time), family functioning (family activities – family cohesion), global item, change in health
Number of items	Parent form PF-98, PF-50, PF-28; child form CF-87
Respondent	Parent or child 11 years and above (but not parallel versions)
Time to complete	20 minutes (approx)

Example item:

"Your child's physical activities"
During the past 4 weeks, had your child been limited in any of the following activities due to health problems?
Doing things that take a lot of energy, such as playing soccer or running.

Yes, limited a lot	Yes, limited some	Yes, limited a little	No, not limited

Reliability *Internal consistency:* Cronbach's alpha range between 0.62 and 0.91 (across 3 samples using the CHQ-CF87)[19]
Validity *Concurrent:* See Landraf *et al.*, or further developments[112]

The Child Quality of Life questionnaire

(CQOL: Graham *et al.*, 1997[113])

Description Multidimensional questionnaire, in which each domain is assessed using one item with multiple choice
Age range 9 to 15 years
Number of domains 15
Name of domains Getting about and using hands, 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
Number of items 15 (each with 3 levels)
Rating scale 7-point Likert scale
Report Child and parent versions
Time to complete 10 to 15 minutes
Example item:

Possible problems include: Clumsy; difficulty running; difficulty walking; unable to control movements; uses a wheel chair; confined to bed.

Over the past month how well have you been in these ways?

As well as any other child the same age ..... Confined to bed

How upset have you been in these ways?

Extremely upset ..... Not at all upset

How satisfied have you been by how you are in these ways?

Very satisfied ..... Not at all satisfied

Reliability *Test–retest:* Good
Validity *Convergent:* 0.63, $p = 0.01$ with Child Global Assessment Scale

Dartmouth Picture and word COOP charts for assessing adolescent health

(COOPS: Wasson *et al.*, 1994,[114] 1995[115])

Description Based on the adult COOP charts. Items are presented on illustrated cards
Age range Adolescent

Number of domains	6
Name of domains	Physical fitness, emotional feelings, school work, social support, family communications, health habits
Number of items	6
Rating scale	5-point Likert scale
Respondent	Adolescent
Time to complete	Not reported
Example item:	

FAMILY COMMUNICATIONS
During the past month, how often did you talk about your problems, feelings or opinions with someone in your family?
1. All of the time
2. Most of the time
3. Some of the time
4. A little of the time
5. None of the time

Reliability	*Test–retest:* 0.77
Validity	*Construct:* Correlations between COOPS and longer measures was 0.62

Exeter Health-Related Quality of Life measure

(EHRQL: Eiser *et al.*, 2000[116] 1999[117])

Description	Based on self-discrepancy theory, it is computer administered with the aid of an interviewer
Age range	6 to 11
Number of domains	Single scale
Name of domains	Not applicable
Number of items	16 (reduced to 12 in a later version)
Rating scale	Two VAS ('very much like me' to 'not very much like me'; 'don't want to be like that' to 'want to be like that')
Respondent	Child
Time to complete	20 minutes
Example item:	

"This is Joe doing P.E. Joe is very fit and healthy and likes doing P.E./games at school".
"Are you like Joe?"
Not at all like me ..... Exactly like me
"Do you want to be like Joe?"
Don't want to be like that ..... Want to be like that

Reliability	*Internal consistency:* satisfactory (actual self = 0.62; ideal self = 0.69; discrepancy = 0.50)
Validity	*Clinical:* distinguished between children with asthma and healthy children

Functional Status II(R)

(FSIIR: Stein & Jessop, 1990[14])

Description	Developed for the measurement of individual child health status and characterising populations. Modelled on the SIP
Age range	0 to 16 years
Number of domains	8
Name of domains	Communication, mobility, mood, energy, play, sleep, eating, and toileting
Number of items	Long (43-item) and short (14-item) version

Rating scale	Two 3-point Likert scale: 'never or rarely' 'some of the time' or 'almost always'. Dysfunction is assessed by whether it is due 'fully', 'partly' or 'not at all' to a health problem
Respondent	Parent
Administration	Interviewer
Time to complete	Less than 30 minutes
Example items	Not given
Reliability	*Internal consistency:* All alpha's > 0.80
Validity	*Construct:* total scales correlated with the criterion variable in every case

Generic Health Questionnaire

(Collier, 1997[118])

Description	Children are asked to relate to the responses of children in a story. QoL is assessed by measuring satisfaction with how life is compared to how one might expect it to be. The sum of the discrepancies provides the QoL score
Age range	6 to 14 years
Number of domains	5 (preliminary)
Name of domains	General affect, peer relationships, attainments, relationships with parents, and general satisfaction
Number of items	25
Rating scale	5-point Likert scale
Respondent	Self-completed
Time to complete	Not reported
Example item:	

Having fun – [the boys] find out that one boy always has fun, one often has fun, one sometimes has fun, another hardly ever has fun and one boy never has fun. Tick the boy most like you.

Always	Often	Sometimes	Hardly ever	Never

Reliability	*Internal consistency:* Cronbach's alpha 0.75 *Test–retest:* not reported
Validity	*Concurrent:* good (correlations were conducted between overall QoL and the general life satisfaction question)

How Are You?

(HAY: Bruil, 1999[119])

Description	Measures general as well as disease-specific aspects of QoL
Age range	7 to 13
Number of domains	6 generic; 2 chronic illness; 2 disease-specific
Name of domains	Generic section: physical activities, cognitive tasks, social activities, social problems, physical complaints, treatment. Chronic illness section: concerns, feelings of inferiority. Disease-specific section: physical complaints, treatment tasks
Number of items	80
Rating scale	4-point Likert scale
Respondent	Child or proxy
Time to complete	30 minutes

Example item:

Have you remembered what you learned at school during the past seven days?
No, never
Yes, sometimes
Yes, often
Yes, very often

Reliability	*Internal consistency:* Cronbach's alphas in the child version ranged from 0.77 to 0.86, and in the parent version 0.86 to 0.93 for the generic section
Validity	*Construct:* correlations with the child version of the HAY scales measuring 'concerns' and 'feelings of inferiority' and a revised version of the Child Attitude to Illness scale were 0.53 and 0.59

KINDL

(Ravens-Sieberer *et al.*, 1999[120])

Description	Generic measure assessing 4 domains of QoL
Age range	10 to 16 years
Number of domains	4
Name of domains	Psychological well-being, social relationships, physical functioning, and every-day life activities
Number of items	40
Rating scale	5-point Likert (1 = never to 5 = always)
Respondent	Child
Time to complete	12 minutes (approx)
Example item:	

During the last week I enjoyed being with my friends	Never	Rarely	Sometimes	Often	Always
	☐	☐	☐	☐	☐

Reliability	*Internal consistency:* each subscale reached a Cronbach's alpha > 0.75, total scale 0.95 *Test–retest:* very good (0.80)
Validity	*Convergent:* very good *Clinical:* poor (no differences between healthy and ill children) *Concurrent:* moderate to strong correlations with SF-36 subscales and a life satisfaction measure

Nordic Quality of Life Questionnaire for Children

(Lindstrom *et al.*,1991,[4] 1993,[121] 1994[122])

Description	Generic questionnaire intended to measure all aspects of children's lives
Age range	2 to 18 years
Number of domains	4 life spheres, each with 5 dimensions
Name of domains	Life spheres: Global, External, Inter-personal, Personal each with the following domains (Physical, Mental, Spiritual, Social, Economic)
Number of items	74
Rating scale	Not reported
Respondent	Child and parent
Time to complete	Not reported
Example item	Not given
Reliability	Under evaluation
Validity	Under evaluation

Pediatric Quality of Life Questionnaire

(PedsQL: Varni *et al.*, 1999,[123] (submitted)[124])

Description	Derived from data collected for measurement of QoL in children with cancer. This measure consists of a generic core and disease specific modules
Age range	2 to 18 years
Number of domains	3
Name of domains	Physical functioning, psychological functioning, social functioning
Number of items	15 (core)
Rating scale	3-point Likert scale (for 5 to 7-year-olds); 5-point Likert scale for 8 to 18-year-olds
Report	Parent or child 5 years and above
Time to complete	5 to 10 minutes
Example item:	

[child report ages 5–7] Social functioning

	Not at all	Sometimes	A lot
It is hard for you to get along with other kids	◯	◯	◯

Reliability	*Internal consistency:* Cronbach's alpha ranged between 0.70 to 0.89 for self-report. Total scale 0.93 (child) and 0.93 (parent)
Validity	*Construct:* the psychological and social functioning scales correlated with standardised scales of emotional distress, perceived competence, and social functioning *Clinical validity:* cancer patients on- versus off-treatment ($p < 0.004$)

Perceived Illness Experience

(PIE: Eiser *et al.*, 1999[125] 1995[126])

Description	This scale was initially developed to assess perceived illness experience in young people with cancer, but may be used with other groups of children/young people with chronic illness
Age range	8 to 18
Number of domains	8
Name of domains	Perceived impact on physical appearance, interference with activity, disclosure, integration in school, peer rejection, parental behaviour, manipulation, and preoccupation with illness
Number of items	34
Rating scale	5-point likert scale
Respondent	Self or parent completed
Time to complete	Not reported
Example item:	

Because of my illness I am not always able to join in with what my friends are doing (1 = disagree to 5 = agree)

Reliability	*Test–retest:* acceptable ($r = 0.92$ for total score) *Internal consistency:* adequate
Validity	*Construct:* total PIE scores correlated with established measures of physical and psychological functioning

Quality of Life Profile – Adolescent Version

(QOLP-AV: Raphael *et al.*, 1996[127])

Description	Emerging concepts of adolescent health considered within a broadened QoL perspective
Age range	Adolescent (14 to 20 years)

Number of domains	3
Name of domains	Being, Belonging, Becoming
Number of items	54
Rating scale	Importance and enjoyment ratings on a 5-point scale
Respondent	Self-complete
Time to complete	Not reported
Example item	Not given
Reliability	*Internal consistency:* all alphas > 0.70
Validity	*Construct:* QoL scores were correlated with self-esteem, life satisfaction, social support and life chances

TACQOL

(Theunissen *et al.*, 1998;[129] Vogels *et al.*, 1998[130])

Description	Generic measure assessing 7 domains of QoL. Incorporates both health status and QoL
Age range	6 to 15 years
Number of domains	7
Name of domains	Pain and symptoms, basic motor functioning, autonomy, cognitive functioning, social functioning, global positive emotional functioning, and global negative, emotional functioning
Number of items	108
Rating scale	3 and 4 point Likert
Respondent	Parent and child (parallel forms)
Time to complete	10 minutes (parents)
Example item:	

[Parents/carers of children aged 6 to 15] Dealings with other children and with you in recent weeks

40. Other children asked my child: to play with them

	Yes	Too little	Never	
At that time my child felt:	Fine	Not so good	Quite bad	Bad

Reliability	*Internal consistency:* Cronbach's alpha 0.71 to 0.89 for the parent version; 0.59 to 0.86 for the child's version
Validity	*Convergent:* completely confirmed using multi-trait multi-method analysis *Clinical:* mostly confirmed according to authors predictions

Warwick Child Health and Morbidity Profile

(WCHMP: Spencer & Coe, 1996[131])

Description	Presents a single summary of a child's health and illness experience
Age range	0 to 5 years
Number of items	16
Number of domains	10
Name of domains	General health status, acute minor illness status, behavioural status, accident status, acute significant illness status, hospital administration status, immunisation status, chronic illness status, functional health status, and health-related quality of life
Rating scale	Four categories of response
Respondent	Parent
Administration	Interview
Time to complete	10 minutes

Reliability	*Test–retest:* kappa range 0.41 to > 0.80
	Inter-rater: perfect on 6 domains; > 0.76 for remaining domains
Validity	*Construct:* validation against clinic records kappa values range from 0.70 to 0.85

The Multiattribute System for Classifying Health Status:

i) Heath Utilities Index Mark 2 (HUI2: Feeny *et al.*, 1992[78])

Description	This multi-attribute system focuses on function rather than performance. The HUI system permits the integration of mortality and morbidity, allows broad comparisons, and facilitates the conduct of cost–utility analysis. The levels of each attribute are meant to be interpreted as developmentally appropriate for age. There are three systems (HUI Mark 1, 2 and 3). The HUI1 is rarely used now
Age range	2 to 18 years
Number of domains	7
Name of domains	Sensation, emotion, self-care, mobility, cognition, pain, fertility
Rating scale	3 to 5 levels of functioning
Respondent	Clinician, parent, (self-complete > 8 years)
Time to complete	A few minutes
Example item:	

HUI Mark 2: Cognition domain
1. Learns and remembers school work normally for age
2. Learns and remembers school work more slowly than classmates as judged by parents and/or teachers
3. Learns and remembers very slowly and usually requires special educational assistance
4. Unable to learn and remember

Reliability	*Test–retest:* intra-class correlation = 0.77
	Internal consistency: not relevant
Validity	*Clinical:* discriminates on (n = 20) and off therapy (n = 8)
	Also brain tumours compared with acute lymphoblastic leukaemia (ALL), Wilms and neuroblastoma[78]
	Survivors of ALL lower scores than normal population[355]

ii) Health Utilities Index Mark 3 (HUI3: Boyle *et al.*, 1995[132])

Description	A utility-based scoring system is being developed for the HUI3. The HUI2 and HUI3 provide complementary sets of information. Together the two systems include 10 dimensions of health status
Age range	Birth to old age
Number of domains	8
Name of domains	Vision, hearing, speech, ambulation, dexterity, emotion, cognition and pain/discomfort
Number of items	8
Rating scale	Each attribute consists of 5 or 6 levels of functioning
Respondent	Parent report for children; self-report for adults
Time to complete	Not reported
Example item:	

Cognition
1. Able to remember most things, think clearly and solve day to day problems
2. Able to remember most things, but has a little difficulty when trying to think and solve day to day problems
3. Somewhat forgetful, but able to think clearly and solve day to day problems
4. Somewhat forgetful, and has a little difficulty when trying to thing or solve day to day problems.
5. Very forgetful, and has great difficulty when trying to think or solve day to day problems
6. Unable to remember anything at all, and unable to think or solve day to day problems

Reliability	*Test–retest:* estimates based on kappa ranged from 0.137 (speech) to 0.728 (vision)
Validity	See Feeny *et al.*[349]

16 Dimensions

(16D: Apajasalo *et al.*, 1996[133])

Description	Derived from the 15D adult measure.[170] The 16D consists of 16 multiple choice questions each representing one health-related dimension. Importance weights and within-dimension desirability can also be calculated
Age range	12 to 15 years
Number of domains	16
Name of domains	Mobility, vision, hearing, breathing, sleeping, eating, elimination, speech, mental function, discomfort and symptoms, school and hobbies, friends, physical appearance, depression, distress, and vitality
Number of items	16
Rating scale	Select the level of functioning on a scale of 1 to 5
Respondent	Self-completed only
Time to complete	5 to 10 minutes
Example item:	

The levels of the 'friends' dimension
1. My state of health has no influence on my getting friends or being with friends
2. My state of health makes getting friends or being with friends a little difficult
3. My state of health makes getting friends or being with friends quite difficult
4. My state of health makes getting friends or being with friends almost impossible
5. My state of health makes getting friends or being with friends totally impossible

Reliability	*Test–retest:* poor
Validity	*Clinical:* discriminated between patients waiting an organ transplant and controls

17 Dimensions

(17D: Apajasalo *et al.*, 1996[134])

Description	Based on the 15D adult version and the 16D adolescent version (see above)
Age range	8 to 11 years
Number of domains	17
Name of domains	Mobility, vision, hearing, speech, breathing, sleeping, eating, elimination, discomfort and symptoms, school and hobbies, friends, physical appearance, depression, anxiety, vitality, ability to concentrate, and learning ability and memory
Number of items	17
Rating scale	Select the level of functioning on a scale of 1–5
Respondent	Self-completed (interviewer administered)
Time to complete	20 to 30 minutes
Example item:	

The levels of the 'friends' dimension
Does your state of health make it more difficult for you to get friends or to be with friends?
1. Not at all
2. A little
3. Quite a lot
4. My state of of health makes getting friends or being with friends almost impossible
5. My state of health makes getting or being with friends totally impossible

Reliability	*Test–retest:* poor
Validity	*Clinical:* patient scores were significantly lower than that of the controls. In addition the 17D discriminated between transplant survivors and children with genetic skeletal dysplasias.

Quality of Well-Being scale

(QWB: Bradlyn *et al.*, 1993;[135] Kaplan, 1989[279])

Description	The QWB enables the computation of QALYs. Children are classified according to objective levels of functioning
Age range	4 to 16 years
Number of domains	3 domains and a list of 27 symptom-problem complexes
Name of domains	Physical functioning, social/role functioning and mobility
Number of items	3 and 27 symptom complexes
Rating scale	Level of functioning
Respondent	Interviewer administered to parent or adolescent
Time to complete	15 to 20 minutes
Example item:	

Mobility scale
- No limitations for health reasons
- Did not ride in a car, or had more help than usual to use public transportation than usual, for age
- In hospital, health related

Reliability	*Test–retest:* not reported for children *Internal consistency:* not reported for children
Validity	*Clinical validity:* distinguishes children with cancer differing in 'treatment toxicity'[135]

Appendix 5

Disease-specific measures of QoL

Asthma

About My Asthma

(AMA: Mishoe *et al.*, 1998[136])

Description	The AMA assesses stressors affecting QoL in children with asthma. It was developed by a multidisciplinary panel of experts
Age range	6 to 12 years
Number of domains	Total score only
Name of domains	Total score only
Number of items	44
Time to complete	15 to 20 minutes
Scale	4-point Likert scales (1 = child never thinks about the stressor; 4 = child thinks about the stressor all the time)
Time frame	None given
Respondent	Child
Administration	Interviewer
Example item:	

As a result of my asthma, I think about:
7. Not being able to do what I want to do in life 1 2 3 4
(1 = Never 2 = Once in a while 3 = Most of the time 4 = All the time)

Reliability	Not reported
Validity	*Internal consistency:* Cronbach's alpha = 0.93 *Concurrent:* moderate, negative correlation between AMA total score and PAQLQ emotional function domain ($r = -0.41$, $p < 0.05$) and PAQLQ overall score ($r = -0.51$, $p < 0.05$)

Childhood Asthma Questionnaires

(CAQ: Christie *et al.*, 1993;[138] French *et al.*, 1994[139])

Description	Examines children's perceptions of both active and passive aspects of living with asthma, together with their perceptions of its severity and any associated distress. There are three age-specific versions
Age range	4 to 7 years (CAQ A); 8 to 11 years (CAQ B); 12 to 16 years (CAQ C)
Number of domains	CAQ A:2; CAQ B: 4; CAQ C: 5
Name of domains	CAQ A: QoL, distress; CAQ B: active QoL, passive QoL, distress, severity; CAQ C: active QoL, teenage QoL, distress, severity, reactivity (derived from Principle axis factor analysis)
Number of items	CAQ A: 14; CAQ B: 22; CAQ C: 31
Rating scale	Smiley faces: CAQ A: 4 point; CAQ B: 4 point; CAQ C: 5 point
Respondent	Child completed
Time to complete	10–15 minutes
Example item:	

11a Do you wheeze (get a whistling noise in your chest)?
Colour one box Yes No

11b. Which face is you when you wheeze? Colour one face

Reliability	*Internal consistency:* CAQ B subscales had Cronbach alpha values of 0.78 (Active QoL), 0.57 (Passive QoL), 0.84 (Severity), 0.67 (Distress). Not reported for other versions *Test–retest:* good
Validity	*Clinical:* distinguishes between asthmatic and non-asthmatic children

Pediatric Asthma Quality of Life Questionnaire

(PAQLQ: Juniper *et al.*, 1996[140])

Development	The PAQLQ was developed to reflect the areas of function that are important to children with asthma. In addition to standardised questions, this measure includes three items concerning activities/hobbies that are individualised
Age range	7 to 17 years
Number of domains	3
Name of domains	Activity limitations, symptoms, and emotional function
Number of items	23
Rating scale	7-point Likert scale (1 maximum impairment – 7 no impairment)
Respondent	Child
Time to complete	10 to 15 minutes
Example item:	

How much did asthma make you feel FRUSTRATED during the past week?
Response options – [Green Card]
1. All the time
2. Most of the time
3. A bit of the time
4. Some of the time
5. A little of the time
6. Hardly any of the time
7. None of the Time

Reliability	*Test–retest:* ICC = 0.95 for stable patients
Validity	*Construct:* significant correlations between domains and clinical measures (FEV1. PEFR, B-agonist use) and domains and feeling thermometer scores

Cancer

Behavioural Affective and Somatic Experiences Scale

(BASES: Phipps *et al.*, 1994,[141] 1999[142])

Description	A biopsychosocial model forms the theoretical basis of this measure. Items were generated as a result of a search through the literature and from a group of BMT nurses
Age range	2 to 20 years
Number of domains	5
Name of domains	Somatic distress, compliance, mood disturbance, quality of interactions, and activity
Number of items	BASES-Parent 38, Nurse 38, Child 14
Rating scale	5-point Likert
Respondent	3 versions suitable for child, parent and nurse report
Time to complete	5 to 10 minutes (parent and nurse version); 2 to 5 minutes (child version)
Example item:	

BASES – PATIENT REPORT

	not at all	a little bit	somewhat	quite a bit	very much
Feeling scared or worried	1	2	3	4	5

Reliability	*Internal consistency:* nurse report – Cronbach's alphas for the subscales ranged from 0.74 to 0.90[141] Child report – total scale was 0.77[142] *Inter-rater:* using paired nurse observations was excellent (median $r = 0.866$). Parent–nurse comparisons median $r = 0.56$ (range 0.11 to 0.93)
Validity	*Clinical:* parent and nurse reports discriminated between patients undergoing allogeneic bone marrow transplantation (BMT) versus autologous BMT Detects differences in BMT outcomes by age, but not gender

The Miami Pediatric Quality of Life Questionnaire

(MPQOLQ: Armstrong *et al.*, 1999[94])

Description	Items were derived from interviews with families of children (aged 1 to 18) undergoing or within one year of completed therapy
Age range	1 to 18 years
Number of domains	3 (factors)
Name of domains	Social competence, emotional stability, and self-competence (derived from factor analysis)
Number of items	56
Rating scale	5-point Likert
Report	Parent (child version under development)
Time to complete	Not reported
Example item:	

Sometimes has problems getting along with others his/her same age

Reliability	*Test–retest*: (approx 1 month) acceptable for total QoL index ($r = 0.94$) and for self-competence ($r = 0.82$); and moderate for social competence ($r = 0.43$) and emotional stability ($r = 0.38$) *Internal consistency*: Cronbach's alpha ranged from 0.76 to 0.88 for the subscales and was 0.89 for the overall scale. Internal consistency of all individual items exceeded the criteria of Cronbach's alpha > 0.70
Validity	*Clinical*: there were significant differences between children with brain tumours versus leukaemia/lymphomas on social competence and self-competence, but not on the emotional stability scale. Also discriminated between children who had received cranial radiation versus no cranial radiation

Pediatric Cancer Quality of Life Inventory

(PCQL-32: Varni *et al.*, 1998[143,144])

Description	The PCQL is based on a definition of QoL as "the impact of disease and treatment on an individual's physical, social, and psychological and cognitive functioning". Items were generated from a search of the literature and open-ended interviews with patients and their families and discussions with healthcare professionals
Age range	8 to18
Number of domains	5
Name of domains	Physical functioning, disease-related and treatment symptoms, psychological functioning, social functioning, cognitive functioning
Number of items	32
Rating scale	4-point Likert scale ("never a problem" = 0 to "always a problem" = 3)
Time to complete	Not reported
Example item:	

Psychological functioning

	Never	Sometimes	Often	Always
I feel afraid	0	1	2	3

Reliability	*Internal consistency:* 0.91 patient form and 0.92 for the parent form
Validity	*Construct:* pattern of correlations between the PCQL-32 and the parent and patient-report standardised measures support the multidimensional conceptualisation of QoL *Clinical:* significant differences between patients on and off treatment on the total scale, as well as for the disease/treatment symptoms and physical functioning scales for both patient and parent report

Pediatric Oncology Quality of Life Scale

(POQOL:Goodwin *et al.*, 1994[145])

Description	Items were generated by health professionals, parents of cancer patients, and patients. The POQOL provides a total score and 3 factor scores
Age range	Preschool to adolescence (5 to 17 years)
Number of domains	3
Name of domains	Physical functioning and restriction from normal activity (Factor 1), emotional distress (Factor 2), response to active medical treatment (Factor 3) (derived from factor analysis)
Number of items	21
Rating scale	7-point Likert scale (1 = Never to 7 = Very frequently)
Respondent	Parent
Time to complete	Not reported
Example item:	

IN THE PAST TWO WEEKS	never		rarely	sometimes	often	very freq.	
8. My child has been embarrassed about physical changes (hair loss, weight change etc.)	1	2	3	4	5	6	7

Reliability	*Internal consistency:* good for total scale (Cronbach's alpha = 0.85)and three subscales (0.87, 0.79, 0.68) *Inter-rater:* mother-father (15 pairs) reliability for total scale $r = 0.87$, and subscales $r = 0.91$, 0.87, 0.75
Validity	*Concurrent:* subscale scores correlate as expected with PPSC, (0.60) CBCL (0.67), but not with the Reynolds Adolescent Depression Scale (0.60; $p > 0.05$) *Clinical:* total score and Factors 1 and 3 discriminated between two groups differing on time since diagnosis

Play Performance Scale for Children

(PPSC: Lansky *et al.*, 1985,[11] 1987[12])

Description	Based on the adult scale of global functional status used to assess QoL in adult cancer patients.[172] The PPSC allows for rapid assessment of performance in childhood
Age range	6 months to 16 years
Number of domains	1
Name of domains	Performance status (play)
Number of items	1
Rating scale	11-point scale: fully active, normal (100) to unresponsive (0)
Respondent	Parent
Time to complete	Not reported
Example item:	

100	fully active, normal
90	minor restrictions in physically strenuous activity
80	active, but tires more quickly
70	both greater restriction of, and less time spent in active play
60	up and around, but minimal active play; keeps busier with quieter activities

50	gets dressed, but lies around much of the day; no active play; able to participate in all quiet play and activities
40	mostly in bed; participates in quiet activities
30	in bed; needs assistance even for quiet play
20	often sleeping; play entirely limited to very passive activities
10	no play; does not get out of bed
0	unresponsive

Reliability	*Inter-rater:* Pearson correlation coefficient between mothers and fathers reports ($r = 0.71$; Lansky *et al.,* 1987)
Validity	*Construct:* correlations between parent rating and nurses' global rating ($r = 0.75$); and between parent and clinician's global rating ($r = 0.92$) *Clinical:* significant differences between inpatients and normal children, and inpatients and outpatients (Lansky *et al.,* 1987)

Epilepsy

Impact of Child Illness Scale

(Hoare & Russell, 1995[146])

Description	Measures the impact of epilepsy on the child and family. Parents are asked to answer each question in terms of frequency and importance
Age range	6 to 17 years
Number of domains	4
Name of domains	Impact of epilepsy and its treatment, impact on child's development and adjustment, impact on parents, and impact on the family
Number of items	30
Rating scale	3-point Likert to rate frequency and importance
Report	Parent
Time to complete	Not reported
Example item:	

6. My child is more moody because of his illness 0 1 2 A B C
(The first part refers to how frequently the problem occurs: never or rarely true = 0; sometimes true = 1; often or really true = 2. The second part refers to how much concern it causes: A = a lot of concern; B = a bit of concern; and C = not much concern)

Reliability	Not reported
Validity	Only face validity reported

Quality of Life in Epilepsy – Adolescent Version

(QOLIE-AD-48: Cramer *et al.*, 1999[149])

Description	Items were derived from several sources, including adolescents themselves. The measure consists of items that tap generic and epilepsy-specific issues
Age range	11 to 17 years
Number of domains	8
Name of domains	Epilepsy impact, memory/concentration, attitudes, physical function, stigma, social support, school behaviour, and health perceptions
Number of items	48
Rating scale	5-point Likert scale
Respondent	Adolescent
Time to complete	Not reported
Example item:	

In the past 4 weeks, how often did you:
29. Feel embarrassed or "different" because you had to take medications?
(Very often = 1, Often = 2; Sometimes = 3; Not often = 4; Never = 5)

Reliability	*Internal consistency:* satisfactory (all subscales > 0.70 except for health perceptions) *Test–retest:* 0.83 for the total score
Validity	*Construct:* the total score correlated moderately to strongly with self-esteem and self-efficacy scales

Arthritis

The Juvenile Arthritis Quality of Life Questionnaire

(JAQLQ: Duffy *et al.*, 1997[150])

Description	The JAQLQ measures physical and psychosocial function and an array of general symptoms
Age range	2 to 18
Number of domains	4
Name of domains	Gross motor function, fine motor function, psychosocial function, and general symptoms
Number of items	74
Rating scale	7-point Likert scale
Respondent	Child report (9 years and above)
Time to complete	5 minutes
Example item:	

How often have you/your child, over the past 2 weeks, had difficulty with the following activities as a result of arthritis or its treatment? (e.g. getting out of bed)
0 = does not apply to 7 = all of the time

Reliability	Not reported
Validity	*Construct*: moderate correlations between JAQLQ and active joint count; stronger correlations were observed with measures of pain *Responsiveness*: discriminated among patients based on physician global assessment of change

Crohn's disease

Quality of Life in Children with Crohn's Disease

(Rabbett *et al.*, 1996[151])

Description	Multidimensional measure of QoL. Requires further work before it can be widely used
Age range	8 to 17
Number of domains	6
Name of domains	Disease and its treatment, social, emotional, family, education, and future aspects
Number of items	88
Rating scale	Not reported
Report	Child report
Administration	Interviewer
Time to complete	Not reported
Example item	Not given
Reliability	Not reported
Validity	Parents views of the severity of symptoms significantly correlated with their children's views regarding rectal bleeding, poor growth, lack of energy and poor appetite

Dermatology

Children's Dermatology Life Quality Index

(CDLQI: Lewis-Jones & Finlay, 1995[152])

Description	Based on the adult Dermatology Life Quality Index,[173] the CDLQI was designed to be of practical use within a busy clinic
Age range	3 to 16 years
Number of domains	6
Name of domains	Symptoms and feelings, leisure, school holidays, personal relationships, sleep and treatment
Number of items	10
Rating scale	4-point Likert (0 = not at all; 3 = very much)
Report	Self report (parental help with younger children)
Time to complete	Not reported
Example item:	

1. Over the last week, how **embarrassed** or **self conscious**, or **upset** or **sad** have you been because of your skin?	Very much	Quite a lot	Only a little	Not at all

Reliability	*Test–retest:* 0.86
Validity	Not assessed

Diabetes

Diabetes Quality of Life for Youths

(DQOL-Y: Ingersoll & Marrero, 1991[153])

Description	The DQOL was developed to assess perceptions of the impact of an intensified regimen on general satisfaction with life and concerns over social and vocational issues related to diabetes. The version for youths was derived from the DQOL for adults
Age range	11 to 18 years
Number of domains	3 (scales)
Name of domains	Disease impact, disease-related worries, and satisfaction with life
Number of items	52
Rating scale	5-point Likert
Report	Self
Time to complete	Not reported
Example item:	

8. How often do you feel good about yourself?	1	2	3	4	5
(1 = never; 2 = very seldom; 3 = sometimes; 4 = often; 5 = all the time)					

Reliability	*Internal consistency:* Cronbach's alpha = 0.85 *Test–retest:* not reported
Validity	Not reported

Headache

Quality of Life Headache in Youth

(QLH-Y: Langeveld *et al.*, 1997,[154] 1999[155])

Description	Derived from existing measurement scales and interviews with migraine patients. This measure was developed for adolescents with migraine or severe headache. Many of the subscales have a generic character, which enables comparisons with other groups
Age range	12 to 18 years

Number of domains	4
Name of domains	Physical functioning, functional status, psychological functioning, and social functioning
Number of items	71
Rating scale	4-point Likert scale and visual analogue (100 mm) to cover the QoL subdomains: 'completely satisfied' to 'completely dissatisfied'
Report	Parent or adolescent
Time to complete	Not reported
Example item:	

My headache bothered me while doing my homework
Rarely/never = 0, sometimes = 1, often = 2, very often = 3

Reliability	*Test–retest: acceptable:* 0.60 (one week interval); 0.47 (6 month interval) *Internal consistency:* satisfactory Cronbach's alpha > 0.70
Validity	*Clinical:* all subscales, except two of the three social functioning subscales discriminated between the headache group and the control group

Neuromuscular disorders

Life Satisfaction Index for Adolescents

(LSIA: Reid & Renwick, 1994[156])

Description	The LSIA is a multidimensional measure of life satisfaction for adolescents with neuromuscular disorders
Age range	12 to 19 years
Number of domains	5
Name of domains	General well being, interpersonal relationships, personal development, personal fulfilment, leisure and recreation
Number of items	35
Rating scale	5-point rating scale (1 = strongly agree to 5 = strongly disagree)
Respondent	Child
Time to complete	Not reported
Example item:	

I am satisfied with the chances I have to try new things
Strongly disagree | Disagree | Neither agree nor disagree | Agree | Strongly agree | Not applicable

Reliability	*Internal consistency:* Cronbach's alpha = 0.52 to 0.88 for the 5 domains
Validity	*Construct:* correlated significantly with the a measure of self-esteem ($r = -0.078$)

Otitis media

Quality of Life for Children with Otitis Media

(OM-6: Rosenfeld *et al.*, 1997[157])

Development	The OM-6 is brief and easy to administer multidimensional measure. Content was based on interviews with caregivers and experts in the area
Age range	6 months to 12 years
Number of domains	6
Name of domains	Physical suffering, hearing loss, speech impairment, emotional distress, activity limitations, caregiver concerns
Number of items	6
Rating scale	7-point Likert (and 10-point overall QoL scale)
Respondent	Parent
Time to complete	30 seconds – 3 minutes

Example item:

Emotional distress: Irritable, frustrated, sad, restless, poor appetite. How much of a problem for your child during the past 4 weeks as a result of ear infections or fluid?

[ ] Not present/no problem [ ] Hardly a problem at all [ ] Somewhat of a problem [ ] Moderate problem
[ ] Quite a bit of a problem [ ] Very much a problem [ ] Extreme problem

Reliability	*Internal consistency*: not reported *Test–retest*: good (>0.70 for all items)
Validity	*Construct*: good correlation between the survey score and a global ear-related QoL rating *Sensitivity*: large sensitivity to clinical change for all OM-6 items after bilateral insertion of tympanostomy tubes

Rhinoconjunctivitis

Paediatric Rhinoconjunctivitis Quality of Life Questionnaire

(PRQLQ: Juniper *et al.*, 1994[158])

Description	Based on the adult rhinoconjunctivitis QoL Questionnaire. There are two versions: one for children and one for adolescents. It is interviewer administered
Age range	Child version: 6 to 12 years; adolescent version: 12 to 17 years
Number of domains	Child version: 5; adolescent version: 6
Name of domains	Child version: nose symptoms, eye symptoms, practical problems, other symptoms, and activity limitations Adolescent version: Practical problems, non-hay fever symptoms, nose symptoms, eye symptoms, patient specific activities, and emotions
Number of items	Child: 23; adolescent: 25
Rating scale	7-point Likert scale
Report	Parent or adolescent
Time to complete	Not reported
Example item:	

(Child version)
How often did your allergies make you feel IRRITABLE (cranky/grouchy) during the past week?
Green response card
1. All the time
2. Most of the time
3. Quite often
4. Some of the time
5. Once in a while
6. Hardly any of the time
7. None of the time

Reliability	*Internal:* ICC = 0.93 *Test–retest:* the PRQLQ picked up no change in QoL in those with stable rhinoconjunctivitis, but did detect change in those children whose rhinoconjunctivitis changed between visits[177]
Validity	The Pearson correlation coefficients between the QoL domains and diary symptom scores were very close to those predicted[158] QoL scores improved with treatment[158]

Short stature

Quality of Life in Children with Short Stature

(Pilpel *et al.*, 1995[159])

Description	Developed to assess QoL in short stature children who are on growth hormone therapy (GHT) with or without underlying disease, and short stature children who are not on GHT
Age range	8 to 18
Number of domains	5
Name of domains	Academic achievement level, leisure activities, physical self esteem, emotional self esteem, relationships with peers and family members
Number of items	45
Rating scale	4-point Likert scale
Report	Child
Time to complete	Not reported
Example item:	

Are you a good student?
Response on a scale from 1 (very good student) to 4 (very bad student)

Reliability	Not reported
Validity	Not reported

Spina bifida

Quality of Life in Spina Bifida Questionnaire

(Parkin *et al.*, 1997[160])

Description	Developed from the viewpoint of the children and their parents
Age range	5 to 12 and 13 to 20
Number of domains	10
Name of domains	Social, emotional, intellectual, financial, medical, independence, environmental, physical, recreational, and vocational
Number of items	44 items (5–12 years); 47 items (13–20 years)
Rating scale	5-point Likert
Report	5 to 12, parent 13 to 20, self-report
Time to complete	Not reported
Example item:	

How much do you feel your child:
Has the opportunity to do everything the other children do at school 1 2 3 4 5

Reliability	*Test–retest:* (2 weeks) ICC 0.78 (child) 0.96 (adolescent) *Internal consistency:* Cronbach's alpha demonstrated good internal consistency for the 5–12-year-olds (0.93) and for the 13–20-year-olds (0.94)
Validity	*Construct validity:* for children HRQOL and global question of child's well-being was $r = 0.63$ ($p = 0.0001$). For the 13 to 20 year age group the correlation between HRQOL and the global question was $r = 0.37$ ($p = 0.0001$)

Spine deformities

Quality of Life Profile for Spine Deformities

(QLPSD: Climent *et al.*, 1995[161])

Description	The QLPSD conceptualises QoL in terms of psychosocial dimensions, pain and function. It was based, in part, on interviews with patients
Age range	10 to 20 years
Number of domains	5
Name of domains	Psychosocial functioning, sleep disturbances, body image, back flexibility, and back pain
Number of items	21
Rating scale	5-point Likert ('strongly agree' to 'strongly disagree'; 'very satisfied' to 'very dissatisfied')
Report	Self
Time to complete	10 minutes (average)
Example item:	

I worry a lot that my back will affect my life in the future

Reliability	*Test–retest:* 0.91 (10 days) *Internal consistency:* 0.88 (0.70 to 0.84 for subscales)
Validity	Varimax loadings for the 5 components of the QLPSD revealed a unifactorial domain of QoL (variance = 51.6%)

Methodology Group

Members

Methodology Programme Director
Professor Richard Lilford
Director of Research and Development
NHS Executive – West Midlands, Birmingham

Chair
Professor Martin Buxton
Director, Health Economics Research Group
Brunel University, Uxbridge

Professor Douglas Altman
Professor of Statistics in Medicine
University of Oxford

Dr David Armstrong
Reader in Sociology as Applied to Medicine
King's College, London

Professor Nicholas Black
Professor of Health Services Research
London School of Hygiene & Tropical Medicine

Professor Ann Bowling
Professor of Health Services Research
University College London Medical School

Professor David Chadwick
Professor of Neurology
The Walton Centre for Neurology & Neurosurgery
Liverpool

Dr Mike Clarke
Associate Director (Research)
UK Cochrane Centre, Oxford

Professor Paul Dieppe
Director, MRC Health Services Research Centre
University of Bristol

Professor Michael Drummond
Director, Centre for Health Economics
University of York

Dr Vikki Entwistle
Senior Research Fellow, Health Services Research Unit
University of Aberdeen

Professor Ewan B Ferlie
Professor of Public Services Management
Imperial College, London

Professor Ray Fitzpatrick
Professor of Public Health & Primary Care
University of Oxford

Dr Naomi Fulop
Deputy Director, Service Delivery & Organisation Programme
London School of Hygiene & Tropical Medicine

Mrs Jenny Griffin
Head, Policy Research Programme
Department of Health
London

Professor Jeremy Grimshaw
Programme Director
Health Services Research Unit
University of Aberdeen

Professor Stephen Harrison
Professor of Social Policy
University of Manchester

Mr John Henderson
Economic Advisor
Department of Health, London

Professor Theresa Marteau
Director, Psychology & Genetics Research Group
Guy's, King's & St Thomas's School of Medicine, London

Dr Henry McQuay
Clinical Reader in Pain Relief
University of Oxford

Dr Nick Payne
Consultant Senior Lecturer in Public Health Medicine
ScHARR
University of Sheffield

Professor Joy Townsend
Director, Centre for Research in Primary & Community Care
University of Hertfordshire

Professor Kent Woods
Director, NHS HTA Programme, & Professor of Therapeutics
University of Leicester

continued

HTA Commissioning Board

Members

Programme Director
Professor Kent Woods
Director, NHS HTA Programme, &
Professor of Therapeutics
University of Leicester

Chair
Professor Shah Ebrahim
Professor of Epidemiology of Ageing
University of Bristol

Deputy Chair
Professor Jon Nicholl
Director, Medical Care Research Unit
University of Sheffield

Professor Douglas Altman
Director, ICRF Medical Statistics Group
University of Oxford

Professor John Bond
Director, Centre for Health Services Research
University of Newcastle-upon-Tyne

Ms Christine Clark
Freelance Medical Writer
Bury, Lancs

Professor Martin Eccles
Professor of Clinical Effectiveness
University of Newcastle-upon-Tyne

Dr Andrew Farmer
General Practitioner &
NHS R&D Clinical Scientist
Institute of Health Sciences
University of Oxford

Professor Adrian Grant
Director, Health Services Research Unit
University of Aberdeen

Dr Alastair Gray
Director, Health Economics Research Centre
Institute of Health Sciences
University of Oxford

Professor Mark Haggard
Director, MRC Institute of Hearing Research
University of Nottingham

Professor Jenny Hewison
Senior Lecturer
School of Psychology
University of Leeds

Professor Alison Kitson
Director, Royal College of Nursing Institute, London

Dr Donna Lamping
Head, Health Services Research Unit
London School of Hygiene & Tropical Medicine

Professor David Neal
Professor of Surgery
University of Newcastle-upon-Tyne

Professor Gillian Parker
Nuffield Professor of Community Care
University of Leicester

Dr Tim Peters
Reader in Medical Statistics
University of Bristol

Professor Martin Severs
Professor in Elderly Health Care
University of Portsmouth

Dr Sarah Stewart-Brown
Director, Health Services Research Unit
University of Oxford

Professor Ala Szczepura
Director, Centre for Health Services Studies
University of Warwick

Dr Gillian Vivian
Consultant in Nuclear Medicine & Radiology
Royal Cornwall Hospitals Trust
Truro

Professor Graham Watt
Department of General Practice
University of Glasgow

Dr Jeremy Wyatt
Senior Fellow
Health Knowledge Management Centre
University College London

Current and past membership details of all HTA 'committees' are available from the HTA website (see inside front cover for details)

Feedback

The HTA programme and the authors would like to know your views about this report.

The Correspondence Page on the HTA website (http://www.ncchta.org) is a convenient way to publish your comments. If you prefer, you can send your comments to the address below, telling us whether you would like us to transfer them to the website.

We look forward to hearing from you.

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 23 8059 5639 Email: hta@soton.ac.uk
http://www.ncchta.org

ISSN 1366-5278