Development of a child-centred, caries-specific measure of oral health-related quality of life

Submitted in accordance with the requirements
for the degree of Doctor of Philosophy

University of Sheffield

School of Clinical Dentistry

Fiona Gilchrist

November 2015

Acknowledgements

Firstly I would like to thank my amazing supervisors, Dr Zoe Marshman, Professor Helen Rodd and Professor Chris Deery, for their dedication to the project, their enthusiasm and their friendship. Their support and collective wisdom has been invaluable over the past seven years, both academically and personally. The unwavering encouragement they provided helped me to negotiate the often difficult path of combining clinical training with academia, allowing me to succeed at both.

My gratitude to the National Institute for Heath Research for funding the project over the last three years, this has allowed me the opportunity to fully commit my time to the study and through this I have been able to challenge myself and learned more than I ever thought possible. I owe a debt of thanks not only to my supervisors but also to Professor Paul Speight, Professor Sarah Baker and Professor Peter Robinson for their advice and support with my NIHR Doctoral Fellowship application.

The participants and their families deserve special mention for their enthusiasm and the insights they have granted me. To all in the Department of Paediatric Dentistry, your support and encouragement has been vital. My gratitude to all who were involved in recruitment to the study, especially Claire Bower, Vanessa Wilson, Jenny Harris and Clair Jenkinson. To all in the Units of Oral Health and Development and Dental Public Health, I appreciate the encouragement, advice and support which has been so generously given. My particular thanks to Tom Broomhead and Ninu Paul for their help with transcription. Further thanks to Professor Sarah Baker for her insights and statistical expertise.

I am hugely grateful to my parents for a lifetime of support and understanding, their willingness to visit Sheffield and for looking after me on visits “home”. To my family and friends, thanks for the encouragement, conversation and fun, it’s been invaluable. Finally, to Gerry, my very best friend and open water swimming buddy, my thanks for technical advice; reference bingo; Scrabble; kindness; laughter and for bravely taking the decision to move to Sheffield during the final stages of this work!

Abstract

Background: Existing oral health-related quality of life (OHRQoL) measures are generic and have not involved children at all stages of development.

Aim: To develop a caries-specific measure of OHRQoL for children.

Design: The first phase involved a systematic review of the three most commonly used child self-report measures of OHRQoL. This was followed by a qualitative study with children, aged 5-16 years, to develop the measure and, finally, a cross-sectional validation study. Necessary ethical approval was granted for the study.

Results: The systematic review included 120 papers and revealed that the three most commonly used existing measures had included children only in the latter stages of development. There was lack of testing for unidimensionality, although other properties were satisfactory. The qualitative study found that children discussed a number of caries-related impacts which affected their daily lives. These were incorporated into a draft measure which was further refined following testing of face and content validity. The questionnaire for validity testing contained 16 items and one global question and was named the Caries Impacts and Experiences Questionnaire for Children (CARIES-QC). Two hundred participants with a mean (range) age of 8.1 (5-16) years took part in the evaluation of CARIES-QC. Four items, which did not fit the Rasch model, were removed from further analysis. The remaining 12 items demonstrated good internal consistency (alpha=0.9) and the total score showed significant correlations with the number of decayed teeth, presence of pain, pulpal involvement, the Child Perceptions Questionnaire (16-item short form) and the global score ($p<0.01$, Spearman's rho).

Conclusion: CARIES-QC demonstrates acceptable validity, reliability and responsiveness using both modern psychometric techniques and Classical Test Theory. Its unidimensionality allows the transformation of raw scores, enabling accurate calculation of effect sizes and change scores following treatment of dental caries.

Frequently used abbreviations

CARIES-QC	Caries Impacts and Experiences Questionnaire for Children
CAT	Computer adaptive testing
CAT-QoL	Child Amblyopia Treatment Questionnaire
CB	Claire Bower, Research Nurse
CCDH	Charles Clifford Dental Hospital
CD	Chris Deery, Professor of Paediatric Dentistry
CHU9D	Child Health Utility 9D
CIS	CARIES-QC interval scale
COHIP	Child Oral Health Impact Profile
C-OIDP	Child Oral Impacts on Daily Performances
COSMIN	Consensus-based Standards for the selection of health Measurement Instruments
CPQ	Child Perceptions Questionnaire
CPQ$_{8-10}$	Child Perceptions Questionnaire for 8-10-year-olds
CPQ$_{11-14}$	Child Perceptions Questionnaire for 11-14-year-olds
CPQ$_{11-14}$-ISF:16	Child Perceptions Questionnaire, 16-item item impact version
CTT	Classical Test Theory
DIF	Differential item functioning
dmft	Decayed, missing and filled primary teeth
DMFT	Decayed, missing and filled permanent teeth
FG	Fiona Gilchrist, Principal Investigator
HR	Helen Rodd, Professor of Paediatric Dentistry
HRQoL	Health-related quality of life
ICC	Intraclass correlation coefficient
IMD	Index of Multiple Deprivation
IRT	Item response theory
JH	Jenny Harris, Consultant in Paediatric Dentistry
MID	Minimal important difference
NHSDEP	National Health Service Dental Epidemiological Programme
OHRQoL	Oral health-related quality of life
PBM	Preference based measure
PMC	Preformed Metal Crown
PREM	Patient Reported Experience Measure
PROM	Patient Reported Outcome Measure
PSI	Person Separation Index
SOHO-5	Scale of Oral Health Outcomes
UK	United Kingdom
VW	Vanessa Wilson, Research Nurse
XGA	Extractions under general anaesthetic
ZM	Zoe Marshman, Senior Lecturer in Dental Public Health

Table of Contents

Chapter Six Development of the measure 108

List of appendices

List of Figures

List of Tables

Chapter One

Introduction

Dental caries is the most prevalent oral disease in children. A recent systematic review reported that approximately 621 million children globally have untreated dental caries (Kassebaum et al., 2015). Despite the prevalence of the disease, there is a paucity of evidence relating to children's perspectives of dental caries or interventions which may be used in its management.

To date, children's perspectives have been captured by the use of child self-report measures of oral health-related quality of life (OHRQoL). However, these measures are designed to capture a range of impacts associated with different orofacial conditions and therefore may not be sensitive enough to assess those impacts which are specifically associated with dental caries (Wiebe et al., 2003). In addition, these measures of OHRQoL only sought children's input at the latter stages of their development and therefore they may not accurately reflect the impacts which are important to them. Furthermore, none of the existing measures have been validated for longitudinal evaluation of the effects of interventions which may be used to manage dental caries. To address these acknowledged limitations, a caries-specific measure of OHRQOL which has fully involved children in its development is required for use in future clinical trials.

Therefore the aim of the research presented in this thesis is to conduct three interlinked studies which will lead to the development of a caries-specific measure of OHRQoL for children.

The objectives of the research are to:

1. Critically review the most commonly used self-report measures of OHRQoL for children against existing quality criteria, the findings of which will be used to inform the development of the new measure.
2. Explore, through qualitative methods, the impacts of dental caries on children.
3. Involve children in the design and content of the measure to ensure that it is meaningful and relevant to them.
4. Test the measure for validity, reliability and responsiveness using modern psychometric techniques and existing criteria.

The thesis is structured as follows:

- Chapter Two is a narrative review of the literature which considers the public health significance of dental caries in terms of its prevalence, its impact on the individual, its impact on society and the availability of preventive measures and various treatment regimens for dental caries. Gaps in the literature regarding children's perspectives of dental caries and interventions used for its management are highlighted.
- Chapter Three describes the rationale, aim and objectives of the research.
- Chapter Four discusses the methodological considerations pertinent to undertaking research *with* children and when developing a new measure of OHRQoL
- Chapter Five reports a systematic review of the three most frequently used child self-report measures of OHRQoL and evaluates their development and validation against existing quality criteria. The review highlights the limited involvement of children in the development of previous measures and lack of modern psychometric techniques to assess their unidimensionality.
- Chapter Six details development of the caries-specific measure, involving children throughout the development process. Impacts reported by children along with the language they used are described. The benefits of actively engaging children throughout development are highlighted.

- Chapter Seven reports validity, reliability and responsiveness testing of the new measure. Both modern psychometric techniques and Classical Test Theory are used to assess the measure’s properties.
- Chapter Eight brings together the findings of the studies contained within the thesis and discusses their implications for clinical care, policy and further research.
- Chapter Nine presents the key conclusions and recommendations which have arisen from the three studies.

Chapter Two

Background

This chapter will discuss the prevalence and impact of dental caries in children and assess the public health significance of dental caries using the criteria suggested by Sheiham and Watt (2003).

2.1 Introduction

Dental caries is the most prevalent oral disease amongst children and can cause chronic pain, local infection and in some cases may lead to hospitalisation due to spread of the infection and systemic illness (Majewski et al., 1988; Benjamin, 2010). A recent systematic review found that globally 621 million children have untreated caries in their primary teeth, reaching peak prevalence at 6 years of age (Kassebaum et al., 2015). No significant decrease in the incidence of the disease was seen from 1990 to 2010. It would therefore appear that the global burden of dental caries is significant. The following sections will discuss the public health significance of caries according to the four criteria suggested by Sheiham and Watt (2003):

1. Prevalence: should be high, or if the disease is uncommon, then it should be serious.
2. Impact of the condition on the individual: the symptoms should have a psychological and social impact on the individual and affect functioning.
3. Impact on wider society: for example, costs of treatment, time off school or work for the population.
4. The condition is preventable and effective treatments are available.

2.2 Public health significance of caries

This section will mainly focus on the prevalence of caries in the UK, as the study will be set in a city in England (Sheffield). The UK has a Human Development Index of 0.89, which is makes it a country with "very high human development" and its Gross Domestic Product (a measure of economic performance) is ranked 5th in the world (United Nations Development Programme, 2013; Knoema, 2015). However, compared to other developed countries, the UK has a high level of income inequality. It is the 4th most unequal in terms of income amongst developed countries and the most unequal in Europe (The Equality Trust, 2015). Sheffield is the 56th most deprived local authority in England (out of 326) (Sheffield City Council, 2010). In Sheffield, 34% of the population live in areas which are amongst the 20% most deprived in England, with 22% living in areas which are amongst the 10% most deprived in England. At the other end of the deprivation scale, 20% live in areas ranked amongst the 20% least deprived areas in England and 8% live in areas in the 10% least deprived (Rae, 2011). Sheffield has a higher proportion of children living in poverty (25%) than the England average (21%) (Public Health England, 2015). Therefore, although Sheffield is not one of the most deprived cities in England due the presence of both affluent and deprived areas, it is one of the most unequal with clear divides between the North East and South West of the city (Sheffield Fairness Commission, 2013). These divides can also be seen when data related to the proportion of children with dental caries are extrapolated for Sheffield wards. For example, within areas in the most deprived quintile, almost 50% of children have caries experience, compared with less than 20% in the least deprived quintile (Public Health England, 2014c). Further details regarding caries experience of children in Sheffield will be discussed in Section 2.2.1.

The next section will discuss the prevalence of dental caries globally and in the UK.

2.2.1 Prevalence of dental caries in children

Dental caries is the most common dental disease in children and young people with untreated caries in primary teeth affecting 9% of children worldwide (Kassebaum et al., 2015). Indeed, data from North America show that caries is five times more prevalent in children than asthma and seven times more prevalent than hayfever (U.S. Department of Health and Human Services, 2000).

Global prevalence

In order to standardise reporting of dental caries prevalence around the world, caries experience is generally expressed as a mean of the number of decayed, missing and filled teeth (abbreviated to DMFT in the permanent dentition and dmft in the primary dentition) for the population (Bodecker, 1939). Criticisms of this method of caries detection are that unless the individual components are expressed individually, it is not clear what each contributes to the overall score. In addition, there are differences in how the "decayed" component is assessed, with some assigning this score only when there is frank cavitation and others where there is visual evidence that the caries extends into dentine. Furthermore, as only lesions into dentine are reported, there may be an underestimation of the prevalence of caries as it fails to account for lesions confined to enamel. More recently developed systems such as the International Caries Detection and Assessment System (ICDAS II), use a more refined measurement system which allows the extent of caries to be more accurately measured (International Caries Detection and Assessment System Coordinating Committee, 2011). The ICDAS II classifies six stages of caries, from the first white spot lesion in dry enamel (code 1) to extensive cavitation involving over half the tooth surface (code 6) (International Caries Detection and Assessment System Coordinating Committee, 2011). Using ICDAS II in epidemiological studies may cause difficulties in comparison with other studies as the majority have used the DMF system (Iranzo-Cortes et al., 2013). However, there are methods which can be used to convert between the two, although there remains some debate regarding how best to do this (Iranzo-Cortes et al., 2013). There remains a need for more sensitive assessment techniques to be adopted to ensure that results are comparable across countries and settings and to provide more accurate results of interventions to prevent progression of dental caries. However, caries data presented in this present study will be as dmft/DMFT as this is still commonly presented in epidemiological studies.

The World Health Organization's (WHO) Global Oral Health Database shows that caries experience in 12-year-olds varies widely throughout the world, with a mean weighted DMFT of 1.86 (World Health Organization, 2015b). The area with the highest mean caries experience in 2015 was South East Asia (DMFT=2.97) and the lowest was the Western Pacific region (DMFT=1.05). Seventy-three percent of countries had a mean DMFT of less than 3 in 12-year-olds which is the level set by the WHO as a global oral health goal. However, within these continents there are large variations between different countries. For example, in 2004 the mean DMFT was 6 in St. Lucia, but 0.6 in Trinidad and Tobago, although direct comparison can be difficult due to differing protocols used for dental surveys (World Health Organization, 2015b).

The difference in caries prevalence between countries appears to be associated with economic and social development. Do (2012) described analysis of global caries trends, indicating that globally caries has moved from being a disease of affluence to a disease associated with deprivation (Do, 2012). Caries prevalence was higher in developed countries pre-1980, however these areas have shown a more rapid decline in the disease over time (Dye et al., 2007; Do, 2012). This trend has been attributed to the availability of fluoride toothpaste, public health measures, improvements in disease management and better living conditions (Spencer et al., 1996; Petersen, 2003; Do, 2012). However, there has been little change in caries experience in countries with lower economic and human development (Do, 2012). As these communities account for the majority of the world's population, strategies for improving oral health in these populations are required.

Prevalence in the UK

National surveys in the United Kingdom (UK) provide information on the prevalence of the disease and are designed to allow analysis of oral health trends. One such survey is the Children's Dental Health Survey which has been carried out every ten years since 1973. The 2013 survey revealed that 31% of 5-year-olds and 46% of 8-year-olds have obvious caries experience (Pitts et al., 2015). The mean dmft was 0.9 and 1.4 for 5- and 8-year-olds respectively. However, this increased to 3.0 in those with dental caries, which is a more realistic estimate of the likely burden of the disease in those affected. Due to changes in the consent process between this most recent survey and the preceding ones, it is not possible to make direct comparisons with findings from previous surveys for these age groups. Almost half (49%) of 5-year-olds and 56% of 8-year-olds had clinical evidence of caries (i.e. enamel caries), indicating preventive measures are required to inhibit progression (Pitts et al., 2015). There remains a large variation for the proportion of affected 5-year-olds between England (31%), Wales (41%) and Northern Ireland (40%) (Pitts et al., 2015). The same is true for 8-year-olds, with 45% of English 8-year-olds having obvious caries experience compared with 55% in Wales and 56% in Northern Ireland. Children who were eligible for free school meals had a higher prevalence of obvious caries experience (41% of 5-year-olds and 57% of 8-year-olds) when compared to those who are not eligible (29% of 5-year-olds and 45% of 8-year-olds) (Pitts et al., 2015). These data suggest that those with higher levels of social deprivation have higher levels of disease, in common with other previous studies (Lader et al., 2004; Schwendicke et al., 2015).

Obvious caries experience in the permanent dentition was present in 13%, 34% and 46% of 8-, 12- and 15-year-olds respectively in the 2013 Child Dental Health Survey (Pitts et al., 2015). Again, a higher proportion of children who were eligible for school meals were affected. However, the prevalence of caries in the permanent dentition has been shown to decrease over the last 20 years with the proportion of 15-year-olds having obvious caries experience (excluding non-cavitated lesions) decreasing from 93% in 1983 to 42% in 2013. Similar decreases have been seen in 12-year-olds (83% in 1983 to 28% in 2013). However, a large percentage of the population still remain affected by the disease and this is heavily skewed towards those with a lower socioeconomic status (Pitts et al., 2015).

In addition to this decennial survey, the NHS Dental Epidemiological Programme (NHS DEP) performs a series of epidemiological surveys. These surveys examine 5-year-olds biennially and other population groups in the intervening years and allow analysis at regional level. In 2011/12 the average dmft in England was 0.94, varying from 0.67 in the South East to 1.29 in the North West of the country and there was an association with deprivation (Public Health England, 2012). The mean dmft in children who had caries experience was 3.38. Up to 13% of 5-year-old children had evidence of sepsis (defined as the presence of a sinus or dental abscess on visual examination of the soft tissues) from carious teeth and the disease remains largely untreated especially in the primary dentition (Public Health England, 2012). The prevalence of sepsis overall was 1.7%, and, as expected was higher in areas where caries experience was high. The caries experience of 5-year-old children in Yorkshire and the Humber region was above average, with a mean dmft of 1.23 and affecting 33.6% of the population. The area with lowest caries experience was North Lincolnshire (dmft=0.60) ranging to the highest in Bradford (dmft=1.98). Locally, Sheffield was reported to have an average dmft of 1.30, with caries affecting 35.8% of 5-year-olds. Similar studies are performed in other areas of the UK. In Wales, the average dmft of 5-year-olds in 2011/12 was 1.59 affecting 41.4% of the population. In Scotland in 2012, there was an average dmft of 1.35 with 33% of the Primary 1 (aged 4-7 years, mean 5.5 years) population affected (Macpherson et al., 2012; Morgan et al., 2013).

In 2008/2009, 33.4% of 12-year-olds in England had evidence of caries experience with an average DMFT of 0.74, although this rose to 2.21 in those with caries experience (Rooney et al., 2010). Again there was significant variation between regions, with children in the North West, North East and Yorkshire and the Humber having mean DMFT above the national average. In Yorkshire and the Humber the average DMFT ranged from 0.74 in North Lincolnshire to 1.37 in Bradford and Airedale with a regional mean of 1.07. In Sheffield, 41.4% of 12-year-olds had caries experience with a mean DMFT of 0.97. Those with caries experience had a mean DMFT of 2.35. During the same period, 11-12-year-olds in Wales were found to have an average DMFT of 0.98, with those with caries experience (42.5%) having a mean DMFT of 2.31. Both of these surveys found reductions in the prevalence and severity of dental caries compared with previous national surveys (Welsh Oral Health Information Unit and Public Health Wales, 2010).

In summary, it would appear that around one-third of children in England have caries experience although there is great regional variation with over 50% being affected in some areas. There is also evidence of an association with deprivation, with a higher prevalence and an increased number of teeth affected in more deprived areas.

The next section will consider the impact of caries on affected children.

2.2.2 Impact on individuals

A number of studies have investigated the impact of caries in children, however, many of these studies have used proxies rather than questioning the children themselves. Studies have understandably tended to focus on pain as the main impact of dental caries. To date, little insight has been gained with regard to the psychosocial aspects of dental disease in children. In this section, the prevalence of dental pain in children will be discussed followed by a review of other caries-related impacts. Studies which have investigated psychosocial status in relation to dental caries will also be reviewed.

Prevalence of dental pain in children

Estimates by proxies

Dental pain was reported in 32% (n=7.5 million) of children, whose parents stated they had caries, in a North American (USA) national health survey (Lewis and Stout, 2010). This prevalence is higher than that found in the most recent UK Child Dental Health Survey, where toothache was reported by parents to have occurred in 14% of 5-year-olds and 18% of 8-year-olds (Tsakos et al., 2015). A study in a clinical population involving children attending for treatment of caries under general anaesthesia (GA) in the North West of

England found a higher prevalence of pain (69%). Of those with reported symptoms, 45% had experienced pain in the previous month and 27% in the previous week (Goodwin et al., 2015). These children had specifically been referred for treatment of their carious lesions and thus provide data for pain experience in a clinical population, rather than the wider population.

It should be acknowledged that these studies may have underestimated the true prevalence of pain as they were based on parental report. Research in the medical field investigating the prevalence of pain in children has also often been based on parental reports (Bille, 1962; Evans et al., 2004; Zernikow et al., 2005; Kröner-Herwig et al., 2007). However, although there are similarities between parental and child responses, they are not identical (Schneider and LoBiondo-Wood, 1992; Doherty et al., 1993; Zernikow et al., 2005; Sundblad et al., 2006).

Child self-reported pain estimates

Several studies have investigated dental pain from the child's perspective using a variety of self-report measures. In Thailand, a national oral health survey found that 39% and 34% of children aged 12 and 15 years respectively, attributed the impacts they were suffering to dental pain (Krisdapong et al., 2009). This is lower than data emanating from Brazil, where a lifecourse study involving 339 children found that between 36-71% of children with caries experience had suffered dental pain by the age of six years and that this increased to 65-85% by the age of 12 years (Bastos et al., 2008). In addition, 11% had suffered pain in the preceding month (Bastos et al., 2008). Pau and colleagues (2007) investigated the prevalence of pain in 187 Greek children and found that 37% reported oral pain in the previous four weeks of which 60% attributed it to a tooth and 31% of those stated that it was specifically from a tooth with a hole in it. Furthermore, a study carried out in Sri Lanka aimed to ascertain the prevalence and impact of dental pain in 576 8-year-old schoolchildren (Ratnayake and Ekanayake, 2005). The questionnaire asked whether pain had been experienced, how often, what caused it and what treatment had been given for the pain? The children's reports revealed that

49% had suffered pain and the most common reason cited by the children for pain experience was “cavity/decayed tooth” as reported by 67% of children.

Finally, in the UK, a study of 589, 8-year-old schoolchildren found that 48% of children reported pain, 8% of whom had experienced pain in the previous four weeks (Shepherd et al., 1999). However, it should be noted that approximately one-third of the children attributed the pain they had experienced to a “wobbly tooth”, therefore it may be that the number of children affected by dental pain due to caries is somewhat lower than the figures suggested by this study. “Toothache” was reported by 18% of 12-year-olds and 15% of 15-year-olds in the latest Child Dental Health Survey (Tsakos et al., 2015). Higher proportions of children who were eligible for free school meals reported that they had had toothache than those who were not eligible for free school meals, indicating an increased oral pain prevalence in those with lower socioeconomic status (Tsakos et al., 2015).

As these studies directly asked children about their experience of dental pain, they may be a more accurate reflection of the prevalence compared to studies which used proxy report. Although these studies were performed in different settings and using a variety of methods, it would appear that at least half of all children with caries have experienced pain. Due to the recall periods and methods employed, this may be an underestimation of the true prevalence of pain in children with dental caries. Despite the differences in whether parent or self-report were used to ascertain the prevalence of pain, where demographic characteristics were investigated, a higher prevalence of pain was found in those with low socioeconomic status (Nuttall et al., 2006; Bastos et al., 2008; Lewis and Stout, 2010; Tsakos et al., 2015).

Impacts related to pain

A variety of other impacts have also been reported relating to caries-induced oral pain. An impact on eating has frequently been reported in population studies assessing the impact of dental caries. In a Thai national survey, 64% of 12- and 15-year-olds reported oral impacts related to eating. However, these may not all have been associated with dental caries, although toothache was the most frequently identified cause of impacts in 12-year-olds and the second most frequent in 15-year-olds (Krisdapong et al., 2009). The most frequent impact reported by Greek children was difficultly eating (40%). It was also the most common impact reported by Sri Lankan children (58%) and children in the UK (73%) (Shepherd et al., 1999; Ratnayake and Ekanayake, 2005; Pau et al., 2007).

Other reported impacts have included loss of sleep, time off school, problems cleaning teeth, speaking and interference with normal daily activities (Shepherd et al., 1999; Ratnayake and Ekanayake, 2005; Pau et al., 2007; Krisdapong et al., 2009). Indeed, Krisdapong and co-workers (2013b) found that 5% of 12-years-olds and 4% of 15-year-olds in Thailand had missed school due to toothache, with a mean of 1.4 days missed. Similar findings were reported in a study in the USA, where 3.9% had missed at least one day of school due to dental pain or infection (Jackson et al., 2011). Other studies have found between 11 and 70% of children missed school due to dental pain (Shepherd et al., 1999; Naidoo et al., 2001; Ratnayake and Ekanayake, 2005). In the USA, 17.3% of absences from school, which were related to oral health were due to dental pain or infection. In addition, absence from school due to pain and infection were associated with poorer school performance, whereas absence for dental treatment was not (Jackson et al., 2011). Therefore, it may be that the presence of ongoing dental pain has effects on children's school performance which is not simply related to school absence.

Caries-related impacts affect daily activities such as eating and sleeping and therefore may have effects on children's general health and development, which may further impact upon their school performance. The next section will discuss the effects of dental caries on general health.

Dental caries and general health

In more recent years, investigators have attempted to evaluate the association between dental caries and general health.

Weight and height

The regulation of growth in children is complex and multifactorial and may be affected by several aspects of dental caries (Sheiham, 2006). Firstly, it may result in decreased intake of food due to pain, as discussed in the previous section. This may result in periods of suboptimal nutrition, thereby restricting growth. Other children may have disturbed sleep which may interfere with glucosteroid production and chronic infection which may, in turn, interfere with metabolic pathways (Sheiham, 2006).

A recent systematic review, involving 17 studies, investigated the relationship between anthropometric measures (i.e. height and weight) and dental caries in children (Li et al., 2015). A meta-analysis was not possible due to the disparity in caries detection methods and the anthropometric measurements used. Two of the included observational studies found that higher caries experience was associated with negative effects on children's growth. However, one-third of the included studies found no significant relationship between anthropometric measurements and caries and more than one-quarter found conflicting data. The authors concluded that the evidence for a relationship between growth, weight and caries was inconclusive. They also recommended that, in order to be able to assess this relationship future research should ensure that caries data are collected in a standardised format by trained and calibrated examiners and that a full range of anthropometric measurements are recorded. In addition, longitudinal cohort studies were recommended so that associations could be accurately assessed throughout the period of children's growth (Li et al., 2015).

The relationship between obesity and dental caries was recently explored in a systematic review and meta-analysis (Hayden et al., 2013). Fourteen studies were included and there was found to be a small overall association between obesity and caries in the permanent dentition, but not in the primary dentition.

There was also a significant relationship between obesity and dental caries in industrialised countries but not in newly industrialised countries. Concurring with the systematic review by Li and colleagues (2015), the authors also found disparities between the methods of quantifying obesity and the way in which caries experience was measured, thus making it more difficult to compare the included studies.

Another systematic review of the literature examining the relationship between body mass index (BMI) and caries included 48 studies (Hooley et al., 2012). One study found both positive and negative correlations between dental caries and BMI. Twenty-one studies found no association between BMI and caries, 16 a positive relationship and nine an inverse relationship. One study found a U-shaped pattern or a difference between age cohorts. Some interesting observations emerged from these different studies:

- Studies which found a positive association between BMI and caries, used methods which allowed more accurate detection of dental caries, whereas those who found negative correlations used field techniques, which may have underestimated caries prevalence
- Positive associations were more likely to be found in highly developed countries
- Studies where a higher prevalence of caries were found were significantly more likely to find an inverse relationship between BMI and caries
- Underweight participants were underrepresented in those studies which found positive associations compared with those which found inverse or no relationship

The authors concluded that dental caries is associated with both high and low BMI and that the precise nature of the associations remains unclear but appears to be non-linear. They also recommended that future studies should examine longitudinal outcomes for children with early childhood caries into adolescence and early adulthood, to further assess these associations (Hooley et al., 2012).

It is apparent from these studies, that there is no predictable effect of caries on weight and height as other factors such as age and socioeconomic factors come into play. However, it is clear that obesity and caries also have common risk factors. There was an inverse relationship with BMI in samples with high levels of caries, which may be attributed to the impacts on eating as discussed in Section 2.2.2. This may also explain increases in BMI that have been noted following dental treatment for dental caries in underweight children (Acs et al., 1999; Monse et al., 2012). Future longitudinal studies may be able to assess why these contrasting relationships are found, enabling targeting of specific interventions according to risk.

Iron-deficiency anaemia

Iron deficiency is the most common nutritional deficiency worldwide and is associated with poor health and nutrition (World Health Organization, 2008). As such, it would not be surprising to find iron deficiency in children with dental caries who may have difficulty eating a varied diet compounded with ongoing chronic infection. The detection and correction of iron deficiency is particularly important in children as it can impair their physical and mental development (World Health Organization, 2007a).

A recent case-control study based in Canada recruited 122 children without caries and 122 with severe early childhood caries (S-ECC) with a mean age of 40 months (Schroth et al., 2013). S-ECC is defined by the American Association of Pediatric dentistry as follows:

> "Under 3 years of age, any sign of smooth-surface caries is indicative of severe early childhood caries (S-ECC). From ages 3 through 5, 1 or more cavitated, missing (due to caries), or filled smooth surfaces in primary maxillary anterior teeth or a decayed, missing, or filled score of ≥4 (age 3), ≥5 (age 4), or ≥6 (age 5) surfaces constitutes S-ECC."

Mean ferritin concentrations did not differ significantly between the groups, however, children with caries had significantly lower mean haemoglobin levels than the controls. Children with caries were also more likely to have lower ferritin levels than the controls. Children with caries accounted for 92% of those who were classified as iron deficient, having both low ferritin and haemoglobin levels. Overall, around one third of the children had iron deficiency anaemia, of which 92% were children with caries. Regression analysis showed that children with caries were more than six times more likely to have iron deficiency anaemia, than those without caries.

These findings concur with those of another Canadian study conducted by Clarke and co-workers (2006) who examined a group of children aged 2-6 years with S-ECC. Several tests were performed to ascertain whether these children were suffering from malnourishment. Low serum ferritin was found in 80% of the children, with iron deficiency identified in 11%. This prompted the authors to suggest that any child with S-ECC should have an assessment of iron levels, as although many of the children in this study had BMI, height, weight and body fat within normal limits, 80% were actually suffering from iron deficiency (Clarke et al., 2006).

In the UK, Rodd and Blankenstein (1995) examined pre-operative haemoglobin, mean cell volume and red cell porphyrin in 109 children aged 6 years or under who attended for dental extractions under general anaesthesia. It was found that there was no statistically significant difference between anaemic and non-anaemic children with regard to sex and social class. However, a significantly higher number of extractions had been prescribed for children with anaemia indicating that anaemic children had a higher caries experience overall (Rodd and Blankenstein, 1995).

There does, therefore appear to be an association between iron deficiency anaemia and dental caries, although this is likely to be similar to general growth and development, in that the aetiology is multifactorial. An awareness of the potential for children with caries to have iron deficiency anaemia is important to prompt appropriate haematological investigations and any deficiencies to be rectified as soon as possible to minimise ongoing developmental problems.

The next section will focus on the impact of caries on children's everyday lives by identifying the psychosocial aspects which may be involved. These aspects have been captured predominantly by the assessment of oral health-related quality of life (OHRQoL) using a variety of measures.

Oral health-related quality of life

OHRQoL concerns assessment of oral symptoms, functional limitations, social and emotional wellbeing associated with oral conditions and will be discussed in further detail in Section 4.6. It is generally determined by administering questionnaires to proxies or to children for self-report. Existing questionnaires are generic (i.e. they are designed to cover a range of oral health conditions) and some have been adapted from adult measures. This section will assess the psychosocial aspects reported by these measures and associations between dental caries and these impacts.

Psychosocial impacts

A number of different questionnaires have been used to assess psychosocial impacts. However, this section will focus only on those which are designed for self-report, to ensure that only impacts reported by children themselves are described. The most frequently used measures will be described in detail in Section 4.6.

Child Perceptions Questionnaire

The Child Perceptions Questionnaire (CPQ) has been widely used in clinical and non-clinical paediatric populations and children with dental caries have been found to have significantly higher total scores to those without caries (Foster Page et al., 2005; Foster Page et al., 2008; Koposova et al., 2012; Martins-Junior et al., 2012b; Paula et al., 2012a). Children with caries have also been found to have significantly different scores to those without caries in some or all of the domains (Do and Spencer, 2007; Goursand et al., 2008; Koposova et al., 2012; Martins-Junior et al., 2012b; Paula et al., 2012a). Positive correlations have been found between the overall score and number of carious teeth (Jokovic 2002, Jokovic 2004, Aguilar-Diaz, 2011, Foster-Page 2012). However, this finding has been disputed by other investigators who have reported no correlation or no statistically significant association with clinical caries data

(Marshman et al., 2005; Robinson et al., 2005; Brown and Al-Khayal, 2006; Gururatana et al., 2011; Kolawole et al., 2011).

Child Oral Impacts on Daily Performances

Studies which have used the Child Oral Impacts on Daily Performances index (C-OIDP) have found that children with dental caries frequently report functional impacts such as difficulty eating, drinking and pain when brushing teeth (Pau et al., 2008; Krisdapong et al., 2012c; Basavaraj et al., 2013; Krisdapong et al., 2013a; Naidoo et al., 2013). Furthermore, children who have reported toothache in the previous three months have been found to have significantly higher scores than those without (Yusof 2012, Castro 2011, Mbwalla 2011) and are more likely to report impacts than those who do not have dental caries (Krisdapong et al., 2013a). Children with caries also report impacts related to smiling, playing, difficulty sleeping or relaxing, emotion and schoolwork (Pau et al., 2008; Krisdapong et al., 2012c).

The 2013 Children's Dental Health Survey also used the C-OIDP to assess OHRQoL in 12- and 15-year-olds (Tsakos et al., 2015). Over half (58%) of 12-year-olds and 46% of 15-year-olds reported at least one impact from their oral health over the preceding three months, although it should be noted that many of these children did not have caries. Therefore it was not surprising that the most frequently chosen impact was "being embarrassed to smile or laugh" which is more likely to relate to their perceived malocclusion, than to caries. However, the next most frequently reported impact was "difficulty eating" which may have been associated with dental caries.

Child Oral Health Impact Profile (COHIP)

Where the Child Oral Health Impact Profile (COHIP) has been used to determine the impact of caries on OHRQoL, significant correlations have been found between dental caries and the overall score (Broder and Wilson-Genderson, 2007; Ahn et al., 2012; Asgari et al., 2013). In addition, significant associations have been found between caries and the oral health, functional wellbeing, social/emotional wellbeing and self-image domains (Broder and Wilson-Genderson, 2007; Ahn et al., 2012). However, relatively few studies using COHIP have been conducted to date.

Scale of Oral Health Outcomes (SOHO-5)

In common with the other OHRQoL measures, the impacts which were most frequently reported using SOHO-5 were associated with eating and sleeping (Tsakos et al., 2012; Abanto et al., 2014). Children also reported impacts related to smiling, playing and speaking (Tsakos et al., 2012; Abanto et al., 2014). Children with caries experience, current or previous toothache, pulpal involvement and with sepsis had significantly higher mean scores than those without these presentations (Tsakos et al., 2012; Abanto et al., 2013a; Abanto et al., 2014).

Studies using other measures

A study carried out in Brazil compared 4-year-old children with no caries with those with S-ECC to assess how they felt about their teeth. The children were asked how they felt about their teeth by showing them a drawing of a happy child and a sad child. When the children were asked if they felt happy or sad about their teeth, 34% (n=26) of the S-ECC group reported that they were sad about their teeth compared with 22% (n=49) in the caries-free group. An odds ratio found that the probability of children with caries feeling sad about their teeth was significantly higher than for those without caries (Feitosa et al., 2005). While this study attempted to involve young children, it is not clear whether the simple measure used by the investigators actually assessed OHRQoL or some other construct.

Similar findings were reported in a USA based study which aimed to investigate the relationship between children's oral health and their OHRQoL and to evaluate their smiles (Patel et al., 2007). Ninety-nine patients, aged 4-12 years, watched a cartoon while they were filmed and then completed a modified Michigan Oral Health Related Quality of Life Scale-Child version, with questions related to smiling. The children's smiles were analysed to assess the degree of mouth opening and the number of teeth shown during smiling. The investigators found a clear relationship between children's oral health status and their smiles as assessed by the children, their parents and an independent objective examiner. Children with caries experience rated their smiles less positively than those with no caries. The independent assessment revealed that children with no caries experience showed more teeth while smiling. However, one limitation of this study was the lack of detail regarding the extent of caries in the anterior teeth.

Association between reported impacts and clinical status

In summary, there have been conflicting findings regarding any associations with clinical status where CPQ has been used. Others have found weak correlations between clinical indicators and OHRQoL scores in adult populations (Cushing et al., 1986; Locker and Slade, 1994). It has been postulated that it may be difficult for OHRQoL measures to assess impacts in children with relatively low levels of disease or that there are other factors which may mediate the impacts such as socioeconomic status, cultural differences and general health (Locker, 1992; Locker and Slade, 1994; Marshman et al., 2005; Cunnion et al., 2010). Using DMFT as the measure of dental caries may be less sensitive than other methods such as ICDAS, as discussed in Section 2.2.1. However, it is likely that many of the impacts associated with dental caries would be associated with more extensive disease, such as that which is cavitated, which would be routinely collected as part of DMFT. However, it would be more informative if studies reported what proportion of DMFT was attributed to the "D" component, as this would give an indication of the level of active caries in the population. As existing measures of OHRQoL for children are generic, and were not designed to be used solely in populations with dental caries, it may be that they are unable to discriminate adequately in these specific populations. Therefore a disease-specific measure may be more sensitive to identifying impacts associated with dental caries and have stronger correlations with clinical data.

It is however clear that dental caries may have a variety of impacts on children's everyday lives and their general health and wellbeing. Therefore it is important to ensure that these psychosocial aspects are considered along with functional aspects and pain, when assessing the overall impact of dental caries. While there are a number of questionnaires available to assess OHRQoL, these have been developed in a variety of ways and used in different settings. Furthermore, it is not clear if they have been validated in all of the populations in which they have been used. A systematic review of these child self-report OHRQoL measures to evaluate the ways in which they were developed and their psychometric properties is thus indicated.

2.2.3 Impact on society

In addition to the impact on the individual, caries may also impact on society. This may have direct societal impacts such as the costs associated with providing treatment, as well as indirect costs stemming from children's absences from school which requires parents to take time off from work. As school absences were covered in 2.2.2, this section will focus on data related to cost.

Very little data exist regarding the cost of treating dental caries in children in the UK and, that which does exist, mainly pertains to treatment under GA. Some data are available for other countries (Australia, USA, the Netherlands) which are discussed in this section, although this is also limited. These relate mainly to the cost of emergency attendances and admission to hospital and therefore underestimate the total costs.

A recent study did however evaluate the cost-effectiveness of three different interventions over a 3-year period in the Netherlands (Vermaire et al., 2014). The mean cost to the health care service was €90 for the non-operative caries treatment and prevention arm (NOCTP), €120 for the increased professional fluoride application programme (IPFA) and €85 for regular dental care. The associated time and travel costs to the family were €227, €356 and €213 for the NOCTP, IPFA and regular care arms respectively. These children had low levels of caries and may therefore not reflect the costs associated with treating children with more extensive caries experience, but do offer some insights into the costs associated with attending for regular dental care.

Nalliah and colleagues (2010) reported figures from the USA based on a total of 120,033,750 attendances to emergency departments, 330,757 were attributed to dental caries. Children accounted for 24,982 of these visits, of which the majority (99.59%) were not admitted for further care. The approximate cost of this was $667.48 per child, which equated to $14.33 million (Nalliah et al., 2010).

A five-year service review was performed in Western Australia to ascertain the reason for paediatric hospital admissions for orofacial conditions and to estimate the costs associated with this treatment (Kruger et al., 2006). Dental caries accounted for 28.3% of all cases with costs of $10 million Australian Dollars, with the highest admission figures seen for pre-school children. Furthermore, conditions of the pulp and periapical tissues accounted for 7.1% of admissions with associated costs of $2.4 million Australian Dollars.

In the UK, the majority of children are seen for regular dental care by a general dental practitioner within primary care settings. The cost of this care is provided by the NHS for those under 18 years of age or 19 years of age when in fulltime education. Sheffield has a higher proportion of children (75.9%) accessing general dental care when compared with the national average (69.1%) (Public Health England, 2015). However, there are inequalities with a smaller proportion of children from areas with higher deprivation accessing general dental care than those from more affluent areas in South Yorkshire (Public Health England, 2015). Primary dental care is also provided at Charles Clifford Dental Hospital in Sheffield and its associated outreach clinics (where the majority of care is provided by dental students). These outreach clinics include general dental practices and salaried dental services clinics.

Access to specialist paediatric dentistry services varies around England with some areas having no dedicated paediatric dentistry specialists (British Society of Paediatric Dentistry, 2015). Children with extensive dental caries, dental anxiety and those who require management of complex dental conditions may be referred from primary care to specialist services for treatment. Children with caries account for a large proportion of these referrals and may subsequently require treatment under general anaesthetic. Outpatient paediatric dentistry services in Sheffield cost £ 937,398 in 2013/14, however this figure does not cover the cost of treatment provided under GA (Public Health England, 2015).

In the UK, approximately 45,000 children aged up to 16 years were admitted to hospital in 2013-2014 with a diagnosis of dental caries, making caries the most common reason for children to have an admission to hospital (Hospital Episode Statistics, 2015). The estimated costs of treatment for children under 18 years of age requiring extractions as hospital admissions in 2013-2014 was cited as approximately £37 million (Department of Health, 2014). In Sheffield, 1495 children underwent hospital admission in 2013/14, the majority of which is likely to be related to the treatment of dental caries (Public Health England, 2015).

The number of hospital admissions for emergency and elective procedures for the treatment of dental caries is costly and figures quoted relate only to the cost of the treatment provided. They do not take into account the financial cost for parents of attending these visits, such as travel costs and time taken off work. Nor do they take account of appointments in primary care or outpatient attendances which are inevitable before the child is finally admitted to hospital. The number of children referred to specialist services for the management of caries appears to be increasing in some areas (British Society of Paediatric Dentistry, 2015). Due to cuts in NHS budgets, there is insufficient capacity to manage the burden of disease, especially in young children. Therefore there is a gross underestimation of the true cost related to the management of dental caries in children. It is clear, that caries does have an impact on society given the large numbers of children who are being admitted to hospital for treatment of dental caries. There are significant costs for the health service and to the economy where parents may be absent from the workplace on multiple occasions.

2.2.4 Prevention and treatment

The final criterion proposed by Sheiham and Watt (2003) relates to whether the disease is preventable and whether effective treatments are available. This section will briefly discuss the prevention of dental caries and clinical effectiveness of the various approaches which can be used to treat the disease in children.

Prevention of dental caries

Prevention of dental caries is achieved by optomising exposure to fluoride, reducing frequency of sugar exposure by improving diet and protecting vulnerable tooth surfaces. This can be achieved by encounters with individuals or as part of community or population based public health measures. Both "Delivering Better Oral Health" and the Scottish Intercollegiate Guideline Network (SIGN) give evidence-based prevention strategies which can be employed and are supported by UK commissioning guidelines (National Institute for Health and Care Excellence, 2014; Public Health England, 2014b; Public Health England, 2014d; Scottish Intercollegiate Guidelines Network, 2014).

Optimising fluoride

Fluoride toothpaste

Ensuring children are using the appropriate strength of fluoride toothpaste twice daily can reduce the incidence of dental caries. A Cochrane systematic review, based on 75 studies, investigated the effects of different strengths of toothpaste in the prevention of dental caries in children and adolescents (Walsh et al., 2010). This review found that the preventive effect of fluoride toothpaste increased with increasing fluoride concentration. Although not statistically significant, it was found that the preventive effect increased with higher caries experience at baseline and with supervised brushing. Overall, the benefit of using fluoride toothpaste in preventing caries was seen when compared to a placebo, however this was only significant when the fluoride concentration was greater than 1000 parts per million (ppm).

An earlier systematic review also demonstrated the caries-protective effect of toothbrushing with a fluoride toothpaste increases with higher frequency of use (Marinho et al., 2003).

Other behaviours which may influence the effect of fluoride include brushing last thing at night as this allows fluoride concentrations to remain high during the night. Not rinsing following toothbrushing has also been found to be associated with decreased caries increment compared to those who do rinse following brushing (Chesters et al., 1992). These findings form the basis of the advice provided by SIGN and the Public Health England, who recommend that children should brush last thing at night and at one other time. Those under three years of age should use a toothpaste containing at least 1000 ppm and older children using a toothpaste containing 1500ppm unless a caries-risk assessment indicates a higher concentration should be prescribed (Public Health England, 2014b; Scottish Intercollegiate Guidelines Network, 2014).

Access to fluoride toothpaste can be improved by the implementation of nursery and school supervised toothbrushing schemes, such as those provided in Scotland and Wales (Welsh Assembly Government, 2008; Macpherson et al., 2010). In Scotland, where this programme has been in place since 2006, reductions in the prevalence of caries and dental inequalities have been seen which have been attributed to the national roll-out of nursery toothbrushing programmes (Macpherson et al., 2013).

Fluoride varnish

There are other vehicles which can also be used to deliver fluoride. Topical fluoride varnish, at a concentration of 22,600ppm, has been shown to reduce the incidence of caries (decayed missing and filled surfaces), with a pooled prevented fraction of 37% in the primary dentition and prevented fraction of 43% in the permanent dentition as reported in the most recent systematic review (Marinho et al., 2013). This review showed that four applications per year did not significantly decrease caries incidence in children when compared to application twice yearly, therefore it is recommended that fluoride varnish should be applied at least twice yearly (Public Health England, 2014b; Scottish Intercollegiate Guidelines Network, 2014). The advantages of fluoride varnish are that it is generally well accepted and can be applied by all members of the dental team where adequate training has been provided.

Water fluoridation

Another vehicle for delivering fluoride is by fluoridation of the water supply. Water fluoridation has been successfully used in many countries; however, strong opposition remains in the UK despite the absence of evidence to support any detriment to general health. A systematic review of the available evidence was conducted and demonstrated that water fluoridation did result in a decrease in caries incidence and that a median of six people would need to drink fluoridated water to render one individual caries-free (McDonagh et al., 2000). This report did not find any evidence to support claims of increased hip fractures or cancer, but there was insufficient evidence to investigate other potential health problems. The available evidence did show a dose-response effect with dental fluorosis, estimating that water fluoridation at a level of 1ppm would result in a prevalence of fluorosis of 48%, with 12.8% being of aesthetic concern. The report concluded that there was a paucity of high quality research in this area, which is surprising given the interest in water fluoridation as a public health measure. Furthermore the Medical Research Council, UK, recommended that further research was required to ascertain the effect of water fluoridation on outcomes such as dental pain and number of extractions required, rather than just relative differences in caries experience between children living in areas with fluoridated and non-fluoridated water supplies (Medical Research Council, 2006).

Diet

Lowering the frequency and quantity of sugar intake can decrease the incidence of caries in children. A recent systematic review, performed to inform the World Health Organization's (WHO) guidelines, included 50 studies which had assessed the relationship between sugar intake and dental caries in children (Moynihan and Kelly, 2014). Of the fifty studies, 42 studies reported at least one positive association between caries and sugar intake. Pooling of data from cohort studies produced a risk ratio of 7.15 when comparing caries prevalence in high and low sugar intake groups. In addition, analysis of included studies of moderate quality demonstrated a reduction in caries incidence when sugar consumption was less than 10% of the total energy intake. The authors concluded that there was consistent evidence to support a relationship between sugar intake and the development of dental caries and that these effects were lifelong. Dietary advice should promote both general and oral health and advocate the reduction in the frequency of intake of sugar and avoidance of sugary drinks at bedtime and during the night. The provision of a diet diary can help to give tailored advice to parents and children regarding dietary modifications which can be made to decrease the cariogenic potential of their diet (Public Health England, 2014b). However, it should be noted that a systematic review investigating the effect of dietary advice demonstrated that it is difficult to change dietary habits without the addition of behaviour changing techniques (Greaves et al., 2011). To date, there have been no studies which have identified the most appropriate behavior changing techniques for use in oral health (Asimakopoulou and Newton, 2015). A complicating factor in changing dietary habits in children, is that they are not responsible for providing their own food and therefore behaviour change has to be tailored to the whole family. The WHO has recently released guidance on reducing free sugar consumption to below 10% of daily energy intake and this will impact on future government policy and perhaps help families to reduce their sugar intake (World Health Organization, 2015a).

Fissure sealants

Fissure sealants (sealants) have been used to protect the occlusal surfaces of molar teeth for over 50 years and can be resin or glass ionomer based. A systematic review of the efficacy of sealants in permanent teeth found that children with resin-based sealants were less likely to have caries in their permanent molars than those with no sealants (Ahovuo-Saloranta et al., 2013). Meta-analysis of nine studies included in this review was highly significant for prevention of caries in first molar teeth at 12, 24, 36 and 48-54 months post-sealant placement. Despite the established benefit of sealants, only 7% of 8-year-old children in England were found to have sealants present in the most recent Child Dental Health survey. This was much lower than the proportion of children with sealants in both Wales and Northern Ireland (24% and 34% respectively) where public health initiatives encourage their placement (Pitts et al., 2015).

Upstream actions

In order to create sustainable improvements in oral health, there should be a focus on addressing the determinants of oral health (Watt, 2007). This requires oral health strategies to be implemented alongside other health interventions using a common risk factor approach. Healthy public policies, legislation and fiscal measures can be used to promote health (Watt, 2007). An example of this is Tobaccco Control where policies to regulate the tobacco industry and introduce legislation to limit the environments where people can smoke have been used in conjunction with information targeted at individuals who smoke (World Health Organization, 2014). A similar model could be adopted to encourage better oral health. This could include tighter regulations on the food industry including a need for clear nutritional labeling or television advertising during children's programming and ensuring that nurseries, schools and hospitals are supported to provide food which complies with nutritional standards (Watt, 2007). A number of these initiatives have been implemented in the UK, such a nutritional standards for schools and children being allowed drinks with no-added sugar during the school day (Long, 2015). However, there have been difficulties with regulation of food labeling in the UK with a number of different types being used making it difficult for consumers to know which products to choose (Malam et al., 2009).

In summary, there are several effective modalities available to help prevent caries and evidence that dietary modification is of benefit. However, children are reliant on their families to provide appropriate fluoride toothpaste, to encourage twice daily toothbrushing, to modify their diets and to access dental services. Population-based schemes may therefore help to reduce the influence of social inequalities on the prevalence of caries when combined with upstream actions.

Treatment of dental caries

The evidence-base for the restoration of carious primary teeth is sparse and conflicting. Some retrospective studies in the UK have reported no benefit in restoring primary teeth as a high proportion of the restorations placed are lost and many carious teeth appear to exfoliate naturally without adverse effects (Tickle et al., 1999; Tickle et al., 2002; Levine et al., 2003). However, the majority of restorations placed in these studies were with glass ionomer cement, which has been shown to have decreased longevity when compared with other materials (Chadwick and Evans, 2007). More recently, a cohort study involving 2654 children found that restored carious primary teeth were less likely to be subsequently extracted than those that remained unrestored (Stephenson et al., 2010). Despite this finding, dental caries in the primary dentition remains largely untreated in the UK (Public Health England, 2012). The lack of an evidence-base for the most effective treatment for young children with caries may be a contributing factor to this non-restorative approach.

Preformed metal crowns (PMCs) have, however, been shown to outperform conventional intracoronal restorations in the primary dentition and are recommended by specialist society guidelines (Randall, 2002; American Association of Pediatric Dentistry, 2008; Kindelan et al., 2008). However, barriers such as lack of training and a desire not to administer local anaesthetic have previously been identified which may account for low provision of these in primary care (Threlfall et al., 2005; Chadwick et al., 2007). More recently, a randomised controlled trial investigating the success of a non-invasive method of placing PMCs (the Hall technique) has shown better clinical outcomes for PMCs than conventional restorations in primary care (Innes et al., 2007; Innes et al., 2011). Children's perspectives were sought regarding which treatment they preferred, but OHRQoL was not measured (Innes et al., 2007). Any reduction in OHRQoL impacts following treatment with the Hall technique has not yet been assessed. However, children generally expressed a preference for the Hall technique crown over the conventional restoration. Other studies have found similar clinical outcomes when comparing conventionally prepared PMCs and those placed using the Hall technique (Ludwig et al., 2014). As the procedure is non-invasive, it is hoped that there will be wider provision of this treatment in primary care in the future.

Further evidence may emerge as to the most clinically effective treatment in the primary dentition as a multi-centre randomised trial (FiCTION) is currently in progress. This study is investigating the effectiveness of different treatment options (prevention alone, biological approaches and conventional restorations) in young children (aged 4-7 years) in a primary care setting (Innes et al., 2013). As such, the findings of this trial are likely to have national and international significance in future planning of dental care for young children. The primary outcome measure is pain, with OHRQoL as a secondary outcome measure. Unfortunately, OHRQoL will be assessed using a proxy measure, as there is currently no existing measure suitable for this age-group, therefore the child's perspective will not be fully explored (Innes et al., 2013).

Management of dental caries in children may sometimes require treatment under GA, especially in young children where dental caries affects several quadrants. Although treatment of the disease may have positive outcomes it is associated with post-operative morbidity and anxiety (Bridgman et al., 1999). Studies have tended to focus on functional outcomes following surgery, with little examination of the patient experience (Low et al., 1999; Hosey et al., 2006). Indeed, a recent systematic review found only 11 studies which had compared children's OHRQoL following oral rehabilitation under GA. Furthermore, studies had used a variety of different measures, none of which had been completed by the child (Jankauskiene and Narbutaite, 2010). Similar findings were reported in another systematic review, where only nine studies were included which had used validated measures of OHRQoL (Antunes et al., 2013). Of those which had investigated oral rehabilitation under GA, none included child-self report. There is therefore, a lack of evidence regarding the effectiveness of comprehensive dental care (i.e. restorations and extractions) or extractions only under GA from the child's perspective. A tool to evaluate these outcomes would be particularly useful, as comprehensive care GA services are being withdrawn in some areas and evidence demonstrating the positive outcomes of this type of treatment may help prevent future service loss.

Dental caries can be treated in a number of ways, however, to date studies have focused on clinical outcomes and there has been little exploration of patient-reported outcomes of these interventions. Inclusion of children's perspectives of treatment would greatly aid the evaluation of the effectiveness of different interventions. Future research is required to ascertain the longitudinal effects of dental treatment on reducing the impacts of dental caries from the child's perspective, as well as the clinical and cost-effectiveness of different treatment modalities.

2.3 Summary

Comparing the available evidence against the criteria proposed by Sheiham and Watt (2003), dental caries appears to meet all four criteria. There is a high prevalence of caries both in the UK and globally. There are associations with deprivation and caries can have severe physiological consequences such as the development of sepsis. It is also clear that dental caries can have a number of other impacts on the individual. These include pain, functional limitations such as impacts on eating, sleeping and speaking as well as effects on children's emotional status, social wellbeing, general development and school performance. Dental caries is preventable and treatable, although there is little evidence regarding the relative effectiveness of different treatment modalities clinically or from the child's perspective. Therefore, it can be concluded that caries has an impact on society and on the individual and can be considered a disease of public health significance.

Further exploration is required of the psychosocial impact of the disease and its treatment from the child's perspective. A number of different generic OHRQoL measures have been used to assess the impacts of caries, however, these may not be sensitive enough to measure impacts related solely to caries. A caries-specific measure suitable for young children would be of benefit in any future trials investigating the prevention or treatment of dental caries, to ensure changes in the impact of caries on children's daily lives are fully assessed.

Chapter Three

Aims and Objectives

3.1 Rationale

The previous chapter described the prevalence and reported impact of dental caries and its public health significance. The narrative review provided evidence of dental caries as a disease of public health significance. Despite this, there is a surprising paucity of research investigating the impacts of the disease from the child's perspective. In addition to impacts reported which may influence general health and development, there have also been a number of impacts described which affect children's everyday lives. These aspects have been assessed using a number of different generic OHRQoL measures, which may or may not have been appropriately validated in the populations in which they have been used. Although one review of OHRQoL measures briefly assessed the content and validation of one of these measures (CPQ for 11-14-year-olds), there has been no robust review of the development process and psychometric properties of these measures against existing criteria (Locker and Allen, 2007).

All of the existing measures of OHRQoL for children are generic and therefore may not be sensitive enough to the impacts associated specifically with dental caries. Development of a measure specifically for this disease is warranted given the prevalence of the disease in children and the lack of evidence to support the most effective treatment from the child's perspective. Indeed, a large multicentre trial (FiCTION) investigating different treatments for dental caries in young children, does not include a patient-reported outcome measure (instead relying on proxy report) due to the lack of a suitable measure for this age group (Innes et al., 2013). Understanding effectiveness from the child's viewpoint will ultimately allow us to tailor our oral health promotion programmes and treatment plans to provide both the most clinically effective and child-centred interventions.

However, in order to develop such a measure, we need to understand what impacts are important to children with caries. No studies were identified in the narrative review which have investigated the language and impacts children themselves describe. Rather, the impacts that have been reported are children's responses to lists produced by adults using adult terminology. It is conceivable that despite the experience of adult "experts", there are impacts which are important to children which have not been described due to this lack of qualitative enquiry. In addition, children may use different terminology than adults and therefore may respond negatively to a question if they do not know or understand the words used. It is therefore important that we ascertain what impacts are important to children with caries, so that we can truly understand the impact of the disease and determine whether treatment actually reduces these impacts.

3.2 Aim and objectives

The aim of this research is to develop a child-centred, caries-specific measure of OHRQoL.

The specific objectives to fulfill this aim are to:

1. Perform a review of the most commonly used self-report measures of OHRQoL for children against existing criteria, to assess strengths and weaknesses which will aid the development of the new measure;
2. Explore, through qualitative methods, the impacts of dental caries on children;
3. Involve children in the design and content of the measure to ensure that it is meaningful and relevant to them;
4. Test the measure for validity, reliability and responsiveness using modern psychometric techniques and existing criteria.

To fulfill the aim and satisfy the objectives, the research will consist of three interlinked studies. The first study will present a systematic review of the most commonly used self-report measures of OHRQoL for children against existing criteria and will satisfy Objective One. The second study will use qualitative methods to investigate the impacts of caries from the child's perspective. These impacts will form the basis of the items for the measure. Further child-centred methodologies will be used to select the final items for inclusion in the measure, to choose the design of the measure and to test face and content validity. This study will be conducted to meet Objectives Two and Three. The final phase of the research will test the measure longitudinally, in a group of children with active dental caries before and after treatment, using modern psychometric techniques. This study will fulfill Objective Four.

There are a number of methodological considerations which should be taken into account when performing this type of research. These will be described in detail in Chapter Four.

Chapter Four

Methodological considerations

Involving children in research requires consideration of a number of specific methodological, sociological and ethical issues. In addition, there are several methodological matters to appraise when developing measures to assess OHRQoL.

4.1 Chapter outline

This chapter will discuss:

1. The sociology of childhood, involvement of children in research and the ethical issues which may arise in this type of research;
2. Children's descriptions of pain and illness and how these may differ from those of adults;
3. The concept of health-related quality of life and guidance on how to construct measures of health-related quality of life;
4. The development and validation of the most commonly used self-report measures of OHRQoL for children.

4.2 Position of children in society

Over the last three decades, there has been a shift in conceptualisation of childhood. Social researchers have changed the focus from seeing children as immature and incomplete, to an appreciation that children are not incomplete adults but competent social actors who are actively involved in shaping their own social worlds (James et al., 1998). Furthermore, there has been a growing appreciation of children's rights, especially with regard to decision-making, following the publication of the United Nations Conventions on the Rights of the Child 1989; Children Act 1989; the Children and Young People (Scotland) Act 2014, and the Children (Northern Ireland) Order 1995 (Parliament, 1989; United Nations, 1989; Parliament, 1995; Scottish Parliament, 2014).

Recent UK health and social care policies have reflected these changes and advise inclusion of children's views in research, audit and service development. The publication of the National Service Framework for Children in 2004, aimed to achieve a change in the way health services were developed and promoted a move away from a disease-centred to a child-centred approach, creating a healthcare system designed around children's needs (Department of Health, 2004). Quality criteria were suggested for children's services and the need for shared decision-making with children and families was advocated. Following this, the Department of Health further promoted the involvement of children in all aspects of their medical care in the publication of "Achieving Equity and Excellence for Children" and "You're Welcome" (Department of Health, 2007; Department of Health, 2010). These papers encourage healthcare providers to listen to children and young people's opinions regarding their treatment. Recommendations are made for young-people friendly services including the development of age-appropriate information materials, to enable children and young people to make informed choices regarding their care. The NHS Confederation reiterated these points in their publication, stating that such engagement with children should be embedded within the culture and be part of the day-to-day activity of health professionals (NHS Confederation, 2012).

4.2.1 Constructions of childhood

In trying to understand childhood, researchers have postulated both psychological and sociological theories.

Psychological theories include five main approaches: physiological; psychodynamic; behavioural; humanistic, and cognitive. The most important of these is cognitive theory, which proposes a four-stage model of child development (Piaget and Inhelder, 1969). This theory is useful when considering which research methods are best to use with children of different ages. Although it should be borne in mind that children develop at different rates and therefore children's ability to understand will be related to their individual experiences (Eiser and Kopel, 2013). The impact of these developmental stages on how children may answer questionnaire items will be discussed in Section 6.1.3.

However, it has been argued that psychological approaches view childhood as a period of development and fail to recognise children as active social agents (Prout and James, 1998). A model was proposed, which involved four major ideal types, to differentiate how sociologists perceived childhood compared to psychologists (James et al., 1998). These were:

- The socially constructed child. This acknowledges that childhoods will be experienced differently depending on the individual child's life experience (e.g. gender, wealth, ethnicity etc.).
- The tribal child. This category describes children as occupying an autonomous world which accentuates the differences between adults' and children's views.
- The minority group child. This category sees the child as being a "minority group" drawing on the power imbalance which exists between children and adults.
- The social structural child. In this category children are viewed as forming a social group which has its own rights and desires.

These categories are not exclusive of each other. Indeed as children have relationships with other children and adults, but also live in a world which has adult constraints placed upon it, all four ideals may co-exist (Morrow, 2008; McLaughlin, 2015).

Some theorists have taken the perspective that as adults were once children, they have an advantage as they can appreciate both views. However, this belittles children's agency and overlooks the fact that as adults, our experiences as children are likely to be quite different to those of today's children (McLaughlin, 2015). By including children's views in our research, we have an opportunity to learn more about how children perceive their experiences, enabling us to better understand the impact of their diseases and the interventions we provide.

4.2.2 Involvement

There are many different ways that children's views can be incorporated into research. Children's involvement in research and social policy has been portrayed as a ladder with eight rungs (Hart, 1992). Three rungs are identified as non-participation (manipulation, decoration, tokenism) and the five which are related to participation are displayed in a hierarchical order (assigned but informed, consulted and informed, adult initiated but shared decisions with children, child initiated and directed and child initiated but they share decision-making with adults as equal partners). Some have criticised this model, due to its hierarchical nature, as this ignores the fact that different levels of participation may be more appropriate depending on the situation and the children involved (McLaughlin, 2015). A description of the terms is shown in Table 4.1. More recently, this ladder has been converted to a wheel by the Children's Commissioner (England), which contains similar descriptions, but removes the hierarchical nature and accepts that participation may be different depending on the circumstances (Figure 4.1) (Children's Commissioner, 2013).

Some research projects will lend themselves to different levels of involvement and the degree of involvement will also depend on the target population. Researchers should try to maximise children's involvement where appropriate. The benefits of involving children in research can be seen within four main areas:

- Benefits to the research itself;
- To its dissemination and evaluation;
- To the young participants;
- For adult researchers (McLaughlin, 2006).

Despite the acknowledged benefits, a systematic review of the pre-2005 oral health literature revealed that only 0.3% of studies included children as active research participants (i.e., children's perspectives were involved throughout the research process or their views were sought) (Marshman et al., 2007). More recently, however, several good examples have been published in the medical and dental literature of projects which have actively involved children in this way (Morris et al., 2007; Freeman et al., 2010; Stinson et al., 2012).

Table 4.1. Quality of children's participation. Modified from Hart (1992).

Quality of involvement	Level of involvement	Definition
Manipulation	Non-participation	Children are used by adults for their own gain. For example; young children carrying political placards with no idea of why they are doing so.
Decoration	Non-participation	Children take some part in a project but have no role in decision making.
Tokenism	Non-participation	Children are asked to be involved but have no say in how or what they communicate.
Assigned but informed	Participation	Adults conceive the project and children volunteer to take part. Children understand why they are participating and their voice is taken seriously.
Consulted and informed	Participation	Adults design and run the project. Children's views are taken into account in decision-making and feedback is provided regarding how decisions have been made.
Adult initiated but shared decision-making with children	Participation	Adults conceive the project but children are involved at all stages of decision-making, planning and initiation.
Child initiated and directed	Participation	Children decide on the project and how to they wish it to progress. Adults may be available to help but do not take charge of the project.
Children and adults share decision-making	Participation	Adults are invited by children to be involved in the decision-making. The idea is conceived by children.

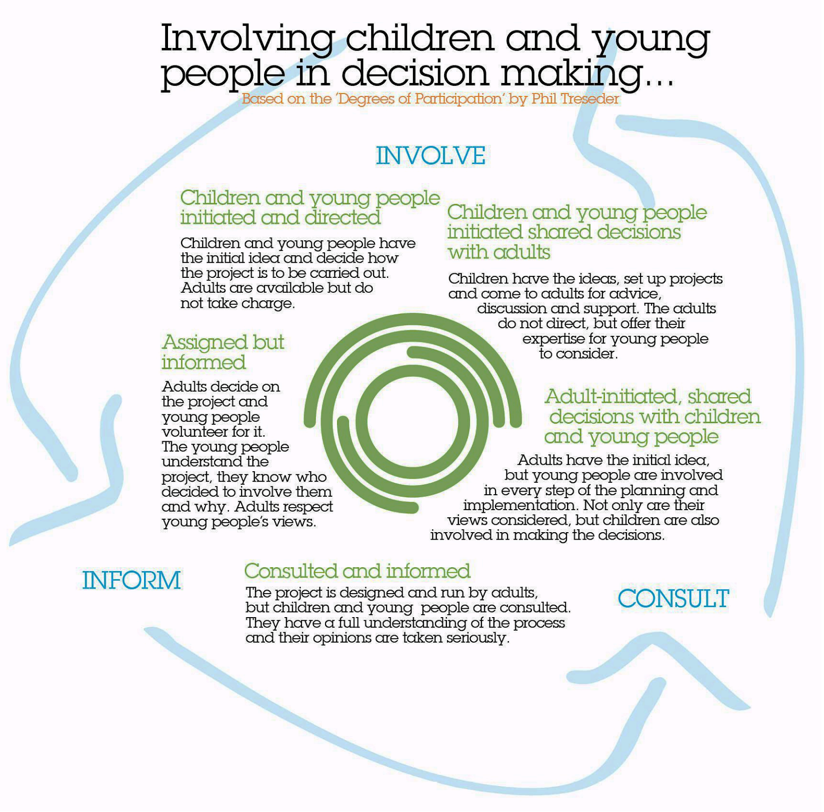

Figure 4.1. Involving children and young people in decision-making (Children's Commissioner, 2013).

Involving children in the development of a research project can aid prioritisation of topic areas, ensure age-appropriate materials are developed and may help with recruitment strategies (McLaughlin, 2006). Dissemination of the research can also be enhanced by the involvement of young co-researchers, ensuring future publications (e.g. patient information leaflets, research reports for participants, etc.) are in a format most likely to be accessed by other young people and that they are user-friendly. Being actively involved in research can help development of reasoning, debating and decision-making skills for the young people involved (Sinclair and Franklin, 2000). In addition, there are benefits to the adult researchers, such as gaining a greater understanding of children and young people's perspectives, learning new skills to facilitate communication and gaining from the enthusiasm young people bring (McLaughlin, 2006).

Children are capable of articulating their views and experiences and should be involved in research wherever possible. However, there are several ethical issues which need to be considered when children participate in research and these will be discussed in the next section.

4.3 Ethical issues

Ethical issues to be considered when carrying out research with children may include: power relationships, consent issues, confidentiality and dissemination of results. These considerations are not unique to research with children, although some aspects require specific attention when involving this population.

4.3.1 Power relationships

In general, our society is adult-centred, and thus there is potential for the power that adults have over children in everyday life to be carried over into research practice (Morrow and Richards, 1996; Harden et al., 2000; Punch, 2002). In particular, there are concerns that children may find it difficult to withdraw from a project once it has started. They may also be less able to explain if they feel uncomfortable with a particular question or may feel pressured to give the answer they think the researcher wants rather than offering their own opinion (Kirk, 2007). These issues exist within research with any population and it is the responsibility of the research team that strategies are in place to ensure that participants are comfortable and are reassured that there are no right or wrong answers. Various approaches have been adopted with children, for example, giving them a yellow card to hold up if they feel uncomfortable with a particular question and a red card if they wish to stop completely (Helseth and Slettebo, 2004). It should also be made clear to participants that they can withdraw at any time and that no one will be cross with them and that they don't have to explain why (Westcott and Davies, 1996). Additionally investigators should be aware of body language which may indicate that a participant is unhappy with the research process (Kirk, 2007).

4.3.2 Consent and assent

Gaining informed consent from participants is mandatory in research. However, this may not be possible with younger children, who may lack the capacity to give this level of consent. Therefore, consent is usually obtained from the adult "gatekeeper" and agreement to participate is gained from the child. This is termed "assent" and is defined as "an expression by the child of their desire to participate in the research" (Helseth and Slettebo, 2004). The use of assent allows children to participate who would otherwise be excluded from research, as they are unable to provide informed consent (Cocks, 2006). Age-appropriate materials must be developed to enable the child to make an informed decision; for example pictures and speech bubbles can be used (Marshman et al., 2012). It should be noted that this process is ongoing and the child's willingness to participate should be checked at regular intervals. The Royal College of Paediatrics and Child Health has produced written guidelines which cover the aspects mentioned above, but also include recommendations for ensuring valid consent, such as: checking families know who to contact if they have questions, that refusal to participate will not prejudice the child's future treatment and that it is stated whether the child will directly benefit from the research (Royal College of Paediatrics and Child Health: Ethics Advisory Committee, 2000).

4.3.3 Confidentiality

Confidentiality can be a complex area when conducting research with children, due to the potential for disclosure of information which may indicate that they or another child are "at risk". For example, where a child reveals information which leads the researcher to suspect there is a child protection issue, confidentiality will have to be broken in order to follow safeguarding protocols. The best interests of the child should always be paramount. It is proposed that the limitations of confidentiality should be discussed with the participant at the outset to ensure that they understand what type of information may be passed on and what will remain private (Alderson, 1995; Beresford, 1997; Davis, 1998).

4.3.4 Dissemination of results

Wherever possible, an age-appropriate summary of the key findings should be provided to participants (Alderson, 1995). Allowing children of a similar age to read the key-findings may help to ensure that the language is appropriate. Consideration should be given to whether participants should contribute to data interpretation or provide additional information (Westcott and Davies, 1996; Glasgow Centre for the Child and Society, undated). This may prove difficult, as it may mean contacting the participants and asking them to commit more of their time to the project. It may however, be feasible for a proportion of the participants to do so depending on the research setting. However, it should be appreciated that the children will have matured and their circumstances may have changed since the initial investigation and therefore what was important at the time they participated, may not be important by the time the research is complete (van Blerk and Ansell, 2007). In addition, opportunities for involving children in dissemination of results are often limited to verbal presentations, which are generally controlled by adults and therefore children have a lack of control over the process. Some have suggested that it would be better to use participatory techniques such as drama or games to allow children to be actively involved in dissemination (van Blerk and Ansell, 2007). It is important to ensure that all published data are anonymised to maintain confidentiality.

Strategies should be in place prior to undertaking any research project involving children to ensure that children who desire to take part are able to and that their safety and wellbeing is protected at all times.

The next section will consider studies where children have been active participants and how they have described their illness and pain experiences.

4.4 Children's description of pain and illness

Children may describe pain and illness in different terms to adults, due to differing vocabulary and concepts of illness. This section will report on the ways that children describe pain and illness. As pain is the most frequently reported impact of dental caries, there is a focus on the words that children use to describe pain.

4.4.1 Children's description of illness

Several qualitative studies have investigated children's experiences of various illnesses. Children in these studies have been able to discuss their knowledge and understanding of their condition, the limitations their illness places on their lives, their emotions and their role in self-care (Sartain et al., 2000; Rudestam et al., 2004; Woodgate, 2005; Bernays et al., 2015; Koller et al., 2015).

Canadian children aged 5-18 years with diabetes were found to be able to discuss their involvement in self-care and report what they thought their role was and at what age they should take more responsibility (Koller et al., 2015). The children were also able to discuss the aetiology of diabetes, although younger children were less able to provide detailed information, they understood that their pancreas was part of their body and that it wasn't working properly. The children in this study also discussed how they felt about having diabetes and the restrictions it placed on their lives, a well as the future consequences of the disease (Koller et al., 2015).

Other studies have focused on children's experiences of being in hospital (Sartain et al., 2000). In a UK study, Children aged 8-14 years discussed what they missed about home while they were in hospital, such as missing playing with friends and falling behind in school. Others discussed how they tried to make the hospital environment more like home by bringing their belongings with them (Sartain et al., 2000).

In a more recent UK study, children aged 7-11 years have also discussed their information needs prior to undergoing surgery and their involvement in the decision-making process (Smith and Callery, 2005). Children discussed their lack of knowledge and had used a variety of methods to obtain further information, including reading information leaflets designed for their parents. Although some of the information they had obtained from watching television actually caused them further anxiety as they involved procedures which did not go to plan. Children had clear ideas of how information regarding their procedure should be presented and the value of pre-admission visits.

These studies demonstrate that children are capable of discussing various aspects of their health including the consequences of their conditions. While younger (less than eight years of age in general) children may use less sophisticated descriptions, they understand and can articulate their experiences. They also are able to evaluate the information that they require prior to undergoing medical interventions and are able to suggest methods which would make this information more acceptable to them (Smith and Callery, 2005).

Children are capable of describing their illness experiences and their information needs. Therefore where possible, they should be involved in the research process from the beginning to ensure their information needs are met (for example, checking patient information leaflets are appropriate) and to ensure their perspectives are included.

4.4.2 Children's description of pain

Pain is described as an unpleasant sensory and emotional response to actual or potential tissue damage (IASP Task Force on Taxonomy, 1994). As such, it is subjective and therefore measures which do not include self-report are likely to yield inaccurate results. Little research exists examining the differences between children's reporting of somatic pain (i.e. that arising from muscles, bone, skin etc.) compared to that of adults. The studies that have been conducted report that the language used is different and that it varies with age and for children who have experienced regular hospitalisation (Savedra et al., 1982; Toole et al., 2000; Harman et al., 2005).

A number of studies have attempted to discover how children describe pain and findings have informed questionnaires and word lists to help children better express their experiences. Savedra and co-workers (1982) conducted a study in children aged 9-12 years of age from a sample of 100 North American children who were in hospital and 114 non-hospitalised school children. The children were presented with a list of 24 words and asked to circle the ones which could be used to describe pain. The words were assigned to three categories correlating to those of Melzack and Torgerson (1971) when formulating the McGill Pain Questionnaire (MPQ):

1. Sensory; i.e. temporal, spatial, pressure, thermal, etc.
2. Affective; i.e. tension, fear, autonomic response
3. Evaluative; i.e. overall pain intensity

The children who were in hospital chose evaluative and affective words more often than the other children. They also used the words "sickening", "like a pinch", "uncomfortable", "horrible" and "tiring" more frequently than the children who were not in hospital. In all categories there were a number of children who selected no words at all; suggesting that perhaps these were words or phrases with which the children were unfamiliar. It is unclear where the words included in the list were drawn from and whether they had been chosen by children or adults. However, this study demonstrated that most children were able to choose some words to describe pain that they had suffered and that these words were different depending on their experiences.

Another North American group used semi-structured interviews to elicit information regarding children's description of pain (Ross and Ross, 1984). Children (n=944) aged 5-12 years took part either in school, clinic or hospital settings. The children were able to give descriptions of pain in 70% of cases and used words such as "stabbing", "burning", "dull" and "pressing". Many children (n=286, 30%) used descriptive sentences to describe pain that they had experienced and some (n=43, 5%) used discriminative comparisons to describe different types of pain. Similar findings were reported in a study in North America (Abu-Saad, 1984). The children used 17 sensory words and one evaluative word. The words were then compared with the score they had given on the VAS. This revealed that words such as "aching", "stinging", “stabbing”, “throbbing”, “hot” and “sharp” were consistently used when pain intensity was rated as greater than 7.0 on the visual analogue scale (VAS). These findings reveal that the majority of children are able to describe their pain experiences, given the opportunity to express it in age-appropriate language. However, it remains to be seen whether it is possible to rate the descriptors that children use according to a pain intensity scale.

Jerret and Evans (1986) attempted, through a series of interviews with 40 Canadian children aged 5-9.5 years who were undergoing treatment for acute conditions, to establish the words children use to describe pain and to categorise them in a similar way to the MPQ. The children were found to use a number of adjectives and phrases to describe pain. The children suggested 60 words which did not feature on the MPQ. These included “weird”, “attacking”, “sausage” and “cymbals clapping”. Sensory words were used by 95% of the children and affective words by 60%. The adults encountered problems trying to rank the intensity of words such as “weird” and “sausage” as they felt that, without knowing the context of these words, it was impossible to rank them. The design could thus have been improved had children been involved in ranking the items as it is clear that adults found this difficult to do, as they were not familiar with the terminology.

Another group in North America attempted to create a word list that could be used to measure pain quality in children (Wilkie et al., 1990). The first part of the study required children to generate a word list to describe pain. One hundred and twenty nine words used by children to describe pain were identified from the literature and these were printed onto individual cards. The children were asked to sort the cards into piles representing: 1) words they knew and used to describe pain; 2) words they knew but would not use to describe pain and 3) words they did not know. To be included in the word list a criterion was set of at least 50% of the children knowing and using the word. In the first student group, 43 words met this criterion and these were used to construct a word list similar to the MPQ. This word list was then tested with 175 hospitalised children who were experiencing pain. The children were also asked to rate their pain intensity on a VAS scale. Each of the words was chosen by at least three children. The mean number of words chosen was 6.9 (range=0–29) and an additional 49 words or phrases were volunteered by the children, however, none of these were suggested by more than five children. There was a significant relationship between the pain intensity and the number of sensory, affective and evaluative words chosen (Wilkie et al., 1990). Therefore children who have experienced significant pain may be more able to describe it than those who have experienced lower levels.

The MPQ contains nine words to describe the temporal (flickering, throbbing etc.) nature of pain, however, the word lists discussed previously were lacking this feature. An American study sought to investigate how children and adolescents describe this aspect of pain (Savedra et al., 1995). One hundred and twelve children participated and it was found that children used very few temporal words spontaneously, however, they were able to choose temporal words when presented with a list. This study demonstrates that children may find it difficult to describe the temporal nature of their pain unless they are given specific prompts. This may be relevant in discovering how children describe dental pain. "Toothache" is often described as a "throbbing" pain by adults but children may not describe it in this way themselves as "throbbing" is a temporal term.

In summary, these studies demonstrate that a word list may be a useful tool to assess pain in children, using language they are familiar with. However, it should also be recognised that children may wish to add their own words in order to explain how they feel (Wilkie et al., 1990). It is important to appreciate the differences which may be present between adult and child pain descriptors and this highlights the importance of involving children in the development of items for questionnaires to ensure the correct terminology is used.

4.4.3 Children's description of dental pain

To date few studies have considered the language that children use to describe dental pain. A study by Toole and co-workers (2000) investigated the language children used to describe the discomfort felt during electric pulp-testing. Seventy-eight children aged between five and thirteen years of age participated and were asked to describe the stimulus felt during both a simulated electrical pulp-test and an actual electrical pulp-test. The children were asked to pick a word from the 56 word list used by Wilkie and colleagues (1990), as described previously. The children were also asked to rate, on a four-point scale, how "tingly" or "sore" the procedure was. The majority (73%) of children chose more words to describe the actual test. The mean number of words used to describe the real test was 9.76 compared with 4.61 for the simulated test. The most common words chosen for the real test were "uncomfortable", "like a sting", "sore" and "stinging". The words which best discriminated between the two tests were "stabbing", "shooting", "stinging", "hitting", "sore", "punching" and "dizzy". On the rating scales the mean score for the term "sore" was 1.23 for the actual test and 0.57 for the simulated test and 1.43 and 0.66 respectively for "tingly". Interestingly, one third of the children described the real pulp-test as "not at all sore" and 24% stated that there was no "tingling". The results of this study are consistent with others carried out using similar scales showing that, as discomfort increases, more words are chosen to describe it. No one word was chosen by all the children and this underlines the importance of ensuring that children are given a wide choice of words which are age-appropriate to describe the quality of their pain. Children with better literacy skills, as assessed by means of the Weschler Objective Reading Dimension (WORD), chose fewer words to describe their pain, which may be due to their greater comprehension allowing them to choose only the most suitable words. However, it was also found that chronological age impacted upon the results, as with increasing age children chose more words to describe their pain, perhaps due to their greater experience. Therefore it would appear to be important to consider both chronological age and literacy when investigating pain quality in children.

A further study carried out at a UK Dental Hospital, investigated the use of language used by children to describe dental treatment (Harman et al., 2005). This study asked children aged 6-17 years to choose words prior to dental treatment and following treatment to assess if there were differences between anticipated discomfort and that which was actually experienced. In addition, participants completed a five-point rating scale to describe how "sore" or "tingly" the treatment was. The Child Dental Anxiety Scale and the Spielberger State-Trait Anxiety Inventory for Children were used to investigate anxiety within the study population and vocabulary was assessed by means of the British Picture Vocabulary Scale. Fewer words were chosen following treatment than had been selected prior to treatment and this trend was even more apparent in the more anxious children. There was no significant difference between the ratings for "sore" or "tingly" before and after treatment. However, the authors did not discuss which words the children used to describe their experience nor were any correlations drawn between the treatment received and the number of words chosen. This would have been of interest as the procedures varied; those which required the administration of local anaesthetic might have been expected to cause more discomfort than those which did not and were minimally invasive, i.e. fissure sealants, denture adjustment.

Overall, these few studies do provide some preliminary information as to how children describe their dental experiences. However, given that the word list was developed for and by North American children it may not be appropriate in a United Kingdom setting. These studies did not allow children to add their own words and therefore some words may not have been chosen because the children did not know them. As children were not able to add their own descriptions, it remains unclear whether the words children use to describe dental pain differ from those used to express pain elsewhere.

The next section will discuss the concept of health-related quality of life (HRQoL) and reviews guidance for the development of measures of HRQoL.

4.5 Health-related quality of life

The WHO defines health as "a state of complete physical, mental, and social wellbeing and not merely the absence of disease, or infirmity" (World Health Organization, 1948). Furthermore it states that "The enjoyment of the highest attainable standard of health is one of the fundamental rights of every human being without distinction of race, religion, political belief, economic or social condition" (World Health Organization, 2007b). These definitions underpin the philosophy that health is not merely concerned with the absence of disease but has more complex interactions with external factors.

The dominant health-related model from the late 19th century can be described as a biomedical model, which has a focus on the pathological processes which affect the body. This neglected to consider what constitutes health and illness. A broader approach, involving these considerations, can be described by a biopsychosocial model of health and health research has focused on this model over the last few decades (Engel, 1977; Alonso, 2004).

4.5.1 The biomedical model of disease

In 1980, the WHO published the International Classification of Impairments, Disabilities and Handicaps as a manual of classification relating to the consequences of disease (World Health Organization, 1980). The model uses a conceptual framework for disability and is described in three dimensions:

1. Impairment: In the context of health experience this is defined as any loss or abnormality of psychological, physiological or anatomical structure or function.
2. Disability: In the context of health experience a disability is any restriction or lack (resulting from an impairment) of ability to perform an activity in the manner, or within the range, considered normal for a human being.
3. Handicap: In the context of health experience a handicap is a disadvantage for a given individual, resulting from an impairment or a disability, that limits or prevents the fulfillment of a role that is normal (depending on age, sex and social and cultural factors) for that individual.

This model has been criticised, as although it aimed to encompass the social aspects of disablement it did not consider the ways in which an unaccommodating environment may impact upon the handicap (Bichenbach et al., 1999). In addition, assessment of population health needs based solely on disease surveillance will fail to account for the impact of the diseases

(Sheiham et al., 1982; Sheiham and Spencer, 1997). From a clinical perspective, interventions which are directed only at correcting physical abnormalities, have limited ability in enabling a full recovery for the patient as they fail to address external factors which may influence that recovery (Engel, 1977).

4.5.2 The biopsychosocial model of health

The biopsychosocial model focuses on health and illness rather than on disease. This recognises the interplay between physiological processes and social factors which influence health and wellbeing. In 2002 the WHO International Classification of Functioning, Disability and Health changed its perspective from the 1980 model to become a "components of health" classification instead of a "consequences of disease" classification (World Health Organization, 2002). This ensures that not only are the biological and physiological features investigated but consideration is also given to how the disease affects participation in society (Figure 4.2).

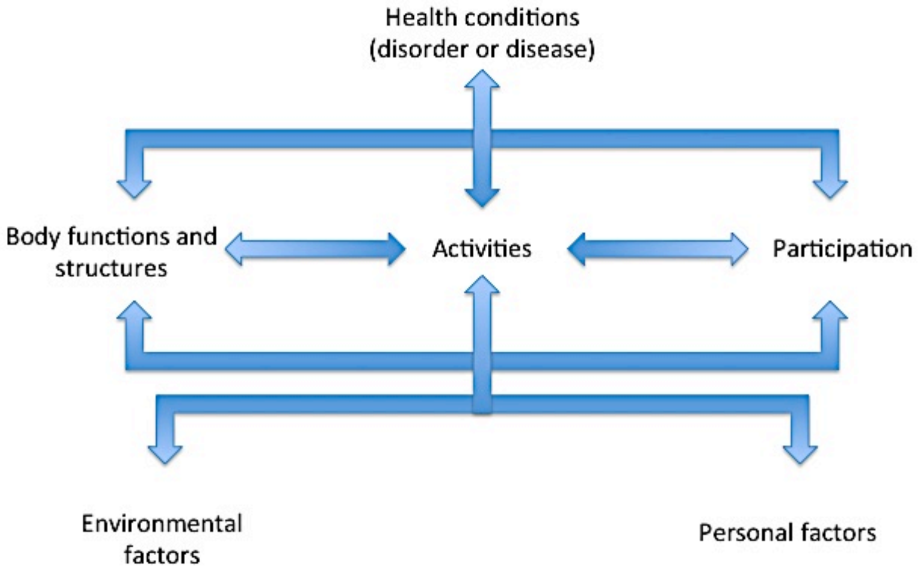

Figure 4.2. Interactions between the components of the International Classification of Functioning, Disability and Health (World Health Organization, 2002).

The assessment of HRQoL aims to assess not only functional but psychosocial factors associated with disease and therefore takes a biopsychosocial rather than biomedical approach. The next section will discuss the concept of HRQoL and how it is measured.

4.5.3 Health-related quality of life

Quality of life has been described as: "An individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their personal goals, expectations, standards and concerns." (World Health Organization, 1948). This can be extended to health and encompass the impacts that individuals may experience as a result of the state of their health. There has been much debate regarding the concept of HRQoL as opposed to "subjective health status". Bullinger (1993) suggested that "subjective health status" should relate to patient-based assessments of health state, which may be termed patient-reported outcome measures (PROMs). Whereas, HRQoL focuses on the impact of perceived health state on an individual's ability to live a fulfilling life and so encompasses external factors in addition to the individual's health state (Bullinger et al., 1993).

Measures of health-related quality of life

As can be seen in Figure 4.2, there may be many aspects of life which may impact upon health. In order to assess the effect of interventions on health, evaluation may be carried out to measure changes in HRQoL using measures specifically designed to evaluate this construct, rather than simply subjective health status.

Different types of measure of HRQoL exist and they can be used in epidemiological studies, clinical trials and to assess treatment success in individuals. To be reliable they must be tailored to meet their objective. Evaluative measures must be responsive and demonstrate longitudinal construct validity, i.e. they must be able to detect small changes within participants and there must be correlation between changes in the new questionnaire and changes in another measure (Guyatt et al., 1993). Generic measures are useful to compare populations and can be used to compare groups with different health conditions, whereas disease-specific measures can be used to track changes in individuals with a specific disease.

Discriminative measures require high reliability and cross-sectional construct validity, i.e. there must be a high ratio of differences between participants due differences in the participants and appropriate correlations between the instrument tested and established scales (Guyatt et al., 1993). Further description of the properties required of a measure of HRQoL will be provided in Section 4.5.4.

The development of child-centred measures of HRQoL is particularly challenging as the meaning of quality of life may change with the developmental stage of the child. Thus, in order for them to be sensitive they should address the concerns of children of different ages. However, if they are to be used in clinical trials to assess change over time, targeting a narrow age range may not be desirable.

To examine changes over time, an instrument which is sensitive to small changes is required. Wiebe and co-workers (2003) found that in randomised controlled trials with a true underlying therapeutic effect, disease-specific instruments were more responsive to change than generic instruments (Wiebe et al., 2003). As the aim of this study will be to produce a disease-specific measure for dental caries, the next section will focus on recommendations for producing disease-specific measures.

4.5.4 Developing disease-specific measures

Guyatt and co-workers (1986) make suggestions for how reliable disease specific measures should be produced and states "items on the questionnaire must reflect areas that are important to those suffering from the disease". To ensure items reflect those important to those with the condition, development of the measure should involve conducting semi-structured interviews with 50-100 affected people in order to elicit all relevant items; the item pool can then be reduced by asking another group with the condition to rank the importance of the items to them. This may still yield a higher number of items than is desired and approaches for further reducing these depend upon the how the measure is intended to be used. For evaluative measures, it may reduce the measure's responsiveness if items are included that would be unlikely to change due to the intervention provided. Each dimension measured should be adequately represented to decrease the variability in response and to minimise the impact of idiosyncratic responses. A method suggested to choose items for inclusion is to multiply the frequency of each item by its mean importance and to choose those ranked highest for the final questionnaire. Item reduction will be discussed in more detail in Chapter Six.

Guyatt and colleagues (1986) also suggest that, in order to accurately chart changes following an intervention it is not appropriate to ask questions with only "yes" and "no" responses as this may not reflect small changes. There is no evidence to support either a Likert scale or a Visual Analogue Scale but these both allow small changes to be seen. Furthermore, it is also suggested that it may be helpful for patients to see their previous responses as this decreases the variability of responses from stable patients and that changes in response to treatment are similar between informed and blind subjects (Scott and Huskisson, 1979; Guyatt et al., 1985). These aspects can be used to develop measures which are relevant to the target population.

To assess the validity of the measure, various psychometric properties must be evaluated. The psychometric properties of reliability, validity and responsiveness can be tested in accordance with recommendations of the Scientific Advisory Committee of the Medical Outcomes Trust (Lohr, 2002). The Scientific Advisory Trust of the Medical Outcomes Trust initially published a set of criteria for assessment of health status and quality of life measures in 1996 (Lohr et al., 1996). These were updated in 2002 to reflect the emerging techniques being used in the development of these measures (Lohr, 2002). They suggest eight key areas for consideration (conceptual and measurement model; reliability; validity; responsiveness; interpretability; respondent and administrative burden; alternate forms and cultural and language adaptations) and criteria against which measures can be reviewed. These are summarised in Table 4.2.

Table 4.2. Attributes and criteria for review of measurement properties (Lohr, 2002).

Attribute	Review criteria
Conceptual and measurement model	• Concept to be measured • Conceptual and empirical bases for item content and combination • Target population involvement in content derivation • Information on dimensionality and distinctiveness of scales • Evidence of scale variability • Intended level of measurement • Rationale for deriving scale scores
Reliability	*Internal consistency* • Methods to collect reliability data • Reliability estimates and standard errors for all score elements (classical test) or standard error of the mean over the range of scale and marginal reliability of each scale (item response theory) • Data to calculate reliability coefficients or actual calculations of reliability coefficients • Above data for each major population of interest, if necessary *Reproducibility* • Methods employed to collect reproducibility data • Well-argued rationale to support the design of the study and the interval between the first and subsequent administration to support the assumption that the population is stable • Information on test-retest reliability and inter-rater reliability based on intraclass correlation coefficients • Information on the comparability of the item parameter estimates and on measurement precision over repeated administration
Validity	• Rationale presented for the particular mix of evidence presented for the intended uses • Clear description of the methods employed to collect validity data • Composition of sample used to examine validity • Above for each major population of interest • Hypotheses tested and data relating to the tests • Clear rationale and support for the choice of criteria measures
Responsiveness	• Evidence of changes in scores of the instrument • Longitudinal data that compares a group that is expected to change with a group which is expected to remain stable • Population(s) of which responsiveness has been tested, including the time intervals of assessment, the interventions or measures involved in evaluating change, and the populations assumed to be stable
Interpretability	• Rationale for selection of external criteria of populations for purposes of comparison and interpretability of data • Information regarding the ways in which data from the instrument should be reported and displayed • Meaningful "benchmarks" to facilitate interpretation of the scores
Burden	*Respondent burden* • Information on (a) average and range of the time needed to complete the instrument, (b) reading and comprehension level, and (c) any special requirements or requests made of respondents • Evidence that the instrument places no undue physical or emotional strain on the respondent • When or under what circumstances the instrument is not suitable for respondents *Administrative burden* • Information about any resources required for administration of the instrument • Average time and range of time required of a trained interviewer to administer the instrument in face-to-face interviews, by telephone, or with computer-assisted formats • Amount of training and level of education or professional expertise and experience needed by administrative staff
Alternative modes of administration	• Evidence on reliability, validity, responsiveness, interpretability and burden for each mode of administration
Cultural and language adaptations or translations	• Methods to achieve conceptual equivalence • Methods to achieve linguistic equivalence • Any significant differences between the original and translated versions • How inconsistencies were reconciled

These guidelines were developed to help the Medical Outcomes Trust to evaluate new measures which were submitted to them, to ascertain which were suitable for dissemination. However, although they provide clear information regarding areas to be assessed, no specific quality standards are given.

More recently, a checklist has been produced by the Consensus-based Standards for the Selection of Health Measurement Instruments initiative (COSMIN) which allows articles reporting on the evaluation of Patient Reported Outcome Measures (PROMs) to be evaluated against defined criteria (Mokkink et al., 2010a). These criteria will be discussed in detail in Chapter Five.

4.5.5 Disease-specific measures

Over the last decade, there have been advances in the production of disease-specific measures of HRQoL measures formulated specifically for children. A systematic review published in 2008 identified 64 disease-specific HRQoL measures for children (Solans et al., 2008). Of these, 42 had been produced since 2001 and 27 different conditions were covered. Ten measures existed for asthma, eight for oncology and seven for epilepsy. CPQ and the C-OIDP were included, however, as discussed later, although these relate to oral health, they are not disease-specific. Most of the measures were developed for children aged five years and over, although some had components relating to very young children. Less than half (43.7%) relied exclusively on child reports, 26.6% on parental report only and 29.6% incorporated child and parent components. One measure included a nurse-reported version. The most common concepts to be explored were emotional wellbeing, friends/social functioning, physical function, symptoms and treatment. This systematic review did not comment on the methods of developing the measures and whether these were child-centred.

Child-centred approaches have been used by some groups to develop disease-specific measures. The Canadian Haemophilia Outcomes-Kids Life Assessment Tool (CHO-KLAT) used a child-centred approach to develop their measure. Firstly, an item bank was created by a review of the literature followed by focus groups with children to develop further unique items. The value of including the children was assessed and the investigators found that the children generated unique items not found from other sources and also suggested rewording (Young et al., 2004).

This child-centred approach has also been used successfully to create disease-specific measures for asthma, foot and ankle problems and inflammatory bowel disease (Christie et al., 1993; Griffiths et al., 1999; Morris et al., 2007). These studies also revealed that items suggested by health professionals were not always relevant to children. The children were also able to suggest alternative wording and response formats which were age-appropriate.

In summary, children have been involved in the development of a number of disease-specific measures and this has proved to be successful, with children suggesting relevant items and response formats. There are clear benefits to including children in this way. The following section will discuss the development and validation of the most commonly used self-report measures of OHRQoL for children.

4.6 Oral health-related quality of life measures

OHRQoL has been defined as "the extent to which oral disorders affect functioning and psychosocial wellbeing" and as the "the impact of oral conditions on daily functioning" (Kressin, 1997; Locker et al., 2000). The application of measures can vary according to the aim of the investigation, for example, they may be used to influence health and social policy, assess the impact of different treatment regimens or be used to analyse change in individual patients over time (Table 4.3). A recent review identified 16 OHRQoL measures, two of which are specifically aimed at children (Locker and Allen, 2007). In this section the development of the three most widely used measures of OHRQoL for children will be discussed.

Table 4.3. Summary of the applications of OHRQoL measures proposed by Robinson and co-workers (Robinson et al., 2003).

Type of application	Description of use
Theoretical	Exploring models of oral health Describing factors influential to health
Political	Demonstrating involvement of the public in healthcare Identifying the public's priorities Advocacy
Practical	Planning, monitoring and evaluating services
Public health	Needs assessments
Research	Evaluating outcomes of healthcare interventions
Clinic-based	Evaluating individual patient care Improving patient-practitioner communication Clinical audit Marketing of services

4.6.1 Child Perceptions Questionnaire

The most frequently used measures are part of the Child Oral Health Quality of Life battery of questionnaires which comprises the Parental Perceptions Questionnaire (PPQ), the Family Impact Scale (FIS) and age-specific versions of CPQ designed for those aged 11-14 years (including short form versions) or 8-10 years (Jokovic et al., 2002; Jokovic et al., 2003b; Jokovic et al., 2004; Jokovic et al., 2006). The development of each of the CPQ instruments will be discussed in detail.

CPQ$_{11-14}$

The CPQ$_{11-14}$ was originally developed as a measure for children with a wide range of dental, oral and oro-facial disorders (Jokovic et al., 2002). The item pool was developed by initially reviewing the existing oral health and child health status measures, followed by revision of this pool by healthcare professionals and parents of children with orofacial disorders. These items encompassed four domains: oral symptoms, functional limitations, emotional and social wellbeing. The modified item pool was further revised following in-depth interviews with 11 child patients. Finally, groups of children from three different disease classifications (dental disease, orthodontic disorders and orofacial conditions) participated in an item impact study to select the items for inclusion in the final measure using a questionnaire with a severity-based response format. The items were ranked within the four domains and according to clinical group. Any items which were above the median in either ranking were chosen for inclusion in the final CPQ_{11-14} measure.

The final instrument asks about the frequency of events in the preceding three months, a contrast to the item impact questionnaire which was used during the development stages. A five-point Likert scale is used ranging from: "Never"=0 to "Every day/ almost every day"=4, for example: *"In the past three months have you had pain in your teeth, lips, jaws or mouth?"*. The final version of CPQ_{11-14} contains 37 items with possible scores ranging from 0-148, higher scores representing higher impacts of oral disease. To test validity of the measure, global ratings of oral health and life overall were included.

The instrument was then evaluated for reliability and validity using a new sample of 123 children, 70 of whom completed the questionnaire again two weeks later. The children who completed the questionnaire a second time were asked to report any changes in their condition and those who experienced changes were excluded from the test-retest reliability part of the study. Internal consistency was found to be acceptable/good with Cronbach's alpha scores ranging from 0.91 for the total CPQ score to 0.64 for oral symptoms. The Cronbach's alpha scores were similar between the clinical groups. The intraclass correlation coefficient (ICC) was 0.90 for the overall CPQ score and ranged from 0.79 to 0.88 for the four domains.

Construct validity was tested by analysing the relationship between the CPQ scores and the global ratings for oral health and life overall. There was a significant positive correlation between these, $p=0.013$ for oral health and $p=<0.001$ for life overall. Significant correlations were found for all health domains with the exception of the oral health rating and the functional limitations score. A positive correlation was found between the number of carious surfaces and overall scale scores with those children with more carious surfaces having higher scores. This was also true in the orofacial group with children with cleft palate alone having lower scores than those with bilateral cleft lip and palate.

The development of this measure demonstrated that it is possible for children to complete self-administered questionnaires as very few had missing answers. This may in part be due to asking children which questions were suitable for inclusion. However, children were involved only at a later stage and the final items and response format were produced by adults in a different format to that used in the previous item impact questionnaire. Inclusion of children in the planning stage may have been beneficial as this would have ensured that the measure truly reflected those items which are important to children and used language which they are familiar with. No information is available regarding the mean number of carious surfaces that these children had and therefore it is impossible to assess whether the correlations apply outwith this population.

CPQ$_{8\text{-}10}$

A measure for 8-10-year-olds (CPQ$_{8\text{-}10}$) was adapted from the CPQ$_{11\text{-}14}$. This was achieved by consulting a child psychologist, a teacher and parents of children of this age group to ascertain which questions were appropriate. The wording was then rewritten by a writer of children's manuals and the teacher. Children were not involved in initial development of this measure. Grammar and readability were assessed using the Flesch reading ease score and a Flesch-Kincaid grade level score. These were found to be acceptable for children of this age. The final measure included 25 questions across the four health domains as discussed previously.

Following self-administration, a group of children participated in qualitative interviews to assess their understanding of the questions, wording and recall intervals. This found that the 9-10-year-old children were able to understand the questions but 10 of the 13 eight-year-olds did not relate the introductory statement: "*In the past four weeks, because of your teeth or mouth….*", to the rest of the question. Therefore this statement was included as part of each question.

Validity of the measure was assessed by 68, 8-10-year-old children who either had caries (paediatric group) or cleft lip and/or palate (orofacial group).

Thirty-nine of the children completed the measure a second time a fortnight later in order to evaluate the measure's test-retest reliability.

The internal consistency of the overall scale was good (Cronbach's alpha=0.89). Within the domain subscales Cronbach's alpha ranged from 0.63–0.88 indicating acceptable internal consistency. Of the 39 children who completed the questionnaire a second time, only 33 were suitable for inclusion in the analysis. The remaining six children were excluded as they had indicated that their oral status had changed in the preceding two weeks. The ICC demonstrated substantial reliability for the overall scale (ICC=0.75), oral symptoms, functional limitations and emotional wellbeing domains. However, the ICC for the social wellbeing domain was low at 0.16.

The measure appeared to respond as one would expect, with those children with higher levels of caries having higher overall scores. However, it was unable to discriminate between the paediatric and orofacial groups. There was a low agreement for test-retest reliability for the social wellbeing domain. The authors explain that this may be due to changes between the two test periods and that social experiences are more likely to change than emotions or symptoms during this period. However, this was not found in 11-14-year-olds and therefore this change may be particular to the younger age group, although no explanation was given regarding reason why social aspects may have changed over a relatively short period. The measure was constructed by adults with little involvement from the children themselves and this may account for its less than ideal properties.

CPQ short forms

In order to facilitate the use of the CPQ in clinical trials and population studies, it was shortened to a 16- and 8-item measure using both item impact and stepwise regression methods. Data from the original item impact study were used to develop CPQ_{11-14}-ISF:8 and CPQ_{11-14}-ISF:16 (Jokovic et al., 2006). These contain two or four questions from each domain respectively. The stepwise regression method of development utilised the data from the evaluation of the full version. The two or four questions which were the best predictors of the overall score were included in the resultant short form measures.

CPQ_{11-14}-ISF:16 and CPQ_{11-14}-RSF:16 share 14 of their 16 items (Table 4.4.) The questions specific to the CPQ_{11-14}-ISF:16 are:
1. Difficult to drink or eat hot or cold foods?
2. Other children asked you questions about your teeth, lips, jaws or mouth?

Those specific to the CPQ_{11-14}-RSF:16 are:
1. Had trouble sleeping?
2. Not wanted to speak or read aloud in class?

In contrast the eight item versions only had two questions in common:
1. Had bad breath?
2. Felt upset about their teeth, lips, mouth or jaws?

The measures were evaluated in terms of their reliability, construct and criterion validity. Analysis of the internal consistency using Cronbach's alpha demonstrated acceptable reliability with scores ranging from 0.71–0.83. All forms performed well in terms of the test-retest reliability with the ICC ranging from 0.71–0.77.

Criterion validity was assessed by correlating the short form questionnaires with the original long version. Almost perfect correlation was found with the exception of the CPQ_{11-14}-ISF:8.

Discriminant construct validity was found to be good as all short forms detected differences between the clinical groups as expected. All short forms had increased precision when compared with the long version with the exception of the CPQ_{11-14}-RSF:16. Children in the group with dental caries who had more than 10 carious surfaces had higher overall mean scores on all versions than those with less than 10 carious surfaces. However, this was not statistically significant and it was not clear why 10 surfaces was chosen as the cut-off point where changes in mean score would be expected.

Table 4.4. Questions in each short form version of the Child Perceptions Questionnaire.

Item	ISF16	RSF16	ISF8	RSF8
Oral symptoms				
Pain in teeth/mouth	✓	✓	✗	✗
Bad breath	✓	✓	✓	✓
Mouth sores	✓	✓	✗	✓
Food caught between teeth	✓	✓	✓	✗
Functional limitations				
Difficulty eating or drinking hot or cold foods	✓	✗	✓	✗
Difficulty chewing firm foods	✓	✓	✓	✗
Trouble sleeping	✗	✓	✗	✓
Difficulty saying words	✓	✓	✗	✓
Taking longer to eat a meal	✓	✓	✗	✗
Emotional wellbeing				
Upset	✓	✓	✓	✓
Felt irritated/frustrated	✓	✓	✓	✗
Felt shy	✓	✓	✗	✗
Concerned what people think about teeth/mouth	✓	✓	✗	✓
Social wellbeing				
Asked questions	✓	✗	✓	✗
Teased/called names	✓	✓	✗	✓
Not wanted to speak/read out loud in class	✗	✓	✗	✗
Avoided smiling/laughing	✓	✓	✓	✗
Argued with children/family	✓	✓	✗	✓

ISF16=16-item item impact version; ISF8=8-item item impact version; RSF16=16-item regression version; RSF8=8-item regression version.

All short forms of the measure demonstrated positive significant correlations with the global oral health and life overall ratings. The correlation was greater for the rating of life overall than for the global oral health rating.

In summary, despite its limitations, CPQ and its short form versions have been used extensively in studies of the oral health of children in many countries and has been translated into several languages (Foster Page et al., 2005; Marshman et al., 2005; Robinson et al., 2005; Brown and Al-Khayal, 2006; Alm et al., 2007; Do and Spencer, 2007; Barbosa et al., 2009; Koposova et al., 2010; Bekes et al., 2011; Wogelius et al., 2011). It has also been used to investigate a variety of conditions including caries, cleft lip and palate, dental trauma, fluorosis and malocclusion (Locker et al., 2005; Robinson et al., 2005; Do and Spencer, 2007;

O'Brien et al., 2007; Zhang et al., 2007a; Marshman et al., 2010; Traebert et al., 2012).

4.6.2 Child-Oral Impact on Daily Performances index

C-OIDP, in contrast to CPQ, was developed to assess dental needs in child populations. This measure also differs in that it is designed to be administered by interview rather than self-completed. It was adapted from the Oral Impact on Daily Performances index for adults and was developed for use in a Thai population (Gherunpong et al., 2004a). The original index was used and face validity assessed by 513, 11- and 12-year-old children from schools in Thailand. Following this, a number of amendments were made: language was simplified; performances explained by giving specific examples; response options limited to four instead of five; recall interval shortened to three months, and pictures added to aid the interview process. Content validity was assessed by interviews with health professionals. The pilot questionnaire was tested for reliability and validity and one item "doing light physical activity" was removed as it had very poor correlations with other items. The final version includes eight performances: 1) eating; 2) speaking; 3) cleaning teeth; 4) relaxing (including sleeping); 4) smiling; 5) laughing and showing teeth without embarrassment; 6) maintaining emotional state; 7) study (including going to school and doing homework) and 8) contact with other people. If children report an impact, they are asked to grade the frequency and severity of this. These are then multiplied to give a cumulative score (range 0-9 for each performance) for that performance. Adding the scores for each performance, dividing by 72 and multiplying by 100 is used to derive an overall score. An alternative scoring system is available which aims to separate "intensity" from "extent". Therefore only the highest severity score is used across the eight performances to give the "severity" (range=1-6) and the total number of performances affected is calculated to give an indication of the "extent" (range=0-8).

The final version was tested using a sample of 110 children aged 11-12 years (mean 11.3 years). Internal consistency was adequate (Cronbach's alpha=0.6) and children with a greater perceived need for dental treatment, scoring more highly (Gherunpong et al., 2004a).

This measure has been used extensively in epidemiological studies across the world and has been translated into several languages and can be adapted to cover specific conditions (Tubert-Jeannin et al., 2005; Mashoto et al., 2009; Bianco et al., 2010).

4.6.3 Child Oral Health Impact Profile

The final measure to be discussed is COHIP, which was developed from the same item pool as the CPQ and is designed to be used in children aged 8-15 years (Broder et al., 2007). Testing of the measure was carried out at four hospital sites in Newark and New York in the USA and at McGill University in Montreal. Following the development of the initial item pool (54 items), face validity was tested with 10 health professionals and 144 carers of children with caries, malocclusion and craniofacial disorders. Following this, items were reworded and 42 items retained. Item impact testing of this revised pool was performed with 155 children and their parents. Participants were asked if they had experienced the event in the past three months and how important they felt it was. Eight items were subsequently deleted and 17 new items developed based on participants' comments, giving a new total of 51 items. A second stage of face validity testing was then undertaken with the new pool. The participants in this stage were 50 children and 55 carers. The second stage of face validity testing resulted in two items being removed and the remaining 49 items were used in the second phase of item impact testing. The final stage of item impact testing resulted in nine items with low impact scores being deleted. However, low impact items related to school were retained as they were deemed to be theoretically important despite the fact that children did not endorse them. Finally, the questionnaire was subject to factor analysis. The measure was administered to 419 children and the final decision on which items should be retained was based on the exploratory factor analysis, correlations between items and discriminatory power. The final measure comprised 34 items across five domains: functional wellbeing; psychological wellbeing; social wellbeing; school and self-image, and incorporated pictures and shading to prevent test-taker fatigue. In common with the CPQ, children are asked if they have experienced events over the previous three months and items are scored on a five-point Likert scale: “never”=0, “almost never”=1, “sometimes”=2, “fairly often”=3 and “almost all the time”=4. The self-image items are scored from “strongly disagree” to “strongly agree”. In addition, there are two questions about treatment expectations and one global rating question.

Reliability and validity of the COHIP was tested in a sample of 523 children aged 8-15 (mean=11.6 years) years (Broder and Wilson-Genderson, 2007). The original sample consisted of 548 children, however, those who completed less than 75% of the items were excluded. The final sample included 157 children with caries, 152 with malocclusions, 110 with craniofacial anomalies and 104 school children. Cronbach’s alpha for the questionnaire overall was 0.91, with a range of 0.68-0.89 across the subscales. Test-retest reliability was analysed using data from 44 participants who reported no change to their oral health in the intervening 3-week period. Test-retest reliability was excellent for the overall scale (ICC=0.84). Discriminant validity was tested and found to be statistically significant for the functional wellbeing and school subscales and *post hoc* testing revealed that the craniofacial group had higher scores than the other groups. The number of carious surfaces correlated with the overall score as did degree of overjet. Correlations were found with the global rating in all four groups confirming good convergent validity.

The development of this measure did involve children, although the extent of this is not clear, as during item impact testing measures were given to the child with their parent and it is not clear who completed these. Items were added in response to participant comments but these were in the minority when compared with the items already generated by the literature review and expert opinion. In addition, some items were retained as they fitted a theoretical construct despite the fact that they were not endorsed by children. This casts some doubt as to the extent to which children’s views were taken into consideration. A more child-centred approach would have ensured that items important to children were included in the initial stages and may have led to a measure which is more meaningful to children.

This questionnaire has not been as widely used as the CPQ but has been translated into Dutch and Persian and used in some epidemiological and clinical investigations (Geels et al., 2008a; Geels et al., 2008b; Calis et al., 2009; Bos et al., 2010; Ravaghi et al., 2011). Recently, a 19-item short form version has been developed, but has not yet been used widely (Broder et al., 2012).

4.6.4 Summary

The OHRQoL measures described above have been extensively used since their development in a variety of studies with varying results. However, although all have involved children to some extent in their development, none have used methods which can be truly described as child-centred. Item pools were created by the researchers from literature reviews, interviews with health professionals and from existing measures for adults, thus leaving children to comment only on existing items. Guidance on production of these measures clearly advises that, in order to be reliable and truly reflect patient views, patients should be involved at all stages.

Although CPQ and COHIP were developed to be used in clinical trials, they have not yet been validated in this context. As they cover a variety of clinical conditions, they are generic instruments and therefore may not be sensitive enough to longitudinal change in order to evaluate changes following clinical interventions for specific conditions (Wiebe et al., 2003). Therefore, the development of a child-centred measure which is caries-specific and sensitive to longitudinal change is required to assess patient-reported outcomes in clinical trials evaluating interventions for dental caries. The measure could also be used to influence public health policy by determining what the impacts of dental caries are for children and how these change following different treatments. This would allow recommendations to made regarding funding of primary care or specialist services. In addition, the measure could be used alongside measures of general HRQoL and preference-based utility measures (indeed, these types of measures have been used in studies investigating children's OHRQoL (Page et al., 2014; Paula et al., 2015) to allow comparisons to be made regarding the impact of dental caries with other health conditions. Although it remains to be seen whether measures of health-related quality of life or preference-based health utility measures are responsive to changes in OHRQoL.

It should also be noted that, despite their widespread use, these measures have not been evaluated using quality criteria such as those suggested by the COSMIN group. To fulfill the objectives of this thesis, each of these measures will be assessed using existing quality criteria in Chapter Five.

4.7 Publications related to the work described in this Chapter

4.7.1 Peer-reviewed journal article

- Gilchrist F, Rodd, HD, Deery C and Marshman Z. Involving children in research, audit and service evaluation. *British Dental Journal* 2013; **214**: 577 – 582.

Chapter Five

Systematic review of oral health-related quality of life measures for children

5.1 Introduction

It is important that measures of HRQoL/OHRQoL are developed in a robust way and validated appropriately to ensure that outcomes are reliable. Measures which have good measurement properties will allow smaller treatment effects to be observed, thus meaning smaller sample sizes can be used in clinical trials and that stronger conclusions can be drawn from the results (Guyatt et al., 1987a). This, in turn, will allow clinicians to choose the best treatments for their patients.

The previous chapter discussed the development of the three most commonly used measures of OHRQoL for children. This chapter will present a systematic review of the literature of the development of these measures and their evaluations to date against recognised criteria. An understanding of the strengths and limitations of these existing measures will aid the development of the proposed new measure ensuring that it is robust and sensitive to clinical changes.

Guidance exists for appraisal of HRQoL measures and PROMs (Gill and Feinstein, 1994; Guyatt and Cook, 1994; Lohr, 2002). Gill and Feinstein (1994) formulated criteria against which to appraise measures which purported to measure HRQoL. These were divided into "Investigator-specific criteria" and "Instrument-specific criteria". Investigator specific items were:

- Did the investigators define what they meant by quality of life?
- Did the investigators state the domains they intended to measure as components of quality of life?
- Did the investigators give reasons for choosing the instruments that were used
- Were the results aggregated into a single score for quality of life?

Instrument specific criteria were as follows:

- Were patients asked to give a global rating for quality of life?
- Was overall quality of life distinguished from HRQoL?
- Were patients able to provide supplemental items?
- Were patients asked to indicate which items were particularly important to them?
- If so, were these importance ratings incorporated into the final score?

In using these criteria, the authors found that the majority of studies analysed failed to meet these standards. However, others have argued that whilst these criteria philosophically meet the requirements of measuring HRQoL, it is difficult to compare across groups when patients can substitute their own items (Guyatt and Cook 1994). Guyatt and Cook (1994), therefore offered a set of less stringent criteria against which measures could be assessed which were further modified by Locker and Allan (2007) following their review of OHRQoL measures . These are summarised as follows:

- Is the stated aim to measure OHRQoL and is this explicit? If so, are these constructs defined and are the constituent domains identified?

- If not, is there an alternative construct measured by the instrument specified and defined and its constituent domains identified?

- Do the investigators specify the contexts in which the measure is to be used? Was it developed for use with groups (as in surveys or clinical trials) or individuals (as in clinical practice)?

- Were the items comprising the questionnaire derived from qualitative interviews with those intended to complete the questionnaire?

- Is there evidence that the aspects of life that the items address are important to those who will be completing the questionnaire?

- Does the questionnaire contain global ratings of health-related quality of life or quality of life?

- How was the measure validated? Was it tested against oral health indicators or were broader indicators that may capture aspects of quality of life used? Is the stated aim to measure OHRQoL and is this explicit? If so, are these constructs defined and their constituent domains identified.

These criteria evaluate the patient-centered nature of the measure but do not address the underlying psychometric properties required to validate such a measure. As discussed in Section 4.5.4, the COSMIN checklist was developed in 2010 (Mokkink et al., 2010a). Following an international Delphi study, a checklist was developed which encompassed all the measurement properties which the panel felt were important and the statistical methods which were appropriate to measure these. The original checklist had a dichotomous scoring system for each item, which was further refined in 2011 and a 4-point scale developed (Appendix A) (Terwee et al., 2012). It is hoped that the use of this checklist will standardise systematic reviews of PROMs. In addition, quality criteria for each measurement property can be used in conjunction with the COSMIN checklist (Terwee et al., 2007). These are shown in Table 5.1.

Table 5.1. Quality criteria based on those proposed by Terwee and colleagues (2007).

Property	Quality criteria*
Content validity	+ A clear description is provided of the aim of the measure, the target population, concepts being measured and involvement of the target population and/or investigators or experts in item selection ? A clear description of the above is lacking or only target population involved or doubtful design or method - No target population involvement 0 No information on target population
Internal consistency	+ Factor analyses on adequate sample size (7x the number of items and >100) and Cronbach's alpha calculated per dimension and between 0.7 and 0.95 ? No factor analysis or doubtful design or method - Cronbach's alpha <0.7 or >0.95 0 No information found on internal consistency
Criterion validity	+ Convincing argument that there is a "gold standard" and correlation >0.7 ? No convincing argument that gold standard truly is "gold" or doubtful design or method - Correlation with gold standard <0.7 0 No information on criterion validity
Construct validity	+ Specific hypotheses were formulated and at least 75% of the results are in accordance with these ? Doubtful design or method - Less than 75% hypotheses confirmed 0 No information on construct validity
Reproducibility	*Agreement* + MIC > SDC or MIC outside LOA or convincing arguments that agreement is acceptable ? Doubtful design or method or above not fulfilled - MIC >SDC or MIC equals or inside LOA 0 No information found on agreement *Reliability* + ICC or weighted Kappa >0.7 ? Doubtful design or method (e.g. time interval not mentioned) - ICC or weighted Kappa <0.7 0 No information on reliability
Responsiveness	+ SDC < MIC or MIC outside LOA or RR > 1.96 or AUC > 0.7 ? Doubtful design or method - SDC > MIC or MIC equals or inside LOA or RR < 1.96 or AUC < 0.7 0 No information on responsiveness
Floor or ceiling effects	+ < 15% of the respondents achieved the highest or lowest scores ? Doubtful design or method - > 15% of the respondents achieved the highest or lowest scores 0 No information found on interpretation
Interpretability	+ Mean and SD scores presented for at least four relevant subgroups of patients and MIC defined ? Doubtful design or method or less than four subgroups or no MIC defined 0 No information on interpretation

Property	Quality criteria*
Content validity	+ A clear description is provided of the aim of the measure, the target population, concepts being measured and involvement of the target population and/or investigators or experts in item selection ? A clear description of the above is lacking or only target population involved or doubtful design or method - No target population involvement 0 No information on target population
Internal consistency	+ Factor analyses on adequate sample size (7x the number of items and >100) and Cronbach's alpha calculated per dimension and between 0.7 and 0.95 ? No factor analysis or doubtful design or method - Cronbach's alpha <0.7 or >0.95 0 No information found on internal consistency

MIC=Minimal important change; SDC=smallest detectable change; LOA=limits of agreement; ICC=intraclass correlation; SD=standard deviation.

+=positive rating; ?=Indeterminate rating; -=negative rating; 0=no information available. *Doubtful design or method = lacking a clear description of the design or methods of the study, sample size smaller then 50 subjects or any other important methodological weakness in design or execution of the study.

The authors acknowledge that the criteria set are somewhat arbitrary and that they will require refinement over time; however, it is hoped that providing quality criteria will allow objective assessment of health status measures. By thus identifying areas of strength and weakness, further refinement of existing measures can be made. In addition, such objective measurement can aid the selection of the appropriate measures for use in clinical studies.

The above guidance overlaps and together provides a robust basis for the systematic review of measures of HRQoL. Those of Gill and Feinstein (1994) and Guyatt and Cook (1994) address the patient-centredness and content validity of the measure while those from the Medical Outcomes Trust and COSMIN focus on the psychometric evaluation. By combining these elements a robust and comprehensive review of measures of OHRQoL will be possible and help to inform the development of the new measure.

5.2 Aim

The aim of this next stage of the thesis is to assess the methodological quality of the development and testing of CPQ, C-OIDP and COHIP, the most widely used child-self report OHRQoL measures.

5.2.1 Objectives

To fulfill this aim, the specific objectives are to:

1. describe the above measures and their use
2. assess the methodological quality and measurement properties against existing criteria

The criteria used were based on those described by Locker and Allen and COSMIN criteria (Locker and Allen, 2007; Terwee et al., 2007; Mokkink et al., 2010a).

5.3 Methods

5.3.1 Search strategy

A systematic search strategy was used to identify eligible studies, using the Mesh terms “child” and “quality of life” in combination with the names or the commonly used acronyms of the three measures. Both MEDLINE (through PubMed) and Web of Science were used to search for articles published up to December 2012. Reference lists of included studies were also searched to identify additional studies.

5.3.2 Selection criteria

Titles and abstracts were read independently by two investigators (FG and ZM) to ascertain whether they met the inclusion criteria. Disagreements were resolved by discussion and, where doubt existed, the full paper was retrieved. A paper was judged to be suitable for inclusion if:

- it used either the CPQ, COHIP or C-OIDP (or versions of them)
- it included participants aged 16 years or younger
- the measures were completed by the participants, not proxies

- the full paper was available in English
- it reported primary data

5.3.3 Data collection

Description of measures and their use (Objective 1)

To fulfill objective one and describe the measures and their use, data were collected relating to:

- The aim of the measure
- The measure used
- Study type (for example; development, validation, cross-cultural adaptation, etc.)
- Population (i.e. clinical, school-based)
- Measurement properties (detailed below)
- Methods used to develop the measure (using the criteria proposed by Locker and Allen (2007))

Data were extrapolated by two teams of two investigators (FG/HR and ZM/CD) for all included studies and analysed by FG. A protocol (developed by FG), with description of the data required to be collected was produced. The data collection spreadsheet was piloted using 10 articles, following which descriptors were added to each of the categories to aid completion. A training exercise was then held by FG with all investigators to ensure consistency of data extraction. Where there was disagreement between investigators, this was resolved by discussion to reach a consensus.

Assessment of the methodological quality of the development and testing of measures (Objective 2)

The COSMIN checklist was used to evaluate the quality of studies that reported the development or evaluation of the original form of the CPQ, COHIP or C-OIDP in the original language (Mokkink et al., 2010a). This tool allows the methodological quality of studies to be assessed against criteria for each measurement property and has been used successfully in systematic reviews of outcome measures (Schellingerhout et al., 2011; Chow et al., 2013; Park et al., 2013; Pusic et al., 2013; Weldam et al., 2013; Haywood et al., 2014). The checklist contains 5-18 items per property which are rated excellent, good, fair or poor, with the lowest score for any item being assigned as the overall score for that property.

Two reviewers (FG and ZM) decided which properties had been assessed in each study and assigned an overall score. A calibration exercise was held prior to data collection to ensure consistency. Disagreements were resolved by discussion between investigators to reach a consensus. Both intra- and inter-examiner reliability were assessed and were found to be excellent (weighted Kappa = >0.9).

Quality assessment rating

The rating system proposed by Terwee and colleagues (2007) was used to assess the quality of the instruments using the results of the studies evaluated by the COSMIN checklist. This allows a positive, negative or indeterminate rating to be assigned depending on the published results.

Measurement properties analysis

Validity, reliability, responsiveness and interpretability of the measures were analysed using the following aspects as discussed in Section 4.5.4 (Terwee et al., 2007):

- Content validity: The degree to which the items in the measure are a reflection of those important to the study population and to the construct under scrutiny.
- Construct validity: this refers to the extent to which scores relate to other measures of a similar concept under scrutiny and should be tested using predefined hypotheses to avoid bias.
- Internal consistency: the extent to which items measure the same construct. In Classic Test Theory (CTT), this is expressed using Cronbach's alpha value. Factor analysis or item response theory should be used to confirm unidimensionality.

- Criterion validity: this relates to whether the scores on a particular measure have a positive correlation with a gold standard. There are no gold standards in the field of OHRQoL and therefore measurement of this is only appropriate when testing a short form against the existing measure.
- Test-retest reliability: the ability of the measure to produce reproducible results in a stable population over time. This should be analysed using the intraclass correlation coefficient (ICC).
- Responsiveness: the ability of a measure to detect clinically important changes over time, for example, after an intervention. The Minimal Important Difference (MID) should be stated and the change score should correlate with the global rating of change score.
- Floor or ceiling effects: these were considered to be present where more than 15% of patients score the highest or lowest score possible.
- Interpretability: the degree to which scores on the measure can be given qualitative meaning. For example, the provision of means and standard deviation of scores of relevant subgroups (clinical diagnoses, age groups, gender).

Best evidence synthesis

A best evidence synthesis was performed to summarise the evidence for each measure based on the methodological quality, consistency of results and the number of studies.

Two reviewers (FG and ZM) assessed the evidence for each measure and assigned a rating. A training exercise was held to ensure consistency. Disagreements were resolved by discussion between investigators to reach a consensus.

The results were defined as:

- Strong evidence: consistent findings in multiple studies of good methodological quality or one study of excellent quality
- Moderate evidence: consistent findings in multiple studies of fair methodological quality or one study of good quality

- Limited evidence: one study of fair methodological quality.

Where there were only studies with poor methodological quality or where statistical methods other than those recommended were used, a lack of evidence was noted.

5.4 Results

The search strategy yielded 653 papers. Four hundred and seventeen were duplicates leaving a total of 236 abstracts. Following analysis of the abstracts, 126 full papers which appeared to meet the inclusion criteria were retrieved. Of these, six were excluded as they did not meet the inclusion criteria therefore 120 papers were included in the final analysis (Figure 5.1). The majority used a version of the CPQ, most frequently the original version of CPQ_{11-14}(Figure 5.1). Most papers reported cross-sectional studies (n = 117) with three of longitudinal design (Figure 5.2). The number of publications using these measures steadily increased from 2008-2011 and reached a peak of 21 in 2011. A decline, perhaps related to delays in indexing of the databases, was seen in 2012 (**Figure 5.3**).

Fifteen studies which had used the original version of the measures in their original language were included in the COSMIN analysis. The following subsections will present findings relating to the evaluation of each questionnaire with the additional COSMIN analysis.

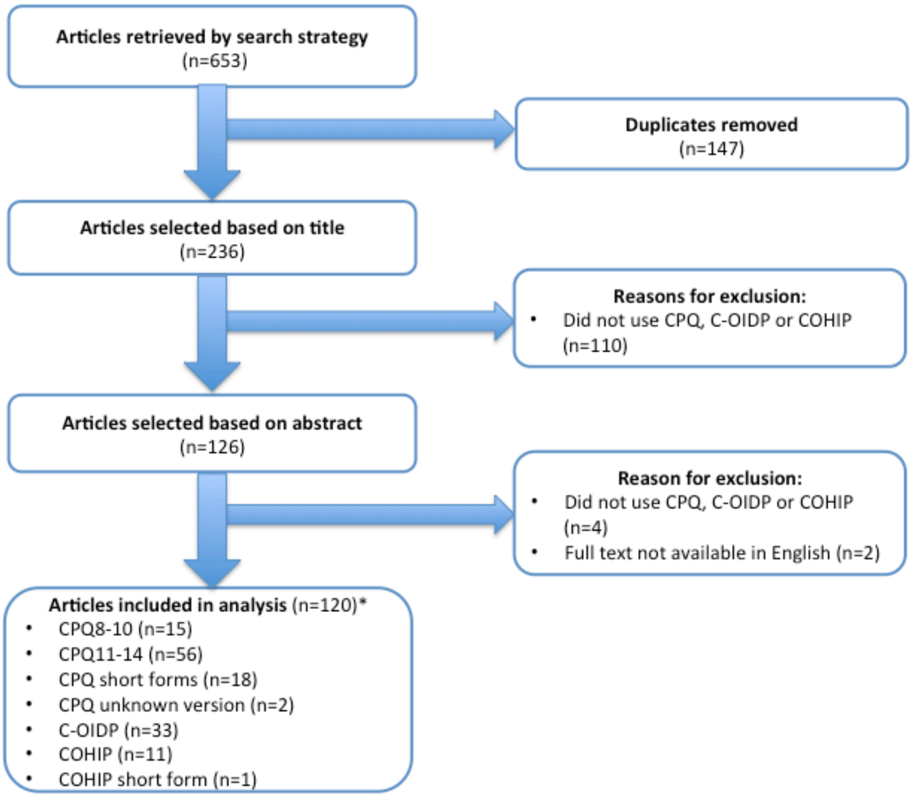

Figure 5.1. Flowchart detailing included articles in main study.

CPQ=Child Perceptions Questionnaire
C-OIDP=Child Oral Impacts on Daily Performances Index
COHIP=Child Oral Health Impact Profile
*Some papers used more than one measure.

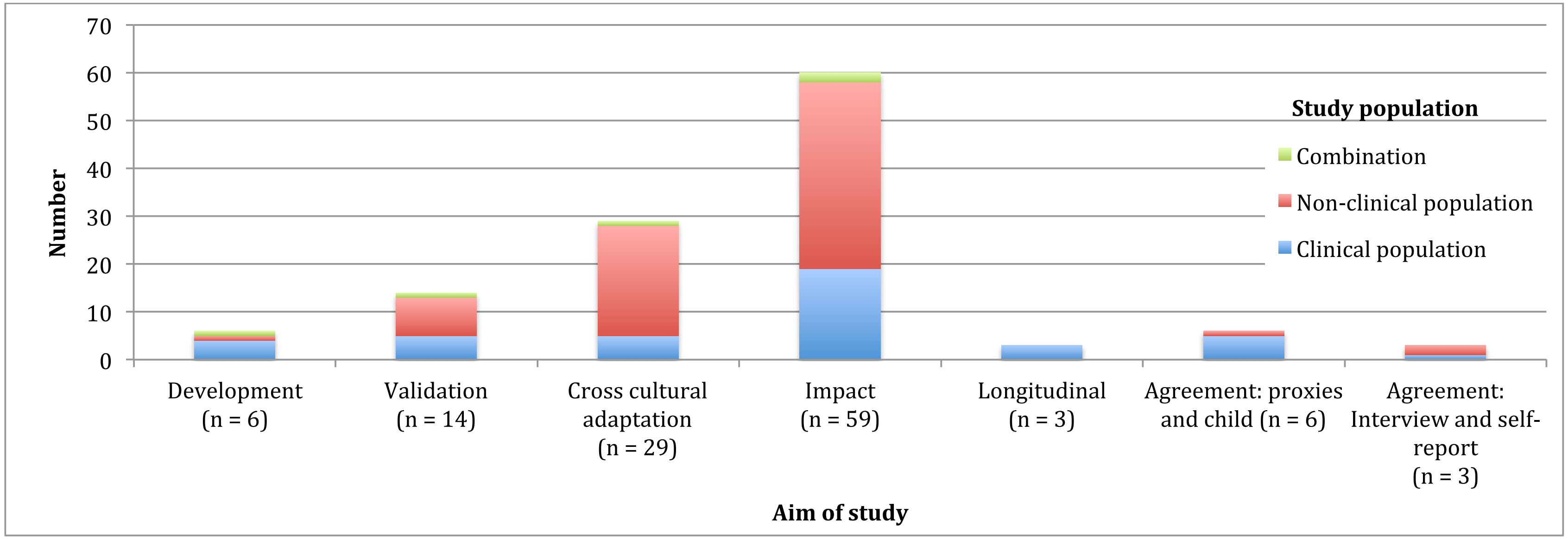

Figure 5.2. Aim of studies described by each paper and characteristics of study population (n = 120).

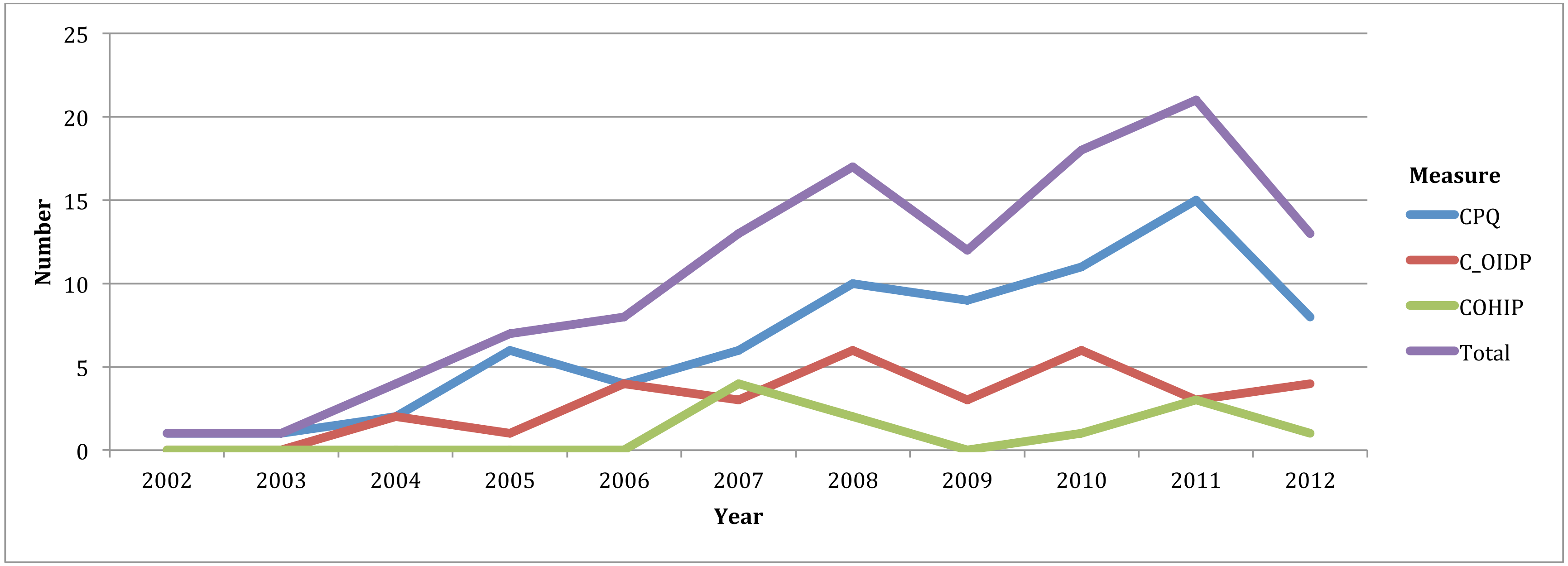

Figure 5.3. Publications per year in total and for each measure.

CPQ=Child Perceptions Questionnaire; C-OIDP=Child Oral Impacts on Daily Performances Index; COHIP=Child Oral Health Impact Profile

5.4.1 CPQ

This questionnaire was developed in Canada and was originally validated in children with caries, malocclusion and craniofacial anomalies (Section 4.6.1). A number of versions have been produced. The original item pool was developed following a review of existing oral health and paediatric measures. This was further reduced following discussion with healthcare professionals, parents of children and children with a variety of oral conditions. The included studies are shown in Appendix B.

CPQ$_{11\text{-}14}$[a]

Description of CPQ$_{11\text{-}14}$ and its applications

The aim of this questionnaire was to "produce a measure which conformed to contemporary concepts of child health and had discriminative and evaluative properties, and which is applicable to children with various dental, oral and oro-facial disorders". Although not explicitly stated, the measure must therefore have been designed to measure change at a group level due to its aims. The measure was validated by comparing scores between groups (caries, malocclusion, craniofacial) and by correlating overall scores with global ratings. Further details are shown in Table 5.2.

Study types/populations

Fifty-six papers used CPQ$_{11\text{-}14}$. Of these, one described development of the measure and seven its validation. Cross-cultural adaptation and validation of these versions were described in 12 studies from Hong Kong, Brazil, Denmark, Uganda, Saudi Arabia, Thailand and Germany. One paper investigated agreement between self- and interview-administered versions and one study reported on the changes in scores during orthodontic treatment. The remaining articles described OHRQoL in cross-sectional population studies and explored the impact of various dental and medical conditions.

---

[a] Studies included in this section are referenced in Appendix B with the suffix [a]

CPQ_{11-14} had been translated into Chinese, Brazilian-Portuguese, Danish, Luganda, Arabic, Thai and German. Further versions in Malay, Finnish, Norwegian and Russian were described but no details were provided regarding their validation.

Measurement properties

Eleven studies reported test-retest reliability with ICCs ranging from 0.6 to 0.94. The test-retest period varied from one week to one month and involved between 14 and 84 participants.

Internal consistency was investigated in 19 studies for CPQ_{11-14} with Cronbach's alpha ranging from 0.81 to 0.95. Factor analysis was used in one study using the Chinese version and showed that the emotional and social wellbeing subscales had a poor fit to the model. No studies had used IRT methods to investigate unidimensionality.

Criterion validity testing was not appropriate for this measure as there is no gold standard. Construct validity was measured in 17 studies using global ratings and clinical data. Positive correlations were found for total scores and the global ratings of oral health and life overall but conflicting results were reported for correlations with total scores and clinical data.

No studies reported face or content validity testing, except during the development and cross-cultural adaptation of the measures.

Specific details regarding floor and ceiling effects were reported in only seven studies, with up to 3 and 5% of participants scoring zero or the maximum scores respectively. Only seven studies discussed the number of omitted items, with five reporting the exact number of missing responses

Two studies reported longitudinal data with one reporting that the MID was a change in score of 4 points. The other study did not discuss what would be considered a clinically important change in score over time.

Mean and subgroup scores, where available, are shown in Appendix B.

Assessment of the methodological quality of the development and testing of CPQ$_{11\text{-}14}$

The CPQ$_{11\text{-}14}$ was studied in four papers in children with dental caries, enamel defects, malocclusion and craniofacial disorders. The original form has been validated in Canada and the UK.

Validity

Hypothesis testing for construct validity was performed in all four studies using correlations with clinical data and global ratings. The methodology was rated excellent in two cases (Marshman et al., 2005; O'Brien et al., 2007) and fair in the other two cases (Jokovic et al., 2002; O'Brien et al., 2006). The results for construct validity were rated positively in all studies. Content validity was considered in one study of fair methodology and rated positively (Jokovic et al., 2002). Criterion validity was not applicable for this measure as there is no gold standard.

Reliability

Internal consistency was analysed in all four studies and the methodology was rated uniformly poor, as the studies did not report testing of unidimensionality by factor analysis or item response theory. Therefore internal consistency was rated as indeterminate, however, it should be noted that all studies reported Cronbach's alpha of between 0.7 and 0.95. Test-rest reliability was performed in three studies, one of which was rated as good (Marshman et al., 2005), one fair (Jokovic et al., 2002) and one poor (O'Brien et al., 2007) and all had a positive ICCs.

Best evidence synthesis (Table 5.3)

Combining the results of the methodological quality with the published results produced strong evidence for construct validity and lack of floor or ceiling effects, limited evidence for interpretability, reliability and content validity and a lack of evidence for internal consistency.

Table 5.2. Characteristics of included measures.

Questionnaire	Age range designed (years)	Number of items	Number of domains	Range of possible scores	Scoring method	Completion method	Recall period
CPQ$_{8-10}$	8-10	25	4	0-100	5-point Likert scale (0-4)	Self	4 weeks
CPQ$_{11-14}$	11-14	37	4	0-148	5-point Likert scale (0-4)	Self	3 months
CPQ$_{11-14}$ short forms	11-14	16 or 8	4	16-item = 0-64 8-item = 0-32	5-point Likert scale (0-4)	Self	3 months
C-OIDP	?	8	1	0-72	4-point Likert scale (0-3)	Interview	3 months
COHIP	8-15	34	5	0-136	5-point Likert scale (0-4)	Self	3 months
COHIP short form	8-15	19	5	0-76	5-point Likert scale (0-4)	Self	3 months

CPQ=Child Perceptions Questionnaire; COHIP=Child Oral Health Impact Profile; C-OIDP=Child Oral Impacts on Daily Performances index.

Table 5.3. Best evidence synthesis per measure.

Questionnaire	Content validity	Internal consistency	Construct validity	Reliability
CPQ$_{11-14}$	+	?	+++	+
CPQ$_{8-10}$	?	+	++	+
CPQ$_{11-14}$ short forms	n/a	?	+	+
COHIP	+++	?	+++	+
C-OIDP	+	?	++	+

+++ or --- = strong evidence of positive or negative result respectively; ++ or – = moderate evidence of positive or negative result respectively; + or – = limited evidence of positive or negative result respectively; ± = conflicting results; ? = unknown due to poor methodological quality of study, n/a = no information available.

CPQ$_{8-10}$[2]

Description of CPQ$_{8-10}$ and its applications

The aim of this measure was not explicitly stated, but it was assumed to be the same as that for CPQ$_{11-14}$. Questions for this version were derived by a child psychologist, teachers and parents from the questions in the CPQ$_{11-14,}$ with no input from children. Testing of the measure took the same approach as described above for CPQ$_{11-14}$ (Table 5.2, page 87).

Study types/populations

Fifteen of the included studies had used CPQ$_{8-10}$. One reported its development, two its validation and three were cross-cultural adaptations in Brazil, Denmark and Mexico. One was a longitudinal investigation of children following atraumatic restorative technique and one study measured agreement between self- and interview-administered questionnaires. The remainder (n=7) described the impact of temporomandibular dysfunction, caries, fluorosis, neutropenia, malocclusion and OHRQoL of cancer survivors.

CPQ$_{8-10}$ is currently available in Brazilian-Portuguese, Danish and Spanish.

Measurement properties

Four studies investigated test-retest reliability with ICCs ranging from 0.67 to 0.96. Retest periods ranged from seven days to two weeks and involved between 33 and 162 participants.

Eight papers reported internal consistency with Cronbach's alpha values ranging between 0.82 and 0.95. One study used factor analysis and found that 18 of the 25 items fitted the mathematical model. No studies were identified that had used IRT methods to evaluate unidimensionality.

Construct validity was tested using correlations between global ratings, proxy measures and clinical data. All showed mainly positive correlations. Criterion

---

[2] Studies included in this section are referenced in Appendix B with the suffix [b]

validity assessment was not appropriate for this measure as there is no gold standard. No studies reported face or content validity testing.

Only one article discussed floor and ceiling effects and reported that none were found. Only two studies discussed missing data, with the number of missing responses reported in each.

One study reported longitudinal data, however, no details of the magnitude of change that would be considered clinically important were discussed.

Details regarding mean and subgroup scores are shown in Appendix B.

Assessment of the methodological quality of the development and testing of CPQ$_{8\text{-}10}$

Two studies reported findings from the CPQ$_{8\text{-}10}$ in children: one involved participants with craniofacial disorders in Canada and one included a school population in Northern Ireland.

Validity

Hypothesis testing for construct validity was performed in both studies and was found to be positive using global ratings (Jokovic et al., 2004; Humphris et al., 2005) and other measures of similar constructs (Humphris et al., 2005). The methodology of one study was rated good (Jokovic et al., 2004) and fair for the other (Humphris et al., 2005). Development of the content of the measure did not involve the target population (i.e. children) and therefore the methodology was rated poor and it was assessed as being negative for quality. Testing of criterion validity was not appropriate for this measure as there is no gold standard.

Reliability

Internal consistency was analysed in both studies, with factor analysis employed in one (Humphris et al., 2005) which was rated fair for methodology and therefore rated positively for the measurement property. The other (Jokovic et al., 2004) was rated poor as there was no analysis of unidimensionality using factor analysis or IRT. Both studies had acceptable

Cronbach's alpha. One study tested test-retest reliability which was rated fair methodologically and given a positive rating for reliability.

Best evidence synthesis (Table 5.3, page 88)
Combining the results of methodological quality with the published results demonstrated there was moderate evidence of positive construct validity, limited positive evidence to support internal consistency, reliability, interpretability and lack of floor/ceiling effects and no evidence to support adequate content validity.

CPQ short forms[3]

Description of CPQ short forms and their uses
Four short forms of the CPQ are available, two with 16 items and two with eight items, each derived from the questionnaires for 11-14-year-olds. Eight- and sixteen-item versions were produced using item impact data from the original study resulting in questionnaires containing two and four items per domain respectively. These are termed the "impact short forms" (ISF:8 or ISF:16). The other versions were developed using the original validation data and by selecting the two or four items contributing most to the coefficient of variation for each domain, and called the "regression short forms" (RSF:8 or RSF:16). All short forms are scored in the same way as the original version with a recall period of three months (Table 5.2, page 87).

Study types/populations
Eighteen studies were identified where the CPQ short forms had been used, including three which used non-standard abbreviated versions. These non-standard versions were not included in the following analysis. One described development, two validation and three were cross-cultural adaptions from Hong Kong, Brazil and Brunei. The remaining studies reported the impact of dental trauma, orthodontic treatment and socioeconomic disparities in OHRQoL.

---

[3] Studies included in this section are referenced in Appendix B with the suffix c

The short forms of CPQ_{11-14} have been translated into Chinese, Brazilian-Portuguese and Malay.

Measurement properties

Three studies investigated test-retest reliability with ICCs ranging from 0.5 to 0.98. All reported periods between tests were two weeks and involved either 34 or 86 participants (one study did not report the number of participants).

Internal consistency was reported in six studies with Cronbach's alpha values ranging from 0.5 (RSF:8) to 0.9 (ISF:16). One study employed confirmatory factor analysis on the Chinese versions of the short forms and found good fit to the model for all four subscales in the RSF:16, ISF:8 and RSF:8 versions. The social wellbeing subscale did not fit the mathematical model for the ISF:16 version.

One study investigated face and content validity of the ISF:16 in an orthodontic population in the UK. This enquiry found a number of the items to be irrelevant especially with regard to the domains of oral symptoms and functional limitations. The participants also felt there were a number of items of importance to them which had been omitted from the measure.

Criterion validity was examined against the full version and found to be positive. Construct validity was assessed using global ratings and clinical data. Positive correlations with total score and global ratings of oral health and life overall were consistently found, however, there were conflicting data for correlations between total scores and clinical status.

Ten studies reported floor effects, ranging from 1.6% for ISF:16 to 13.3% for RSF:8 where a proportion was stated. Two studies reported ceiling effects ranging from 0 to 5%. Three studies reported the number of missing responses.

Mean and subgroup scores are shown in Appendix B.

Assessment of the methodological quality of the development and testing of CPQ short forms

Two studies investigated all short forms of the CPQ$_{11-14,}$ one in a clinical population in Canada including children with caries, malocclusion and craniofacial disorders and the other in a school in New Zealand. The third study investigated face and content validity of the ISF:16 in children undergoing orthodontic treatment in the UK.

Validity

Hypothesis testing for construct validity was undertaken in two studies using clinical data and global ratings (Jokovic et al., 2006; Foster Page et al., 2008) both of which had fair methodology and were rated positively. Criterion validity was tested in both studies against the original measure and was found to be positive with a fair methodology in both studies. The investigation of content validity (Marshman et al., 2010) had excellent methodology and found that some items were irrelevant to the target population and therefore this aspect was rated negatively.

Reliability

Both studies (Jokovic et al., 2006; Foster Page et al., 2008) which investigated internal consistency were rated poor for methodology and were subsequently given an indeterminate rating for the measurement property as neither used factor analysis or IRT methods to confirm unidimensionality. Only one study (Jokovic et al., 2006) analysed test-retest reliability which was given a positive rating and graded as having fair methodology.

Best evidence synthesis (Table 5.3, page 88)

Combining the elements from the methodological quality rating and the published results, moderate evidence was found to support the lack of floor/ceiling effects, limited positive evidence for construct validity, reliability and interpretability. There was no evidence to support adequate internal consistency. There was strong evidence that content validity was inadequate in an orthodontic population for the ISF:16 version in the UK.

5.4.2 C-OIDP[4]

Description of C-OIDP and its applications

This measure was developed from the existing Thai version of the Oral Impacts on Daily Performances index (OIDP) (Section 4.6.2). It can be assumed that the aim of C-OIDP was to be a socio-dental health indicator (measuring the social effect of dental conditions) based on the theoretical model of oral health consequences, that underpinned the OIDP (Adulyanon et al., 1996). Validity was tested using correlations with two global questions (perceived oral health problems and perceived treatment need). Further details are shown in Table 5.2 (page 87).

Study type/populations

Thirty-three papers reported use of C-OIDP. One reported development and two its subsequent evaluation. Ten articles described cross cultural adaptation and subsequent validation in the United Kingdom, Malaysia, France, the Sudan, Tanzania, Spain, Italy, Brazil and Peru. Two studies investigated the level of agreement between self- and interview-administrations and one gave an account of changes in impact following treatment of caries. The remainder described the impact of various oral and medical conditions.

C-OIDP is available in English, French, Malay, Arabic, Kiswahili, Spanish, Portuguese and Italian. A further paper described its use in Hungary, however, no details were given regarding the validation of this version.

Measurement properties

---

[4] The studies included in this section are referenced in Appendix C

Eighteen studies reported test-retest reliability with ICCs ranging from 0.7 to 0.98. Test intervals ranged from the same day to three weeks and involved between 18 and 106 participants.

Five studies reported internal consistency for C-OIDP with Cronbach's alpha values ranging between 0.79 and 0.91. No studies were identified that employed factor analysis or IRT methods to evaluate unidimensionality.

Construct validity was tested using perception of treatment need, satisfaction with oral health, dental appearance and clinical data. Significant correlations were found with perceived need for treatment, oral hygiene and satisfaction with oral health. Testing of criterion validity was not appropriate for this measure. There were no studies which reported investigation of face or content validity.

There were no studies which reported the incidence of floor or ceiling effects. Only one study stated the completion rate, stating there were no missing responses.

One study was longitudinal in design, however, no data were available regarding changes in score which were considered clinically significant.

Appendix C provides mean and subgroup scores, where available.

Assessment of the methodological quality of the development and testing of C-OIDP

This questionnaire had been analysed in two studies involving Thai school children and these studies were evaluated using the COSMIN criteria.

Validity

Development of the content of the measure was rated as fair and given a positive rating (Gherunpong et al., 2004a). Hypothesis testing for construct validity was undertaken in two studies using clinical data (Bernabe et al., 2009) and perceived oral health need (Gherunpong et al., 2004a; Bernabe et al., 2009) with good methodology. One study found positive findings, while the other found positive findings using a condition-specific version. Criterion validity testing was not appropriate for this measure as there is no gold standard.

Reliability
Testing of internal consistency and test-retest reliability were tested in one study (Gherunpong et al., 2004a). Assessment of internal consistency was rated poor methodologically due to lack of testing for unidimensionality and therefore given an indeterminate rating, although Cronbach's alpha was adequate. Reliability testing was rated fair for methodological quality and given a positive rating for the published results.

Best evidence synthesis (Table 5.3, page 88)
Moderate positive evidence was available to support construct validity, there was limited evidence of positive reliability and interpretability and no evidence for internal consistency.

5.4.3 COHIP[5]

Description of COHIP and its applications
This instrument was designed for use in clinical situations to discriminate between children with different clinical conditions and with differing clinical severity (Section 4.6.3). It was intended for use in research and in clinical practice. Validity was tested using comparisons between groups (caries, malocclusion and craniofacial), and in those with differing levels of clinical severity. Correlation with other instruments and the two global rating questions (treatment expectations and effect on life overall) were also undertaken. Further details are available in Table 5.2 (page 87).

Study types/populations
Eleven articles included COHIP. One study reported development of the measure and two its validation. Four described cross-cultural validation in Korea, Iran and the Netherlands. The remaining studies described investigation of the impact of orthodontic treatment, correlations with self-

---

[5] The studies included in this section are referenced in Appendix D

reported dental aesthetics, the impact of cleft lip and palate and concordance between child and caregiver's scores.

COHIP has been translated into Dutch, Korean, Malay and Farsi. Finnish, Norwegian and Russian versions have also been reported but there was a lack of validation data for these translated measures.

Measurement properties

Two studies assessed test-retest reliability with ICCs ranging from 0.84 to 0.88, one using a two-week and the other a three-week interval between tests. The number of patients involved was not defined in either study.

Internal consistency was evaluated in seven studies with Cronbach's alpha ranging from 0.79–0.91. Factor analysis was used to evaluate the Brazilian version of COHIP and found that only the oral symptoms and the emotional wellbeing fitted the mathematical model. No studies used IRT methods to evaluate unidimensionality.

Construct validity was measured in four studies, using correlations with global ratings of general and oral health, clinical data and parental scores. Statistically significant correlations were found between global ratings, number of carious surfaces and degree of overjet.

As there is no gold standard, testing of criterion validity was not appropriate for this measure. No data were available for face or content validity outwith the initial development stage.

Two studies reported the proportion of floor (0-0.4%) and ceiling effects (0%). Three studies discussed missing data, with proportions of missing responses ranging from 0.4% to 55.7%.

There were no longitudinal studies and therefore there are no responsiveness data available for this measure.

Mean and subgroup scores are shown in Appendix D.

Assessment of the methodological quality of the development and testing of COHIP
Three studies investigated the COHIP in clinical and school populations in Canada and USA.

Validity
The methodology for development of the content of the questionnaire was rated as excellent and found to be positive (Broder et al., 2007). Hypothesis testing for construct validity was investigated in two studies (Broder and Wilson-Genderson, 2007; Dunlow et al., 2007). These were found to have excellent methodology and were rated positively.

Reliability
One study (Broder and Wilson-Genderson, 2007), investigated internal consistency, this had a good methodology but did not test unidimensionality in this population and was therefore rated indeterminate. The same study (Broder and Wilson-Genderson, 2007) investigated test-retest reliability and was rated fair with a positive result.

Best evidence synthesis (Table 5.3, page 88)
There was strong positive evidence of adequate content validity and construct validity and limited positive evidence of reliability and interpretability. There was no information provided regarding the presence of floor or ceiling effects. Although factor analysis had been performed during the development of the measure, to aid item reduction, further investigation of the unidimensionality of the scale had not been performed and therefore internal consistency was rated as indeterminate.

COHIP short form

Description of COHIP short form and its uses
Recently, a 19 item short form of the COHIP has been developed by using confirmatory factor analysis to remove items with weak loadings. This version

had not been tested independently at the time of this review (Broder et al., 2012).

Assessment of the methodological quality of the development and testing of COHIP short form

This measure has been evaluated only in the study used to develop the measure (Broder et al., 2012). Data from the original version were used to evaluate the measure.

Validity

Hypothesis testing for construct validity using clinical data, parental and global ratings revealed positive results with a fair methodology. Criterion validity was not assessed despite the fact that the original form would be considered the gold standard.

Reliability

Confirmatory factor analysis was used to determine the items for inclusion in the short form, therefore the measure was given a positive rating for internal consistency with a fair methodology. However, IRT was not used to confirm unidimensionality.

Best evidence synthesis.

This was not evaluated due to the limited evidence for this measure at the time of this review.

5.5 Discussion

This review evaluated the three most commonly used generic measures of OHRQoL for children against existing criteria and the findings will inform the development of the proposed new measure. The CPQ_{11-14} was found to be the most frequently employed measure, followed by the C-OIDP.

5.5.1 Development and validation

Although the aim of the measures seemed implicit from the outset, OHRQoL was not defined in any of the papers describing their development. As there is great debate about whether measures of this type can really capture aspects of quality of life, it is important to define exactly what it is that will be measured (Locker and Allen, 2007). Some authorities have suggested that measures such as these may be more appropriately termed "subjective health status measures" (Locker and Allen, 2007). However, as the measures addressed in this systematic review include items which cover aspects of daily life they may be considered to measure OHRQoL (Bullinger et al., 1993). The incorporation of global quality of life and oral health-related quality of life may allow patients to express their own feelings towards these concepts (Prutkin and Feinstein, 2002). Analysis of this information, together with the numerical scores for the measure, may provide a way to ascertain the meaning of the scores derived from these instruments (Locker and Allen, 2007).

In addition, further qualitative investigation may be required to ensure that measures cover the full range of issues which are important to children. Individuals with the relevant conditions should be involved in item generation (Broder et al., 2007). Although children were involved in the development of some of these measures, they did not fully participate in item generation and therefore impacts which are important to children may have been omitted. Indeed, Marshman and colleagues found that orthodontic patients felt that some of the questions in the CPQ$_{11-14}$-ISF:16 were irrelevant or difficult to understand (Marshman et al., 2010). Participants in that study also commented that a frequency based response format was less relevant than one which was based on severity. This was the only study to investigate these aspects outwith the development process and therefore it is not possible to generalise these findings. However, further investigation of face and content validity may be useful in other settings. It should be noted that other investigators working with children, have implemented severity based response formats following children's involvement (Christie et al., 1993; Otley et al., 2002; Stevens, 2009; Carlton, 2013a). As both CPQ and COHIP rely mainly on frequency scores, this may impair their ability to adequately reflect children's views. It is important that these aspects are considered during the development of the caries-specific measure to ensure that items are relevant and that the response format mirrors how children describe the impacts. This strategy has previously been employed in the development of a preference-based measure of health-related quality of life (Stevens, 2009).

It has been suggested that quality of life measures should include both "positive" and "negative" items to encompass all aspects which may impact upon wellbeing (Broder, 2007). Indeed, it has been suggested that the inclusion of positive items may aid identification of factors relating to coping or resilience which might otherwise be difficult to ascertain (Edwards et al., 2002). Of these measures reviewed in this chapter, only COHIP incorporates positive items. These statements include "I am happy with my teeth" and "Felt that you were attractive". Both items were suggested during focus groups with parents and further endorsed by children. While positive items may be useful in measures for some conditions, the benefit of positive items in a caries-specific measure is unclear, as the aim of treatment is to improve negative impacts such as pain, rather than to improve coping or self-esteem.

Four studies investigated change following an intervention and reported changed mean scores (Zhang et al., 2008; Foster Page et al., 2010; Mashoto et al., 2010; Paula et al., 2012a). Discussion of whether these changes were clinically meaningful was included in only one of the studies (Foster Page et al., 2010). The COSMIN group recommend that, in order to validate a measure's responsiveness, correlations between the change in score and a comparator instrument (e.g. patient's perception of whether their condition has changed) should be assessed. This was not performed in any of the studies, therefore it cannot be assessed whether the reported changes found were valid. (Abanto et al., 2013b; Amato et al., 2014; Brosens et al., 2014; de Paula et al., 2015; Turton et al., 2015)

To evaluate responsiveness in a clinically meaningful way it is essential to calculate the minimal important change or difference (MID). This can be done by comparing global ratings to assess when patients perceive change to have occurred and their overall questionnaire score. Thus the MID can be defined as "the smallest difference in score, that a person perceives as important" (Guyatt et al., 1987a). Foster-Page and colleagues (2010) reported the MID, but not correlations between the change scores and global scores. Since the original study was completed, further studies have reported on the responsiveness of CPQ_{11-14}, CPQ_{8-10}, CPQ_{11-14} 16-item short form, C-OIDP and SOHO-5 (Abanto et al., 2013b; Amato et al., 2014; Brosens et al., 2014; de Paula et al., 2015; Turton et al., 2015). Several of these studies reported the MID or effect sizes (Abanto et al., 2013b; Amato et al., 2014; de Paula et al., 2015; Turton et al., 2015). However, none of these studies investigated correlations between change scores and the global transitional judgement question. Therefore although CPQ_{11-14}, CPQ_{8-10}, CPQ_{11-14} 16-item short form and C-OIDP have been used longitudinally they have not yet been validated for use in this way according to the methodology suggested by the COSMIN criteria (Mokkink et al., 2010a).

Disease-specific measures have been found to be more adept at detecting these clinically important changes as the questions specifically address issues associated with one disease (Wiebe et al., 2003). As CPQ, C-OIDP and COHIP are generic, they may be unable to identify subtle changes following interventions. A caries-specific measure should be sensitive to these changes and would allow the effects of different treatment regimens to be compared from the child's perspective.

5.5.2 Types of study

In the main, the use of self-report OHRQoL measures for children has been restricted to validation, cross-cultural adaptation and the description of impacts in various dental and orofacial conditions. There is very little information available about the impacts which may change following interventions for dental caries and that which does exist may not accurately reflect changes as the measures have not been validated for this purpose. Thus, many of their potential applications, such as those described in Table 4.3 (page 60) have not yet been pursued. For example, the theoretical models which the measures are based on have rarely been evaluated. Exploration of this facet may improve our understanding of what these questionnaires really measure (Baker et al., 2010). In addition, their influence on policy has yet to be seen. It has been suggested that difficulty in interpretation, due to uncritical reporting of scores, has contributed to their lack of use by policy makers (Tsakos et al., 2013).

5.5.3 Properties

Methodological quality was assessed for 15 studies, most of which involved $\mathrm{CPQ}_{11\text{-}14}$. The majority of studies were rated as excellent, good or fair in relation to assessment of test-retest reliability, hypothesis testing for construct validity and content validity. However, lack of testing of internal consistency using factor analysis or item response theory (IRT) meant that most studies were rated poor for this property. Factor analysis and item response theory allow redundant items to be removed, thus shortening the questionnaire and removing duplication, thereby creating more sensitive instruments and reducing participant burden. It should be noted that such techniques have been employed in studies using versions of the measures which have been subject to cross-cultural validation. However, these methods have not been consistently applied to the original forms which were included in this analysis (Geels et al., 2008a; Lau et al., 2009; Baker et al., 2010; Traebert et al., 2010a; Wong et al., 2011).

No studies in the original review were identified that had used IRT methods to assess unidimensionality. However, since the review was undertaken, one study has assessed unidimensionality using an IRT method (Yau et al., 2015). This study found that ISF:8 and RSF:8 versions of CPQ appeared to have acceptable unidimensionality but that items in the oral symptoms domain were not discriminating (Yau et al., 2015). However, studies which have used factor analysis on versions with more items have found that the measures may not be unidimensional (Humphris et al., 2005; Geels et al., 2008a; Lau et al., 2009). Unidimensionality is an important property where measures are intended to assess change, as this allows conversion of the raw ordinal score into an interval scale. Use of IRT methods can also identify items which are redundant due to local dependence and those which function differently in different groups (Tennant and Conaghan, 2007). Yau and colleagues (2015) found that there were items in CPQ$_{11-14}$-ISF:8 and CPQ$_{11-14}$-RSF:8 which were subject to differential item functioning (DIF) with females with similar levels of impact answering these questions differently to males (Yau et al., 2015). This has an impact on the overall score obtained as it will not adequately reflect the level of impact as female scores will be artificially inflated. Rasch analysis (a form of IRT) will be used in the present study to ensure that the caries-specific measure is unidimensional, is free of DIF and that any redundant items are removed.

Best evidence synthesis shows strong positive evidence for construct validity for both CPQ$_{11-14}$ and COHIP, indicating that they measure appropriately according to the construct they intend to measure. However, in this part of the analysis, COHIP had only been evaluated in two studies, both during its initial validation, and thus limited to one population.

Positive evidence for test-retest reliability was found for all measures indicating that they are reliable in stable populations. Strong evidence of content validity was best for COHIP, due to the rigorous process implemented in its development. Although as previously discussed, this could have been improved by involvement of children in the initial item generation, rather than at the item impact stage.

The measures evaluated in this review were developed before the publication of standards such as the COSMIN checklist. Therefore, some elements, such as analysis using item response theory, were not included in many studies and adversely affected their overall ratings.

5.5.4 Interpretability of the studies

The COSMIN group recommend that, to allow assessment of interpretability and generalisability, studies should include demographic information (age, gender, location and population type), disease characteristics, the range and mean scores, the proportion of participants who had floor and ceiling effects and the methods used to select the sample (Mokkink et al., 2010a). Many studies lacked clinical information about the participants, for example severity of malocclusion or dmft/DMFT for the population. This makes it difficult to compare results between populations or to allow any understanding of what the range of scores might be in a given sample. In addition, little information was provided regarding the proportion of missing items and how these were dealt with. Missing data can cause bias and reduce the power of the study but may also indicate that participants either do not want to answer a question or do not understand it (Fayers et al., 1998). Finally, very little information was provided on the time it took participants to complete the questionnaires. This is an important aspect particularly for young children who might find it difficult to concentrate for long periods of time.

5.5.5 Limitations

The systematic review presented in this chapter assessed the three most commonly used self-report measures of OHRQoL for children. There are other measures which have been developed to measure OHRQoL in children, however these were not included as they have not been as widely used. Only articles published in English were included and therefore there may be studies which have used these measures which were not identified. The analysis of methodological quality was only undertaken for those studies which had used the measure in its original form. As two of the measures were developed in English (CPQ and COHIP), it is unlikely that any studies reporting the validation of these were omitted. However, studies reported in other languages may have evaluated unidimensionality using factor analysis or IRT methods and these were not identified. Further studies have been published since the original review was undertaken which have used IRT methods and measured responsiveness (Abanto et al., 2013b; Amato et al., 2014; de Paula et al., 2015; Turton et al., 2015; Yau et al., 2015). However, IRT was used only to assess the Chinese version of CPQ11-14-ISF:8 and CPQ11-14-RSF:8 and although MID and effect sizes have been calculated, the measures have not been validated for longitudinal use using the methodology recommended by the COSMIN group (Mokkink et al., 2010a).

5.5.6 Novel aspects

This is the first study to report a comprehensive review of child self-report measures of OHRQoL. The COSMIN checklist was used to objectively review the measurement properties of each of the measures. Although this checklist has been used in other systematic reviews of health measures, this is the first time it has been used to assess OHRQoL measures (Schellingerhout et al., 2011; Chow et al., 2013; Park et al., 2013; Pusic et al., 2013; Weldam et al., 2013; Haywood et al., 2014). This study demonstrated the value and feasibility of using the COSMIN checklist and quality criteria. There was good intra- and inter-rater reliability (weighted Kappa=>0.9), which was higher than has been reported previously (Mokkink et al., 2010b). This may be attributed to the training and calibration sessions which took place prior to employing the criteria.

5.5.7 Summary

This review has provided valuable information which will inform the development of the proposed caries-specific measure. There is a lack of evidence that the existing measures are responsive longitudinally which strengthens the rationale for the development of a disease-specific measure. It will be important to consult children prior to testing the measure to ensure that the items are relevant and that the response format is based upon their description of the impacts of the disease. The COSMIN checklist will be used as a framework to evaluate the measurement properties of the questionnaire to ensure that sample sizes and statistical methods are appropriate. Rasch analysis will be used to supplement Classic Test Theory to ensure that the measure is unidimensional and that redundant items are removed to minimise participant burden.

5.6 Publications arising from the work described in this chapter

5.6.1 Peer-reviewed journal article

- Gilchrist F, Rodd HD, Deery C, Marshman Z. Assessment of the quality of measures of child oral health-related quality of life. *BMC Oral Health*. 2014; 23;14:40.

5.6.2 Published abstract

- Gilchrist F, Rodd H, Deery C, Marshman Z. A systematic review of measures of oral health-related quality of life for children. *International Journal of Paediatric Dentistry* 2012; 23 (s1): 1-256. Presented at the International Association of Paediatric Dentistry Conference, Seoul 2013. (Winner BSPD Max Horsnell Prize).

Chapter Six

Development of the measure

6.1 Introduction

The previous chapter discussed the strengths and limitations of existing measures of OHRQoL in children. While existing measures are generic, there remains a need for a caries-specific measure of OHRQoL which is responsive to changes in the participant's clinical state and thus will aid the evaluation of different approaches to the prevention and management of dental caries.

This chapter will discuss the development of a caries-specific measure of OHRQoL. The generation of items for the measure, reduction of these to ensure the most relevant are included, the formulation of the response format and the face and content validity of the measure will be outlined in the following text.

The measure will be based on the definition of OHRQoL as proposed by Locker and Allen (2007) and will therefore encompass "The impact of oral diseases and disorders on aspects of everyday life that a patient or person values, that are of sufficient magnitude, in terms of frequency, severity or duration to affect their experience and perception of their life overall". To ensure the resulting measure is relevant and meaningful to children with dental caries, it will follow the guidance offered by Guyatt and co-workers (1996) (Section 4.5.3) and will also meet the criteria proposed by Locker and Allen (2007) (Section 5.1) (Guyatt et al., 1986; Locker and Allen, 2007). Guyatt and colleagues (1986), recommend a seven-stage process (Figure 6.1). The first stage involves selection of the items following interviews with participants with the disease to elicit what is important to them. Following the generation of the items, it may be desirable to reduce these to a manageable number, ideally by asking a different group of participants to rate the importance and frequency of occurrence of each item. Those which occur most frequently and are rated most important should be retained. Once the final item selection has been made, the design and response format can be chosen before testing the measure for face and content validity. When face and content validity have been verified, the measure can then be tested for construct validity, reliability (internal consistency and test-retest) and responsiveness.

The first four stages (the remaining stages will be described in Chapter Seven) will be used to develop the measure, using child-centred methodologies to ensure that the broad concepts proposed by Locker and Allen (2007) are met. These are as follows:

1. Measures should be patient-centred
2. They must include aspects of daily life which are important to the patients involved and which may be affected by the disorder under investigation.

6.1.1 Stage 1: Item selection

The recommendation of Guyatt and co-workers (1986) is that a sample of 50-100 participants with the disease should partake in a semi-structured interview to elicit the items which are important to them. As discussed in Chapter Three and Chapter Five, the development of existing measures of OHRQoL involved parents and health professionals generating the items, which children were then invited to comment upon. This approach results in a measure using adult terminology, which children may find difficult to comprehend.

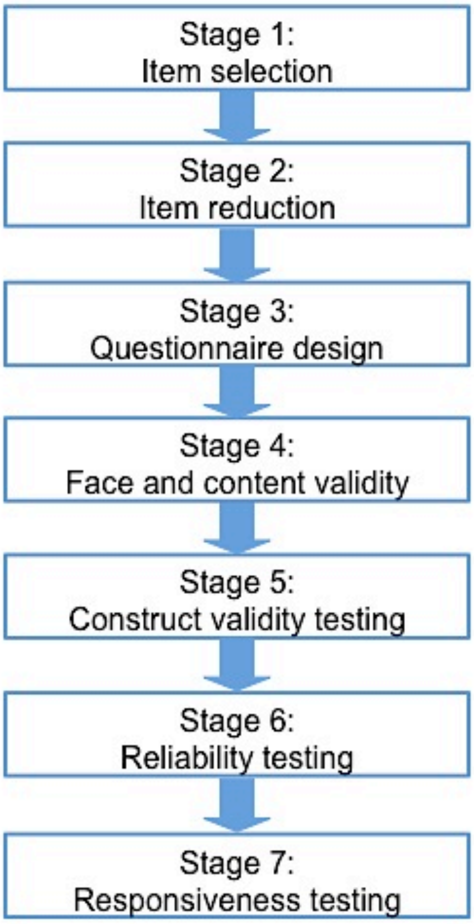

Figure 6.1. Questionnaire design stages adapted from Guyatt and colleagues (1986).

Increasingly, there is a recognition that children have their own unique view and perspectives and that they should have the right to express these (United Nations, 1989). Contemporary sociological research, as discussed in Section 4.2.1, has used methodologies which draw out children's own perspectives, rather than relying on the views of others (Christensen and James, 2000). This shift has resulted from a change in viewing children as passive objects to a realisation that children are experts on their own lives. The COSMIN group recommend that patients with the disease under scrutiny should be involved in judging the relevance of the items (Mokkink et al., 2010a). The US Food and Drug Administration (FDA) also recommends that patients and lay persons should be involved in developing measures designed for use in clinical trials (U.S. Department of Health and Human Services et al., 2009). This involvement of the target population is particularly important where children are concerned to ensure that the correct vocabulary is used and that those impacts which are important to children themselves are included. It has been suggested that the content validity and relevance is increased where those with the disease under scrutiny are involved in item generation (McColl, 2005). Indeed, techniques such as focus groups and in-depth or semi-structured interviews with children have been used successfully to elicit this information in the development of HRQoL and health utility measures (Griffiths et al., 1999; Young et al., 2004; Detmar et al., 2006; Morris et al., 2007; Stevens, 2010; Carlton, 2013b). Riley and co-workers (2004) reported that children as young as six years of age are able to self-report on their own health provided the questionnaire used is age-appropriate, therefore there is justification for developing a measure specifically for this young population.

The use of focus groups in social research studies has increased over the past decades and can be used to gather information in a naturalistic environment, as they stimulate more conversation between participants, than in-depth interviews (Krueger and Casey, 2000; Ritchie et al., 2003). The focus group allows participants to interact with each other therefore generating further information. Where children are involved it is advisable to have small groups to allow everyone to participate and to have children of similar ages in each group (Hoppe et al., 1995).

In contrast, interviews allow participants to talk freely and in privacy. This can result in richer information being produced and allows younger, quieter children to have their say. Interviews should ideally be performed outside of the clinical environment and where it is most comfortable for the participant, which is often their own home. A combination of these techniques will be used to develop the items for the measure in this study.

6.1.2 Stage 2: Item reduction

Guyatt and co-workers (1986) recommend using a questionnaire to identify item frequency and importance. Juniper and colleagues (1997) adopted this method to produce a shortened version of the Asthma Quality of Life Questionnaire and found that this approach included more items of importance to people with asthma. The $CPQ_{11\text{-}14}$ was also developed using this approach, whereby an item impact score is calculated by multiplying the proportion of children with the impact by the item's mean importance (Jokovic et al., 2002). The potential limitation of this method is that impacts which have been experienced by few participants but have a high "bother" rating may be omitted (Guyatt et al., 1986). Therefore a modified version of this method will be used in this study to ensure that items which are deemed to be very important are not excluded from the final measure. Other methods which can be used are factor analysis and Rasch analysis, where items which do not fit the mathematical model may be removed (Guyatt et al., 1986).

6.1.3 Stage 3: Questionnaire design

There is little in the literature describing how best to approach questionnaire design with young participants. The specific design and layout of the questionnaire depends on the questions involved and the reading ability of the target population. There are several factors which should be considered when developing scales for use in children (Matza et al., 2004).

Cognitive development

There are several complex cognitive stages involved in answering questions (Tourangeau and Rasinski, 1988; Scwartz and Sudman, 1996):

1. Comprehending the question
2. Retrieving relevant information from memory and formulating an answer
3. Choosing the appropriate response category
4. Evaluation of the chosen answer and editing for social desirability
5. Communicating the final answer

Children at different stages of cognitive development, as described by Piaget and Innhelder (1969) (see Section 4.2.1), will have differing abilities when formulating their answers. Children in the "intuitive thought" stage which occurs around 4-7 years of age, generally have limited language skills. Therefore questions should be simple and clear and it is advised that the words that children use should be employed to aid understanding (Borgers et al., 2000). Children at this stage also have short attention spans, which has consequences for the reliability of their answers if they are not engaged by the measure (Borgers et al., 2000). Around the age of 8 -11 years, children are entering the "concrete operational stage", their language is still developing and therefore simple, unambiguous language should be used. Finally, children in the "formal thought" stage (11-16 years) can start to use measures which are similar to those used in adult surveys, however, they will still need to be engaged and therefore the measure should be interesting to them (Borgers et al., 2000). Therefore, investigating face validity is essential to ensure that children can understand the language used and their input into the design may help to make the format interesting to prevent boredom.

Likert scales

It has been suggested that eight-year-olds can reliably use 5 to 7 point Likert scales to rate their health status, whereas younger children tend to choose extremes (Matza et al., 2004; Riley et al., 2004). Indeed, young children expressed a preference for three-point response scales when consulted during the development of the KIDSCREEN HRQoL questionnaire (Detmar et al., 2006). The Child Health Illness Profile and the Child Asthma Questionnaire-B use graduated circle sizes and bricks for their response options. This format was suggested by children as being something they were accustomed to doing at school to represent different quantities, which may aid questionnaire completion (Christie et al., 1993; Riley et al., 2004).

Recall periods

Younger children may find it difficult to accurately recall information over longer periods of time. Eight-year-old children have been shown to be able to accurately assess events over a four-week period. However, some young children who participated in the development of KIDSCREEN preferred a shorter recall period, as they felt that four weeks was too hard for them to remember. Whereas, adolescents in the same study preferred recall periods to be stated as "last week", "generally" or "recently" (Detmar et al., 2006).

Others have suggested that it may be beneficial to anchor the time frame to a specific event to aid recall in younger children (Juniper et al., 1997; Matza et al., 2004).

Length and formatting of the instrument

Generally older children will be able to maintain attention for longer periods. Larger print and spacing can aid completion by younger children. Piloting the questionnaire with the target age group will help to ascertain any changes which need to be made to simplify the language or format (Matza et al., 2004). It may be desirable to have multiple versions of the questionnaire for different age groups. This may allow consideration of aspects of the construct which may differ between children of varying ages. The Childhood Asthma Questionnaire has versions for 4-7-year-olds, 8-11-year-olds and 12-16-year-olds (Christie et al., 1993; French et al., 1994). However, this makes longitudinal evaluation difficult as the same questionnaire might not be appropriate at the end of the investigation. Recently, the short-form version of CPQ for 11-14-year-olds was found to perform appropriately in 5-8-year-olds (Foster Page et al., 2013a). Therefore it may not be necessary to have different versions for different ages, depending on the length and structure of the measure.

Existing OHRQoL measures for children use frequency based response options, with little explanation offered as to why these were chosen. Marshman and co-workers found that young people with a malocclusion felt that items in the CPQ$_{11-14}$-ISF:16 should be based on severity rather than frequency. They also felt that the use of response options such as "sometimes" were difficult to evaluate (Marshman et al., 2010). Stevens (2009) also found that children discussed the majority of dimensions in terms of severity and therefore the response scale for the Child Health Utility 9D (CHU9D) was based on this (Stevens, 2009). Likewise, the Child Amblyopia Treatment questionnaire (CAT-QoL) is also based on severity, reflecting how the children described the dimensions (Carlton, 2013a).

Stevens (2009) reported a method used to rank response options during the development of the CHU9D (Stevens, 2009). This involved giving participants an envelope with the potential options and asking children to rank them, those which children found difficult to rank accurately were removed. Where two options were ranked equally, children were asked to express a preference. This method ensures that children understand the response levels, and therefore more accurate scores may be obtained on the final questionnaire.

6.1.4 Stage 4: Face and content validity

Face validity assesses whether the questions make sense to the participants. Analysis of this, with subsequent necessary modifications, increases the likelihood of questions being completed accurately. Where the wording of the question is ambiguous, participants may omit the question or guess at an answer. Content validity examines the extent to which the construct is comprehensively covered by the measure (Section 5.1). This allows the identification of items which may be important but which have not yet been elucidated to be included, and for items deemed irrelevant to be deleted.

These elements can be tested using cognitive debriefing, where the participants are asked to explain why they have chosen the response (think-aloud technique) and what they think the question means (verbal probing) (Brod et al., 2009). Follow-up questions can be used to determine language comprehension, whether the response format makes sense and whether there are any items of importance which have been omitted or any irrelevant ones which have been retained. This form of questioning may allow insights into aspects which might not have been spontaneously volunteered by the participant (Beatty and Willis, 2007). Verbal probing may be more suited to assessing language comprehension, whilst think-aloud may be more useful to address why the response option has been chosen (Brod et al., 2009). This is an iterative technique which continues until consensus is reached that the measure is acceptable (Turner et al., 2007). Decisions to change the measure are usually taken after two or three participants have raised the same issue (Brod et al., 2009). A sample size of seven to ten participants is usually sufficient for consensus to be reached (Leidy and Vernon, 2008). This technique has been used successfully in the development and cross-cultural validation of HRQoL measures for use with children (Price et al., 2009; Carlton, 2013a; Young et al., 2013).

6.2 Aim

The aim of this part of the study is to produce a caries-specific measure of OHRQoL for children aged 5–16 years of age.

To fulfill this aim, the specific objectives were:

1. To describe the impacts of dental caries from the child's perspective
2. To evaluate how children describe these impacts to aid formulation of the questions and response format
3. To determine which of these impacts concern children the most
4. To evaluate which design and response format children prefer
5. To assess face and content validity of the items

6.3 Method

There are several stages to this part of the project and the method and results for each will be discussed in sequence. The following inclusion and exclusion criteria apply to each stage, with the exception of questionnaire design and face validity, where children with and without dental caries will be included:

Inclusion criteria

- Children and young people aged 5-16 years of age at recruitment
- Children who have active dental caries or who have had previous experience of dental caries
- Children who were patients at the Charles Clifford Dental Hospital or attended a primary care dentist
- English speaking

Exclusion criteria

- Children with other pre-existing medical conditions (American Society of Anesthesiologists' (ASA) Physical Status Classification System Grade 3 or greater (American Association of Anesthesiologists, 1974)

- Children with dental conditions other than dental caries (e.g. hypodontia, cleft lip and palate)
- Children with severe learning difficulties who would be unable to participate with the intended activities even with support

6.3.1 Ethics

This part of the study was approved by South Yorkshire Research Ethics Committee (Reference number: 11/H1310/3). Consideration was given to ethical issues which may arise during these stages such as power imbalance, confidentiality and safety as discussed in Section 4.3.

6.3.2 Stage 1: Item generation

Recruitment

Potential participants were identified from clinic lists and patient case-notes by FG. Children were purposively sampled to ensure that a representative sample was included. Children were recruited from both primary (Salaried Dental Services, Firth Park Clinic, Sheffield) and secondary care (Paediatric Dental Department at the Charles Clifford Dental Hospital, Sheffield) by FG. Recruitment continued until saturation was achieved (Sandelowski, 1995). Guyatt and co-workers estimated that around 50 participants may be required, however, previous studies have indicated that the inclusion of approximately 20 participants may be adequate (Guyatt et al., 1986; Marshman et al., 2009; Marshman et al., 2010).

For recruitment to the focus groups, children who were scheduled to attend for treatment on the day of the proposed focus group were sent an invitation by post with an enclosed reply slip and reply-paid envelope. They were informed that lunch would be provided and that the children would receive a £5 gift voucher. Once participants were identified, focus groups were split into age-appropriate groups.

For recruitment to the interviews, participants were purposively sampled from both primary and secondary care clinics to ensure that children with different levels of dental caries, and treatment modalities, were represented. Eligible children and their parents were asked if they would consider participating in the interviews by FG. If they expressed a desire to find out more they were given an information pack with a reply slip together with a stamped, addressed envelope.

Method

The sociology of childhood described in Section 4.2.1 guided the conduct and analysis of the interviews and focus groups. In order to achieve this and to prioritise the child's voice, the analysis took a narrative approach, so that rather than trying to "verify" what children said, the focus was on how they described their experiences and what they meant to them (Holland et al., 2008). Interviews and focus groups were held with children aged 5-16 years who had dental caries or experience of dental caries. Clinical data were also collected including their dmft/DMFT, presence of anterior caries, pulpal involvement or pain and where appropriate treatment experience by FG. DMFT/dmft were assessed by FG using data from case notes in combination with radiographic evidence where available. Caries was judged to be present where there was caries into dentine clinically or radiographically. A tooth was judged to have pulpal involvement where clinical and/or radiographic assessment revealed signs of infection (sinus, swelling etc.) and/or where it was judged restoration of the tooth would require root canal therapy. This method was used as there was no risk of underestimation of caries which would potentially require treatment. Potential participants and their parents were given age-appropriate information sheets detailing the project and containing contact details and what to do if they had any concerns. These information sheets had been developed with children to ensure that they used age-appropriate language. An example is shown in Appendix E. Written informed consent was sought from parents and children were asked to give their assent (Appendix F).

To address the ethical issues described in Section 4.3, all children were informed that they could stop taking part at any time and an explanation was given about why the interviews were being recorded. All children were asked to provide a pseudonym to maintain confidentiality.

Focus groups

The discussions centred around the participants' experiences of dental caries and its treatment and were facilitated by FG and ZM. This activity informed development of the topic guide for the in-depth interviews. The focus groups took place in June 2011 within the School of Clinical Dentistry, University of Sheffield and it was anticipated that these would last approximately 45-60 minutes. A separate area with refreshments was available for accompanying parents to wait during the session. The session was audio-taped (Olympus Digital Voice Recorder WS-812) and sound recordings were transcribed verbatim and analysed concurrently.

Interviews

The aim of the interviews was to elicit participants' experiences of dental caries and its treatment using a topic guide developed from the information obtained during the focus groups. Each interview was conducted by FG and took place at a time and venue chosen by the participant. The interviews were recorded and transcribed verbatim and analysed using the method described below. Recruitment continued until saturation was reached. Interviews were held between November 2012 and July 2013.

To ensure compliance with the ethical considerations discussed in Section 4.3 the following steps were taken. Participants were informed that they could stop the interview at any time without consequence. In addition, FG continually monitored the participant's body language and responses to ensure that that the participant remained comfortable with the conversation. Confidentiality was discussed with participants. Any disclosure which raised concerns would be handled according to normal safeguarding protocols.

It was anticipated that the majority of interviews would take place in the participant's home. To ensure safety of the interviewer, FG informed ZM, HR or CD of the time of the interview, anticipated finish time and provided location details. A working mobile phone was carried at all times and FG informed ZM, HR or CD when the interview was complete.

Qualitative data analysis

Analysis of qualitative data should result in a detailed description that identifies patterns and develops explanations, while remaining faithful to the data in its original form, thus allowing the participants' narratives to be explored as described previously (Sandelowski, 1995; Ritchie et al., 2003; Holland et al., 2008). In keeping with the new sociology of childhood, the analysis took the participants' accounts at face value without imposing any constructs on the views as expressed by the children. The most appropriate method to meet the aim of this study was framework analysis, which is concerned with classifying data by organisation according to themes and categories that emerge from the data. It has developed from social policy research to facilitate handling large volumes of data (Ritchie and Spencer, 1994).

The analytical approach involved the following stages (Ritchie et al., 2003):

1. *Identifying initial themes*

The two transcripts of the focus groups were read and notes made independently by FG and ZM on the general themes. ZM had previous experience of qualitative data analysis and FG had attended a three day course on this topic (National Centre for Social Research, London). These general themes were further refined to develop themes for the topic guide for the interview stage. HR and CD also read these transcripts to ensure that no potential themes for the topic guide had been omitted. Next the two focus group transcripts were read and notes made along with the first three interview transcripts by FG and ZM independently. This allowed further themes to be generated which were then added to the topic guide for subsequent interviews. This process continued, and as new themes emerged these were added to the topic guide to be included in future discussions.

Finally, the 18 transcripts (relating to two focus groups and 15 interviews) were read and notes made by FG and ZM independently. These notes were discussed and disagreements resolved. Following this, HR and CD analysed the transcripts to ensure no themes had been omitted. Once developed, the themes were grouped into a number of main and sub-themes. Initially, a number of themes were identified which were further refined following discussion between the four investigators (FG, ZM, HR, CD).

2. *Labelling the data*

Each section of the transcripts was labelled with an index number to represent the theme to which the data related by FG. These were then refined to add any themes previously missed after discussion with ZM, HR and CD.

3. *Sorting the data by theme*

NVivo (QSR International, Melbourne, Australia) was used to code the themes which emerged from the data. Data with the same index number was brought together by FG and verified by the other investigators (ZM, HR, CD).

4. *Synthesising the data*

Thematic charts were created for each of the main themes retaining the context and language used in the data. These were discussed by all four investigators (FG, ZM, HR, CD). This information was used to generate potential items for the measure and the potential response format options, retaining as much of the children’s language as possible.

6.3.3 Stage 2: Item reduction

Recruitment

Consecutive children attending as new patients with dental caries or who were having treatment for caries at the Charles Clifford Dental Hospital were identified and recruited to this stage by FG. Children and their parents were given age-appropriate information leaflets. Parents gave informed consent and all children were asked to give their assent. Clinical data were not collected for this group of children, however, they were considered to form a representative group of children with dental caries attending the clinic during the recruitment period. Children were recruited to this stage between January 2014 and February 2014.

Method

Children were shown cards by FG, with the identified items on them and asked to indicate which they had experienced and which had affected them most frequently and had the greatest severity (Figure 6.2). Their answers were recorded on a data collection sheet (Appendix G) and the items with the highest frequency and the greatest severity were retained. Previous studies have used samples of over 75 participants, where there have been a large number of items generated and there was no use of factor analysis or Rasch analysis (Griffiths et al., 1999; Jokovic et al., 2002; Broder et al., 2007). As items were to be deleted from the questionnaire following Rasch analysis, item reduction in this study was more limited and with a smaller sample, allowing discussion with participants about the items. Following the completion of this stage, draft questionnaires were developed to form the basis of the discussions in Stage 3.

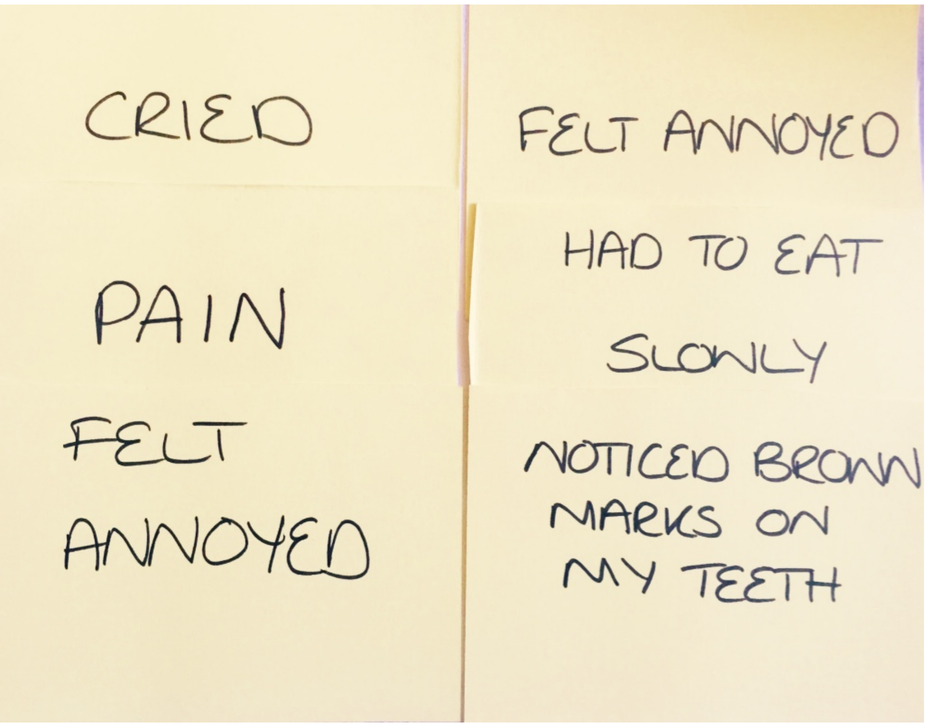

Figure 6.2. Example of cards used in item reduction stage.

6.3.4 Stage 3: Questionnaire design

Recruitment

These children were recruited by FG through personal contacts and had either been patients at the Charles Clifford Dental Hospital or attended a general dental practitioner for routine care. This stage took place in January 2014.

Method

Children who did not participate in the interviews or focus groups were asked by FG to comment on the design of the questionnaire and the format of the response scale. They were also asked to rank the words used in the response options from best to worst, to ensure that the responses made sense to children. This stage was designed to ensure that the resulting measure was easy for children of all ages to complete and that participant burden was minimised. These children may or may not have had dental caries and, as such, no clinical data were collected.

Six potential questions were formulated using two different methods. The first version (V1) contained a lead-in with three statements as response options. This type of response format has been used in the CHU9D and the CAT-QoL successfully (Stevens, 2009; Carlton, 2013a). Participants select their preferred option using a tick in a box (Appendix H), while the second version (V2) had a statement with three response options which could be circled to indicate their answer (Appendix I). A similar response format is used in the Oxford Foot and Ankle Questionnaire for Children measure, with tick boxes rather than circling (Morris et al., 2008). However it was felt that a circling option was something which could be considered. Children were asked to complete both questionnaires and state which format they preferred. They were also shown other designs and response formats from other questionnaires (such as those discussed in Section 6.1.3) to see if they thought they were better than the options they had been presented with.

Children were also asked to rank the potential response options from worst to best. Finally children were asked whether they felt coloured or white paper was best for the measure. These sessions were audiotaped (Olympus Digital Voice Recorder WS-812) and transcribed verbatim.

6.3.5 Stage 4: Face and content validity

Recruitment

Children attending the Charles Clifford Dental Hospital or Firth Park Salaried Dental Service Clinic were recruited by FG to this stage of the investigation. Children with and without dental caries participated in the analysis of face validity. All children who participated in the analysis of content validity had active caries or previous caries experience. Analysis of face and content validity took place during the month of March 2014.

Method

It is important to ensure that potential participants understand the items in the final measure and that no important items have been omitted. Children were asked to complete the questionnaire and discuss why they had chosen their answer, so that it could be assessed whether they actually understood the question. This approach has previously been used with both adults and children to explore face and content validity of questionnaires (Mallinson, 2002; Marshman et al., 2010). Changes to the wording of the questions was carried out where necessary to clarify the content. Additional items were added where it became clear that an item of importance had been omitted. Clinical and some demographic information for those children who participated in this part of the analysis was not collected, as some of those who participated in the analysis of face validity did not have caries. Stages 5, 6 and 7 will be discussed in Chapter Seven.

6.4 Results

6.4.1 Stage 1: Item generation

A total of 75 patients were approached to take part in either an interview or focus group (Figure 6.3). Of these, six children were invited to take part in the focus groups. Five children and their parents returned their slip indicating that they would like to participate in the focus groups (response rate 83%).

Forty-one families were given packs with reply slips for the interviews. Only five were returned. It was therefore decided that where a family had expressed an interest, a follow-up telephone call would be made with their permission. This resulted in the recruitment of a further 10 children. Ten declined to take part, six were unable to be contacted, one failed to attend the interview appointment and one cancelled. The overall response rate was 27%. However, in the group where a follow-up telephone call was arranged, the response rate was 36%.

Two focus groups (Focus group 1: 2 children both aged 12 years; Focus Group 2: 3 children aged 8, 8 and 9 years) were convened to ensure the ages of the children in each were similar. The accompanying parents (n=3) of the children in the younger group were included as this aided the discussion of events with the children. The focus groups began by asking the children to think of a pseudonym and to discuss why they liked this name to act as an ice-breaker. Props to allow the children to draw or write were available, however, these were not used as the children were happy to talk and share their experiences immediately. The focus groups were between 12 and 14 minutes in length. The information provided by these children formed the basis of the topic guide (Appendix J) for the interviews. These were supplemented with themes which were identified from other OHRQoL measures to ensure all aspects of OHRQoL were explored during the interviews. The transcripts of the focus groups were analysed along with those from the interviews.

Interviews were held with 15 children by FG following which saturation was reached. Fourteen interviews were held at the child's home, with the majority wishing to have their parent present during the interview. The participants ranged in age from 5-13 years of age and, despite the difficulties in recruiting participants, covered a wide range of ages and clinical presentations. The characteristics of the sample are shown in Appendix K. The interviews varied in length from 6 minutes to 16 minutes, with a mean duration of 6.58 minutes.

Data saturation was reached after two focus groups and 15 interviews, with no further impacts being discussed.

The transcripts were initially analysed to gain a broad understanding of the areas covered. The children described narratives of the impact of having caries and the treatment they had received. Some of the younger (those aged 5-7 years in general) children did not always express themselves in sentences but would nod or shake their head in response to questions or provide “yes” or “no” answers. Gestures where noted were included in the transcripts. Their responses have not been quoted but their views included in the analysis where possible.

The overarching theme related to pain, within this theme there were three subthemes: impacts related to pain, methods which were used to reduce pain and emotional aspects resulting from pain. Another theme related to the aesthetics of dental caries was also identified. These themes will be explored using quotes from the children to illustrate the impacts experienced and how the children described them.

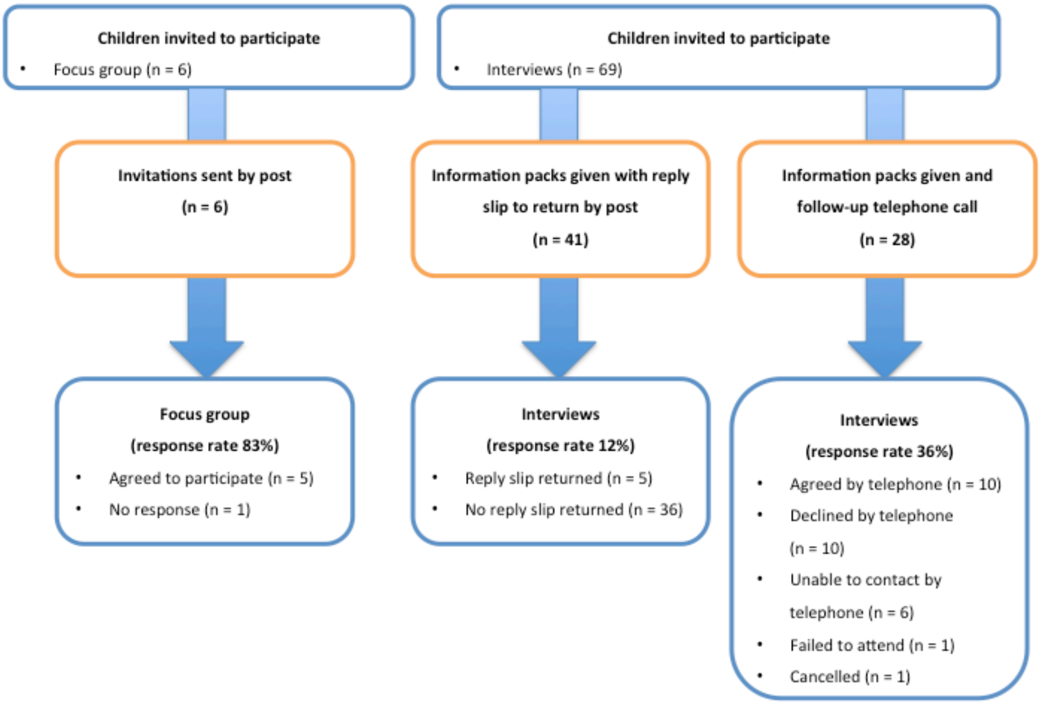

Figure 6.3. Recruitment details for focus groups and interviews.

Pain

Not surprisingly pain was the impact that most bothered children. Most children described pain as *"hurting"* or *"it hurts"*. Whereas others described the pain as toothache. Some were prompted by their parents to call it toothache and some weren't sure if it had a particular name.

> *"Well, I wouldn't call it toothache, but if that's the name then, that's what I would call it" (Mark, aged 8 years)*

When children were asked how they would describe toothache, as opposed to another type of pain, they described the quality and severity of the pain:

> *"Like, it kept going on." (Liam, aged 13 years)*

> *"… it's where your tooth aches and hurts quite a lot." (Brodie, aged 9 years)*

Children described the pain that they had experienced in a variety of ways, some using analogies to other sensations of discomfort:

> *"… it felt like you wanted to itch it and pull it out." (Mark, aged 8 years)*

> *"It was like sharp." (Liam, aged 13 years)*

> *"… like it prodding." (Lily, aged 12 years)*

The length of time the pain had lasted was mentioned only in relation to it stopping because they had taken analgesia. The location of the pain was described by some.

> *"It's the tooth, and then, my gum bit below." (Lily, aged 12 years)*

Others described the pain as "earache", which may be due to referred pain or because the child had heard of "earache" and used this to refer to any type of pain.

> *"It were mainly earache." (Leah, aged 6 years)*

Children were probed about the nature of the impacts and whether it was the frequency of the pain that bothered them most or whether it was how severe it was. For example, they were asked what bothered them most, was it how often their tooth had hurt or how much it had hurt. Those who hadn't experienced dental pain were asked the same question about pain in another part of their body. Generally it was the severity of the pain that bothered children.

> *"Would it be how much it bothered you or would it be how often it hurt would it be the main thing for you do you think?" (FG, interviewer)*

> *"Erm, probably the how much it hurt." (Jack, aged 11 years)*

> *"Erm, how bad." (Elizabeth, aged 7 years)*

However, some children did describe the frequency as bothering them.

> *"Does it matter how often something hurts or is it how bad it hurts?" (FG, interviewer)*

> *"It's how often." (Isabelle, aged 7 years)*

The severity of pain was described as hurting "a lot" or "so much'.
However, "a lot" was also used to describe the frequency of pain:

> *"I keep getting toothache a lot." (Lily, aged 12 years)*

However, pain was typically described in terms of severity rather than frequency.

Impacts related to pain

A number of impacts related to pain were reported, some of these were caused by pain and others were activities which were hampered because of the potential to incite pain, for example, eating.

Children reported a variety of difficulties encountered such as not being able to bite with their sore teeth and not being able to eat some foods or finish their meals:

> *"It did stop me eating on my teeth." (Wayne, aged 9 years)*

> *"Yeah, like, I can't eat some food…" (Lily, aged 12 years)*

> *"… but I didn't always eat it all." (Mark, aged 8 years)*

In addition, they reported that foods often got stuck in the cavities in their teeth which caused pain, with some avoiding certain foods to stop this happening.

> *"I couldn't eat apples cos that, cos skin kept going in." (Brodie, aged 9 years)*

The other main impact associated with pain, was being woken from sleep or not being able to sleep.

> *"Yeah, I couldn't get to sleep." (Jack, aged 11 years)*

> *"I had like half an hours sleep then I kept waking up and it started hurting again." (Lily, aged 12 years)*

Children who had pain during the day and at night, felt that it was worse at night. This lack of sleep caused the children who had experienced this to sometimes feel tired the next day at school and some felt it interfered with their schoolwork.

> *"Yeah, I get tired at school." (Lily, aged 12 years)*

Other impacts which were experienced included not being able to do planned activities such as going out with friends or family, difficulty talking and pain when brushing their teeth.

> *"When it's that bad, I just sit there and go, like that, and I don't eat and talk." (Lily, aged 12 years)*

In terms of impact relating to eating, again descriptions were a mixture of severity and frequency. Severity was described as "a bit", "a little bit" and "a lot", with frequency being described as "sometimes" especially regarding how often food got stuck in their teeth. Discussions were predominantly about the severity of the symptoms they experienced or by how much they had to moderate their eating.

"When you were eating did you have to eat more carefully, or slower or anything like that?" (FG, interviewer)

> *"A lot slower." (Mark, aged 8 years)*

"Apart from sleeping, was there anything else that it stopped you from doing?" (FG, interviewer)

> *"Erm, not too much, it hurt when I ate a little bit." (Jack, aged 11 years)*

Steps taken to reduce pain and associated impacts

Children used a variety of methods to alleviate or prevent pain. Children who experienced pain while eating, discussed methods they used to prevent symptoms or food getting stuck in their teeth. These included changing the way they ate or as mentioned previously, avoiding foods which were likely to cause pain.

> *"Just, er, tried to eat carefully." (Jack aged 11 years)*

> *"I just had to eat on the other side of my mouth." (Precious, aged 11 years)*

Participants also discussed eating more slowly, but the predominant strategy was to eat on one side of their mouth, thus avoiding the painful tooth. Where there was bilateral pain, this was not possible so soft foods like soup or sandwiches would be consumed.

"Medicine" was also seen as something which could be used to alter the pain. Children discussed this as a generic term and did not always distinguish between analgesia and antibiotics. Indeed, it is likely that the younger children were not aware of the type of medicine prescribed. However some participants remarked that having "medicine" didn't always relieve the pain when referring to analgesia, whilst others found it difficult to swallow tablets.

> *"Well, cos it was hurting, I, I had some medicine." (Annie, aged 8 years)*

> *"Yeah, I don't like taking tablets." (Liam, aged 13 years)*

Medicine was also seen as something that could reduce the other impacts associated with pain, such as sleeping:

> *"Well, before I went to bed my mum gave me some medicine." (Precious, aged 11 years)*

Both parents and children reported issues where children required analgesia during school hours, as schools required consent to administer pain relief or parents having to take time off work to bring analgesics to school.

"Does it stop you doing your schoolwork quite as well?" (FG, interviewer)

> *"Yeah so I go erm, medical tutor" (Lily, aged 12 years)*

"And do they give you some medicine?" (FG, Interviewer)

> *"No, they ring my mum." (Lily, aged 12 years)*

These impacts were not generally described using severity or frequency, although children did discuss that "sometimes" they took these measures to alleviate or avoid pain but this seemed be related to whether they needed to, rather than how often they did it.

Emotional issues

Children reported feeling "annoyed" by the pain in their teeth or by the effect the pain was having on their lives.

> *"Yeah, annoyed cos it's annoying me that I can't do much things cos of pain, like I can't eat and sleep and stuff." (Lily, aged 12 years)*

Children also reported feeling “sad”, “grumpy” and the pain causing them to cry. “Annoyed” was typical term used to describe how children felt about their teeth. Some children who did not use the word “annoyed”, were asked if they knew what “annoyed” meant and could readily give a definition for it.

> *“Annoying means, em, you don’t like it much and …” (Danesh aged 6 years)*

There was a sense that some children felt that it was unfair that they were suffering and that they were *“worse than most people”.* This sense of injustice was reiterated by parents who mentioned that their children often said things like *“why does it have to be me”*, when they were suffering pain from their carious teeth.

Aesthetics

Although several children had anterior caries, the way their teeth looked did not seem to bother them. Children who had caries in their anterior permanent teeth did comment on the appearance of their teeth, stating that it was the appearance of “holes” before they were filled that had been noticed, but that they had not been worried about this. Others reported not being bothered about the colour of the carious tooth but that it “looked better” following treatment.

> *“Well after it’s done, it looked better.” (Brodie, aged 9 years)*

Relationship between the severity of caries and impacts

There appeared to be no relationship between caries experience and the impacts experienced, indeed some children with a single carious tooth reported more severe impacts than those with multiple carious teeth. Children who were in pain or who had experienced a significant episode of pain, were able to describe the impacts they had experienced in more depth than those who had not experienced pain or ongoing discomfort. Interestingly, even children who had not reported any symptoms at their clinic visit, generally had experienced impacts related to food getting stuck in their teeth or having to eat on the other side of their mouth. No differences were observed between the impacts and the age of the participant; older children (8 years and above) were generally more able to describe the impacts in detail, whereas many of the younger (5-7 years) children would gesture or use one-word descriptions.

The transcripts contained invaluable information about the language children used to describe the impacts they experienced. Each impact that the children described is shown in **Table 6.1**. It was decided that "a bit", "a little", "sometimes", "very much" and "a lot" would potentially form the basis of the response format as these descriptors had been most commonly used by the participants. The next stage would verify which of these words would be most suitable. As the children discussed similar issues regardless of age, it was decided that only one measure would be formulated incorporating the language used by the younger children. The next stage determined which items were most relevant to children and these were included in the questionnaire to be tested for face and content validity.

6.4.2 Stage 2: Item reduction

Twenty-five children were invited to participate by FG in this part of the study, three declined to participate (two due to distress following radiographic examination and one due to time constraints). Therefore twenty-two children with caries participated in this stage (a response rate of 88%). The mean (range) age of participants was 9.4 (4.8-15.7) years with 41% (n=9) male and 51% (n=13) female. The recruited children were a representative sample of the patients attending paediatric dentistry clinics with a variety of different ethnicities and living in areas with varying levels of deprivation. Sixty-eight percent (n=15) of the children who participated in this part of the study were from the most deprived quintile in England.

Table 6.1. Items generated from interviews and focus groups.

Potential items
• Pain (hurts) • Difficulty eating some foods • Having to eat on one side • Getting food stuck in teeth • Being kept awake by pain • Feeling annoyed • Having to take medicine • Pain during toothbrushing • Having to eat more slowly • Having to eat more carefully • Crying • Front teeth looking brown • Feeling tired • Not being able to do schoolwork • Difficulty talking • Feeling grumpy • Not been able to do normal activities

The mean (range) number of caries-related impacts was six (0-12). Only one child had not experienced any of the impacts listed, this child had suffered no symptoms or any other negative effects of the disease. Pain and getting food stuck in their teeth and having to eat carefully were the items which had affected most children (n=16, 72.7%). Not being able to talk and not being able to do normal activities had each been experienced by only one child (4.5%). The results are shown in Table 6.2.

Seven children (31.8%) could not identify what had been the worst impact that they had experienced. Of those who responded, seven (46.7%) children reported that pain was the worst thing that had happened to them. The remainder chose getting food stuck in their teeth (n=4, 26.7%), being kept awake (n=3, 20%) and having to eat on the other side of their mouth (n=1, 6.7%) as the worst impact. Seven children (28%) could not identify which impact had happened most frequently. Of those who responded, getting food stuck in their teeth was chosen by six (40%) as the impact they had suffered most frequently. The worst item correlated with the most frequently occurring item in five cases (Table 6.2).

Having difficulty talking and not been able to do normal activities were endorsed the least. Since these items were not chosen as having the highest impact or occurring with the greatest frequency, it was decided to omit them from the measure to be used for face and content validity testing.

Table 6.2. Number of children who endorsed each item and items reported as having the highest and most frequent impacts.

Potential item	Number (%) with experience of impact (n = 22)	Number (%) stating item had highest impact (n = 15)	Number (%) stating item occurred most frequently (n = 15)
Pain	16 (72.7%)	7 (46.7%)	4 (26.7%)
Getting food stuck in teeth	16 (72.7%)	4 (26.7%)	6 (40.0%)
Having to eat more carefully	16 (72.7%)		1 (6.7%)
Having to eat on one side	15 (68.2%)	1 (6.7%)	3 (20.0%)
Having to take medicine	12 (54.5%)		
Having to eat more slowly	11 (50.0%)		1 (6.7%)
Crying	9 (40.9%)		
Feeling annoyed	8 (36.4%)		
Being kept awake	7 (31.8%)	3 (20.0%)	
Difficulty eating some foods	6 (27.3%)		
Feeling tired	6 (27.3%)		
Feeling grumpy	6 (27.3%)		
Not being able to do schoolwork	5 (22.7%)		
Pain during toothbrushing	4 (18.2%)		
Front teeth looking brown	3 (13.6%)		
Difficulty talking	1 (4.5%)		
Not been able to do normal activities	1 (4.5%)		

6.4.3 Stage 3: Questionnaire design

Ten children with a mean (range) age of 9.9 (6.4-15.5) years took part in this stage of the investigation which sought to elicit their preferences for the questionnaire format. All children who were invited to participate agreed to take part. Seven were male and three female. Four children lived in the least deprived quintile, three in the second least deprived quintile and three in the middle quintile, according to national statistics. All but two of the children preferred V2 of the questionnaire, with one having no preference although he felt that V1 looked more "official". However, all children felt that V2 was easier to read. Indeed, whilst observing children complete the two versions, it was clear that the V2 had fewer words to read and therefore seemed easier for the children to complete. This version reduced participant burden, thus making it more accessible to those with low literacy levels. The children also preferred V2 over the other examples they were shown.

Several suggestions were made to improve clarity, such as changing the word "example" on the front page as this was difficult for younger children to read. It was also suggested that an additional box around the example on the instruction page would make it stand out. Rewording of the sentence introducing the questionnaire was also suggested to make it easier to understand. It was also suggested that "Please circle one answer" should be positioned at the top of each page to remind participants of what they had to do. The question relating to "eating on the other side of your mouth" was reworded to make the sentence shorter as several children struggled to read it in the original format as it spanned two lines of text.

Children were asked which of the words "not at all", "a bit" or " a little" and "a lot", "sometimes" and "very much" they preferred and were asked to rank them from worst to best. Most children felt that "a bit" and "a lot" were words that they used most often with their friends and all were able to rank these consistently in the correct order. "Very much" and "sometimes" were also popular but more so with the older children. One child stated that he would never use "a little" and only used "very much" when he was talking about food! The younger children also took longer to think about how to rank "sometimes" or "a little" and "very much", but the majority were able to do this.

Some children felt it did not matter what colour the paper was. However the majority of those who did express an opinion preferred coloured paper as they felt it made it easier to read or made it look more interesting.

It was therefore decided to use V2 with the response options: "not at all", "a bit" and "a lot" on coloured paper in keeping with the amendments suggested by the participants.

6.4.4 Stage 4: Face and content validity

A total of eight children took part in the analysis of face validity, with four of those with active caries also participating in the analysis of content validity (Table 6.3). Twenty-five children with active caries or caries experience took part in analysis of content validity, including two with learning disabilities. Four children took part in the face validity investigation only (three girls aged 8.9, 10.6 and 11.9 years and one boy aged 12 years). Four boys (aged 6.5, 9.4, 9.4 and 12.5 years) took part in both face and content validity. Twenty-one children with a mean (range) age of 8.2 (5.2-14.4 years) years took part in the analysis of content validity only. All children who were invited to participate agreed to take part. Thirty-eight percent (n=8) were male and 62% (n=13) were female. The majority of children lived in the most deprived quintile in England (48%, n=12). Twenty percent (n=5) lived in the second most deprived quintile, 16% (n=4) in the middle quintile, 4% (n=1) in the second least deprived quintile and 12% (n=3) in the least deprived quintile. Six children who participated in the content validity analysis were recruited from a primary care clinic, the remaining children were recruited from secondary care clinics at the Charles Clifford Dental Hospital. The participants were a representative sample of the children who attend the clinics at the Charles Clifford Dental Hospital and Firth Park Clinic and included children with a range of different ethnicities and living in areas with varying levels of deprivation. The testing took an iterative approach, with amendments being made to the questions during the period of testing to allow these alterations to be evaluated by other children.

Table 6.3. Characteristics of participants involved face and content validity stage (n = 29).

Study stage	Gender		Mean (range) age (years)	Recruitment site	
	Male	Female		CCDH	CDS
Face validity only (n = 4)	1	3	10.84 (8.8 – 11.96)	4	0
Face and content validity (n = 4)	4	0	9.45 (6.45 – 12.5)	4	0
Content validity only (n = 21)	8	13	8.24 (5.19 – 14.35)	15	6
Overall (n = 29)	29	32	9.17 (4.82 – 15.72)	55	6

CCDH = Charles Clifford Dental Hospital
CDS = Firth Park Clinic

Two participants who had caries affecting their anterior teeth stated that they had noticed holes in them. They felt that a question related to this should be included. One question was added to the questionnaire for subsequent face and content validity testing. This was "I worry that I can see holes in my teeth".

Following further discussion with other participants it became clear that children were answering the question about having holes in their teeth positively, even if they could not see holes in their teeth. They stated that although they couldn't see any holes, it worried them that they might. This question was reworded to remove the word "worry" to: "I can see holes in my teeth" to avoid this confusion. Some children said that their teeth looked "black" not "brown", therefore the wording of this question was altered to: "Do you think your front teeth look brown or black?". Subsequent testing of these questions showed that children understood them and answered as expected. The word “grumpy” was changed to “cross” as children felt that this was more indicative of how they felt.

A global question was also formulated, based on that used in CPQ: “Overall, how healthy are your teeth?” (Jokovic et al., 2002). Response options were “Not healthy”, “A bit healthy” and “Very healthy”. The term “healthy” had been discussed with some children early in the face and content validity stage of development. Children were able to define healthy and seemed to understand it in relation to their oral health. However, during testing it became clear that the participants were not answering this question as expected. When asked why they had answered “a bit healthy” or “very healthy”, those children who were able to offer an explanation stated that their response was based upon how often they brushed their teeth, how often they ate fruit, whether they sometimes forgot to brush their teeth or whether they ate sweets sometimes when they knew that these were not a healthy snack. None of their answers actually related to the current condition of their teeth, so even children who reported numerous impacts did not report that their teeth were not healthy. In fact, no child chose the option “not healthy”. There appeared to be a stigma associated with this term, which may be due to children being taught about the importance of health and to admit to not being healthy may not be viewed as socially acceptable. As this question had poor face validity, two further global questions were tested. The first was: “How much do your teeth bother you” and the second was: “How much of a problem are your teeth for you?”. Both questions had the same response format as the other questions. The second global question was felt to be better understood and was therefore included in the final version of the questionnaire.

Further minor amendments were made to the introductory sentence and to the example on the front page following feedback from the participants. No further items were added and no items were felt to be irrelevant. Therefore the final version contained 16 questions, scored 0-2 (“not at all”=0, “ a bit”=1 and “a lot”=2) and one global question using the same response format. The questionnaire was printed on light lilac paper, as this colour was preferred by most children.

The resulting questionnaire (Appendix L), named the Caries Impacts and Experiences Questionnaire for Children (CARIES-QC) took approximately 2-5 minutes to complete depending on the reading ability of the child, with some younger children requiring assistance to read it. CARIES-QC, including the instructions which are intended to be read with a parent, has a Simple Measure of Gobbledygook (SMOG) index of 5.3 and a Flesch-Kincaid grade score of 1.5 indicating a reading age of 6-10 years (Edit Central).

6.5 Discussion

This chapter has described the development process of CARIES-QC using child-centred methodologies. The input of children throughout the development of the measure was critical and has ensured that future young participants will find it easy to complete as it uses words which they are familiar with and with a simple response format. The following discussion will now reflect on this first phase of the study and will critique the findings.

6.5.1 Outcome

The aim of this part of the study was to produce a caries-specific measure of OHRQoL for children aged 5-16 years of age in accordance with the definition of OHRQoL. This was achieved through multiple stages and a measure has been produced which is able to be understood by children in the target age range and which covers the aspects of dental caries which are of concern to children.

The specific objectives were to:

1. Describe the impacts of dental caries from the child's perspective
2. Evaluate how children describe these impacts;
3. Determine which of the reported impacts were of most concern to children;
4. Evaluate which design and response format children prefer
5. Assess face and content validity of the measure.

By employing the structured approach advocated by Guyatt and co-workers (1986), all of these objectives were fulfilled and they will be discussed in further detail in the following sections.

6.5.2 Development of the measure

Ability to discuss the impacts

Children who participated were competent in discussing the caries-related impacts that they had experienced. It was noted that younger (aged 5-7 years) children often gave shorter answers with less depth, but they were still able to clearly describe their experiences. The interviews and focus groups were relatively short, this was due to the focus of these being narrow. In addition, it is recommended that interviews should be short when involving young children (Borgers et al., 2000). Previous studies, described in Section 4.4.1, investigating children's perceptions of their own illnesses have also found that children are capable of discussing how their condition affects them (Beales et al., 1983; Walsh and Bibace, 1991; Herrman, 2006). Piaget and Innhelder (1969) in their description of cognitive theory (Section 4.2.1) reported that children from two to seven years are in the "pre-operational stage" and they only place meaning on the parts of their body that they can see. They also find it hard to understand that there are parts of their bodies which they cannot see that might have an effect on their lives. Whereas from seven years (concrete operational stage) they start to be able to appreciate cause and effect from their own experience and by around 12 years are capable of abstract thought (formal operational stage) (Piaget and Inhelder, 1969). These concepts are linked to how children view illness and it has been found that children early in the pre-operational stage think of illnesses as having developed magically. As they get older they gain understanding about the transmission of "germs" and finally acquire a more sophisticated understanding of illness involving internal structures and psychological aspects (Vacik et al., 2001). However, more recent work has demonstrated that although there may be some systematic manner in which this knowledge is acquired, there are likely to be individual differences in understanding between children depending on their experiences

(Eiser and Kopel, 2013). Although children's understanding of caries was not explored in this study, several of these concepts could be seen in the way that older (aged 8 years and above) children described the impacts as having emotional as well as physical effects without probing. The older children were also able to appreciate that they were sometimes given medicine by their parents to enable them to sleep, whereas the younger children knew that medicine would stop it hurting but didn't discuss the wider context.

Previous studies have reported that as children get older their description of pain becomes more sophisticated (Gaffney and Dunne, 1986; Harbeck and Peterson, 1992; Crow, 1997; Franck et al., 2010), however, others have found this not to be the case (Ross and Ross, 1984). In this study, most of the children used the term "hurts". This was found to be one of the most common words used by North American Children to describe pain (Mills, 1989; Stanford et al., 2005). The word "hurt" was also found to be commonly used in a UK study where parents were asked about their child's use of pain words (Franck et al., 2010). There did appear to be some differences between the description of pain by the older children who used more descriptive terms (*"… it felt like you wanted to itch it and pull it out" Mark, aged 8 years; "It was like sharp" Liam, aged 13 years and "… like it prodding" Lily, aged 12 years).* There was however, a lack of temporal terms, such as pulsing or throbbing, which has been found previously and was discussed in Section 4.4.2 (Savedra et al., 1995). Dental pain is often described by adults as a "throbbing" pain and therefore it is interesting that none of the children in this study described their pain in this way. There is little in the literature to account for how children learn the names for different diseases. It has been reported that children learn to describe illness from what they learn from their family context (Campbell, 1975). However, young children often use terms such as "headache" or "tummy ache" inappropriately to describe pain in other areas of their bodies (Liossi et al., 2012). It is therefore not surprising that some children were unaware of the term "toothache".

The use of the word "earache" to describe the pain was interesting as it was not possible to ascertain whether children thought that the pain had been near their ear or whether this was just a word they used to describe any pain around that area. Parents who were asked how their children expressed pain discussed that often they would use terms like "tummy ache" or headache" to describe pain in other parts of their bodies (Liossi et al., 2012), therefore the use of these types of description may be part of how children learn to express their pain experience.

Some other reported impacts were, perhaps, not surprising as it is well known that eating and sleeping can be affected by dental pain. However, the restrictions that the children described in terms of diet due to being unable to eat some foods because they got stuck in their teeth or because they were too hard to bite, is significant when it is considered that this may have been an issue for some time. Children who described eating-related impacts had used some avoidance method, either avoiding certain foods or eating differently, to cope with the pain they were experiencing. It was interesting to note that when some of the children were completing the questionnaires for the content validity stage, they answered that they had experienced some impacts related to eating, but their parents were unaware of this. This may indicate that the children had found a coping mechanism and therefore had not discussed it with their parents, only reporting more acute symptoms. This also highlights the importance of involving children in the item generation process but also the value of self, rather than proxy, reporting which has been highlighted elsewhere in the child pain and HRQoL literature (Zhou et al., 2008; Eiser and Varni, 2013).

Most of the younger children required more probing to elicit their responses, but the information they provided was invaluable in ensuring that the language used in the final version was appropriate for younger as well as older children. Although they described the impacts in less depth, their experiences were largely similar to the older respondents, supporting the development of a single measure. A previous study found that children aged five to eight years were able to complete the short form version of CPQ which was designed for 11-14 year-olds (Foster Page et al., 2013a).

Language

The children generally discussed the impacts in terms of severity rather than frequency which is similar to the findings of others (Stevens, 2009; Marshman et al., 2010; Carlton, 2013a). Whilst observing children complete the questionnaire, some seemed to consider the frequency and the severity with which an impact has occurred to help them arrive at their answer; for example, how often food gets stuck combined with how much it hurts when it does. These response options are similar to those described by Stevens (2009) and Carlton (2013b) following evaluation of the words children had used in their interview. Although these studies all took place in Sheffield, the words used by the participants have appeared in other measures of HRQoL developed in other areas of the UK and in other countries for both adults and children (Guyatt et al., 1987b; Bradley et al., 1999; Wille et al., 2010). However, further testing of the measure in other areas and countries will confirm if any alteration is required.

A lack of face validity was discovered with the question “I worry that I can see holes in my teeth” when it was discussed with children. Many children reported that they had answered “a bit” or “a lot” as although they could not see holes in their teeth, theoretically they would be worried if they could see them. Therefore the wording of this question was changed to improve face validity.

The format of the global question also posed some challenges. Initially, a question which was similar to that used in CPQ was formulated (“How healthy are your teeth?”). It became clear during the testing of face validity, that children were not completing this question as expected. Their answers seemed to have little to do with the health of their teeth and be more related to their general health or oral health practices. This may be due to how children perceive health at different ages. Natapoff (1978) found that young children described health as series of specific practices such as eating meats and vegetables, exercising and being clean, when compared to older children who were more concerned with whole body states such as feeling good and being in good physical shape (Natapoff, 1978). In addition, young children did not believe that you could be part healthy and part unhealthy, therefore if children thought they were generally well, it would not be possible for their teeth to be unhealthy. Children are taught about health and wellbeing in schools and learn through the media, and there may be some social stigma associated with not being healthy, therefore children may be reluctant to admit that their teeth are not healthy. This may, in part explain weaker correlations with the CPQ total score and the global rating of oral health than with the global rating of life overall (Jokovic et al., 2002; Jokovic et al., 2004; Robinson et al., 2005; Goursand et al., 2008; Aguilar-Diaz and Irigoyen-Camacho, 2011; Gururatana et al., 2011; Kolawole et al., 2011).

Despite these minor alterations to CARIES-QC, very little else required modification during the face and content validity testing stage. This may be attributed to ensuring that the words children used were included in the questionnaire, thereby aiding comprehension. The importance of asking children why they have picked a particular answer is essential in ensuring good face validity and proved extremely valuable in rewording the questions so that they made sense to the participants.

Questionnaire format

Children were involved in choosing between several different designs of questionnaire, two of which they could complete. The majority preferred the version with a statement and separate response scale (V2) where the desired option was circled as there was less to read. It was interesting to watch the children complete the questionnaires, as the statement-based version (V1) took much longer for them to read. They would read all of the words in each response option before deciding upon a response, whereas as adults we may skim read the options before honing in on the preferred option. Younger children in particular, did not appear to realise that there was a format to how these were written and carried out the same procedure for each question. It therefore took them almost twice as long to complete as the alternate version. As children with dental caries come predominantly from areas of high social deprivation or from minority ethnic groups where English is not the first language, reducing the number of words which have to be read is desirable.

Indeed, while assessing face and content validity some parents struggled to read the questions to their children.

A three-point scale was chosen to aid completion by younger participants and also to reduce the overall volume of reading. The CAT-QoL initially had a five-point Likert scale, however, following Rasch analysis this was redesigned as a three-point scale as it was clear that, despite the ranking exercise which had been used (Section 6.1.3), children could not reliably use the five-point version (Carlton, 2013c). Gherunpong (2004a), in the development of the C-OIDP also found that a four-point scale was more suitable for children than a five-point scale, while parents who commented on the SOHO-5 also felt that five responses would be difficult for children to comprehend (Gherunpong et al., 2004a; Tsakos et al., 2012). Three-point Likert scales have been successfully used in other measures of HRQoL for use with children (Varni et al., 2001). However, reducing the number of response options may decrease the instrument's ability to detect change and attenuate its test-retest reliability due to the limited number of values which can be chosen (Weng, 2004). The reliability and responsiveness of CARIES-QC will be discussed in the next chapter.

A recall period was not included in the measure, as it was felt that this would complicate the process for young children. Others have suggested relatively short recall periods should be used (Matza et al., 2004). However, it was felt that limiting the recall period to a week for example, would not be useful due to the often sporadic nature of symptoms from carious teeth.

6.5.3 Study design

Participants and recruitment

Recruitment to the interview stage of the study was challenging and yielded a relatively low response rate (27%). The initial approach involved giving the family an information pack to take home with a reply slip and prepaid envelope for its return. The majority of these patients were seen as new patients and therefore they may have had other information given to them on the day (consent forms, information sheets, diet diaries etc.) and may have overlooked or forgotten to post the reply slip. There was a sense that perhaps the families were overloaded with information at this time. Non-return of reply sheets is a common occurrence in national dental surveys even when individuals are sent reminders by the school (Public Health England, 2012; Public Health England, 2013). However, in this study the response rate did improve slightly with a follow-up telephone call. As the parents were acting as "gatekeepers" it is difficult to know whether it was the child's decision not to participate or that of their parents. Reasons given by the parents were that: the child wasn't interested, they would find it difficult to organise a convenient time due to work or their children's extracurricular activities or that they had an illness in the family that would make it difficult to participate. It is clear that families are busy and have many competing activities to arrange. Therefore it may be best in similar populations to perform interviews along with scheduled appointment visits, in a room away from the clinic to make the environment more comfortable and less intimidating. This was not possible in the majority of cases for the present study as most of the eligible participants were returning for treatment under GA only. Despite the low response rate, bias is likely to be low as children living in areas with varying levels of deprivation, of different ages and with varying clinical presentations were recruited. However, it is unknown if those who chose not to participate may have had different experiences than those who did choose to take part. This is an acknowledged limitation of the study design and may have introduced some bias within the study group. Interestingly, recruitment to all other stages was found to be high. This is likely to be due to minimising the burden for the families by asking them to participate during scheduled clinic appointments.

Ethical issues

During the planning of these stages consideration was given to how to reduce power imbalance, confidentiality and interviewer safety as described in Section 4.3. FG undertook a three-day training course in in-depth interviewing techniques (National Centre for Social Research, London), which ensured adequate preparation for the technical, theoretical and ethical aspects. No ethical or governance concerns arose during the study conduct.

As FG had approached the children and their families in the clinical environment, they were aware that they were being interviewed by a dentist. However, FG was not involved in their care nor actively involved in clinical activities when they were invited to participate, this perhaps made FG's clinical role less apparent. In addition, although the participants were purposively sampled to have a range of presentations, FG did not look at their clinical notes until after the transcripts had been analysed to minimise any preconceptions about their experiences. The interviews were either held in the participant's homes or in a non-clinical area, ensuring distance from the clinical setting. Children and their parents chose where to conduct the interviews within their homes and were given a choice of seating areas when these were conducted in non-clinical settings. This aimed to give the participants some control over the situation and to reduce the power imbalance. The children did not seem to be inhibited by the fact that FG was a dentist and were happy to discuss their experiences of dental caries and the treatment they had received to manage it. In fact, although comments about the dental profession and the treatment they received were generally positive, both children and their parents were openly critical at times, indicating that they did not feel inhibited by FG's status as a dentist.

During the interviews and focus groups, no issues were raised which required safeguarding procedures to be followed. The majority of interviews took place in the participant's homes, there were no safety issues and the interviewer was made to feel very welcome at all times.

6.5.4 Limitations

A limitation of this part of the study was that the participants were from only one city in the UK and therefore the language used may be specific to children from Sheffield. Colloquialisms such a "kinda" (to describe severity) and "dint" (to describe a hole) were not included in the measure. Therefore it is hypothesised that the words used and experiences are generalisable but further testing would be required to confirm this. The children who participated in the focus groups and interviews had a variety of clinical presentations; some were in pain while others were not, some had experience of dental treatment for caries, whilst others had limited dental experience and they had differing numbers of carious teeth with varying extents. It can therefore be assumed, that although they were predominantly recruited from a secondary care environment, the experiences they discussed have representational and inferential generalisability, as they reflected the different presentations seen in clinical practice (Ritchie et al., 2003). It would be expected that the experiences of children in this population would be similar to those in the rest of the UK and indeed further afield; although there may be slight differences in the way children from other areas or countries may describe these impacts. However, it should be noted that these findings may be different from those found in a population sample where there are lower levels of caries experience.

The reliability and validity of the qualitative data may be affected by the large number of non-responders and as it was not possible to compare these non-responders with the participants due to ethical constraints, it is unclear whether these non-responders may have differed from those who did respond in some way. However, the fieldwork and analysis were conducted in a robust and consistent way, ensuring the reliability of the data obtained (Ritchie et al., 2003). The transcripts were analysed by dentists, this may have imposed a clinical perspective upon the data, although none of the analysts were aware of the child's clinical presentation which attempted to minimise the impact of this. The findings of the interviews and focus groups are not unexpected and relate to issues that children report to be symptoms of dental caries which would also tend to imply their reliability and validity. This study did not set out to validate children's descriptions, rather it aimed to discover the language they used in order to create questionnaire items that were meaningful and relevant to them.

Only children who could speak English were included in all parts of the study, as undertaking qualitative data collection through interpreters is difficult and costly. As the aim was to discover the words which children used, it would have been inappropriate to include those provided by adult interpreters. Children from families where one or both parents did not speak English were included, where appropriate consent was attained, to ensure that as many children from ethnic minorities were involved as possible.

The focus groups involved a small number of children as is advised when conducting focus groups with young children (Hoppe et al., 1995). This may limit the range of views which are expressed, but may aid the inclusion of young, quiet children. Although the number of children in the focus groups was small, the issues that individual children raised were then discussed with the group to elicit whether they had similar experiences. In this study, the data from the focus groups were used, in part, to inform the topic guide and therefore acted as a guide to important themes and the vocabulary which could be explored in the interviews. The interviews were also relatively short in length, this reflects the narrow focus of the topic (i.e. impacts related to their experience of caries only) and the fact that some children had experienced few impacts as a result of dental caries. It should also be noted that the recorded duration of the interviews did not include time spent providing introductory information, gaining of consent, choosing of pseudonym or post-interview debriefing and therefore relates entirely to the time spent talking about the topic. Further in-depth qualitative enquiry may allow greater understanding of the impacts discussed in this study and their relative importance to be explored in more detail.

The focus groups and interviews were guided using the sociology of childhood as a theoretical underpinning (Holland et al., 2008). This allows the data to be taken at face value and does not impose any restrictions upon the data. This approach was chosen as the aim was to understand how children describe the impacts of caries in order to create questionnaire items which reflected these in the language children used. Theoretical models commonly used within oral health-related quality of life research, such as the Wilson and Cleary model (Wilson and Cleary, 1995), would have imposed a structure upon the data rather than allowing the focus to be on how children described their experiences. Further analysis of the interview data using such a model may allow further insights to be gained in the future, however, this was outwith the remit of this study.

A further limitation of the study was that some demographic and caries experience data was not collected from children in Stages 2, 3 and 4 as consent for this was not sought. As these children were recruited consecutively, they were felt to be representative of the children who attended the clinics at Firth Park and the Charles Clifford Dental Hospital who had caries experience. There may be reasons why this would not be the case and therefore it is not possible without this data to confirm that these participants were a representative sample, although the deprivation data would suggest that they were. The children were however, a representative range of ages and they suggested changes to the wording and addition of items which were then further discussed with other children. It can therefore be assumed, that a range of views were obtained.

6.5.5 Novel aspects

This is the first study to involve young children at all stages of the development of a measure of OHRQoL. The value of this approach is evidenced by the limited number of changes required during the face and content validity stages, as not only the concepts the children discussed were used but also their terminology, enabling the questions and response format to be understood. It has also demonstrated that young children are capable of discussing their experiences of dental caries.

Discussion of the items with children throughout the study and the reasons for their responses provided insight into how children evaluate the different options. This was particularly important during development of the global question, as it became clear that children were not choosing the expected responses when asked about the "health of their teeth". This may explain why, when CPQ has been used in other studies with children with caries, the global question on life overall has been shown to have a higher correlation with the total score, than with the global rating of oral health (Jokovic et al., 2002; Foster

Page et al., 2008). This finding has important implications for development of future measures which may wish to include a global measure of oral health.

However, development of a measure in this way does take considerable time and was described by Guyatt (1986) as the "Rolls Royce model". This may explain why some of these steps have been omitted in the development of existing OHRQoL measures for children (Guyatt et al., 1986).

6.5.6 Summary

A 16-item questionnaire has been developed to measure the impact of caries using child-centred methodologies, and has been found to have good face and content validity. Children have been involved at all stages of development which has resulted in a measure which is relevant to children and has acceptable readability for the intended audience. The next stage will test the measure for reliability, validity and responsiveness.

6.6 Publications arising from the work presented in this chapter

6.6.1 Peer-reviewed journal article

- Gilchrist F, Rodd HD, Deery C, Marshman Z. The impact of dental caries on children and young people: what they have to say. *International Journal of Paediatric Dentistry* 2015; 25: 327-338.

6.6.2 Published abstract

- Gilchrist F, Deery C, Rodd HD, Marshman Z. Child-reported Impacts of Dental Caries. *Journal of Dental Research* 2014; 93(Sp Issue B): 1200.

Chapter Seven

Evaluation of CARIES-QC

7.1 Introduction

The previous chapter discussed the development of CARIES-QC using child-centred methodologies. This chapter will report on the testing of this measure in terms of its validity, reliability and responsiveness (Stages 5-7 Figure 6.1, page 110).

The findings of the systematic review presented in this thesis found that item response theory (IRT) had not been used to evaluate any of the included measures. Unidimensionality is an important property as CARIES-QC has been developed to measure change following interventions for the management of dental caries and this can only be evaluated accurately where an interval scale can be created. To ensure that CARIES-QC is unidimensional, a form of IRT (Rasch analysis) will be used. A description of this technique will be described in detail in this chapter. As discussed in Chapter Five, the COSMIN checklist can be used to ensure robust evaluation of these elements and has been used successfully in the development and validation of several PROMs (Thorborg et al., 2011; Kwakkenbos et al., 2012; Malliaropoulos et al., 2014). The COSMIN group recommends that a number of aspects should be reported and evaluated, these will be detailed in the following sections and CARIES-QC evaluated using this framework (Mokkink et al., 2010a).

7.1.1 Item response theory

Where studies use an item response theory method, the type used should be stated. IRT encompasses a group of models which examine latent trait theory. Trait level estimates depend on both the participant's responses and the properties of the items administered (Embretson and Reise, 2013). This produces an item characteristic curve which models the participant's response to an item and their level on the construct measured by the scale (Edelen and Reeve, 2007). An example is shown in Figure 7.1. Several models are available where polytomous scales are used. IRT and the Rasch model vary in their approaches. IRT models examine the data and find a model which fits whereas Rasch provides one mathematical model to guide the production of a linear scale (Cano and Hobart, 2011). Production of a linear scale enables accurate calculation of change scores and therefore may be more suitable for measurements relating to health interventions.

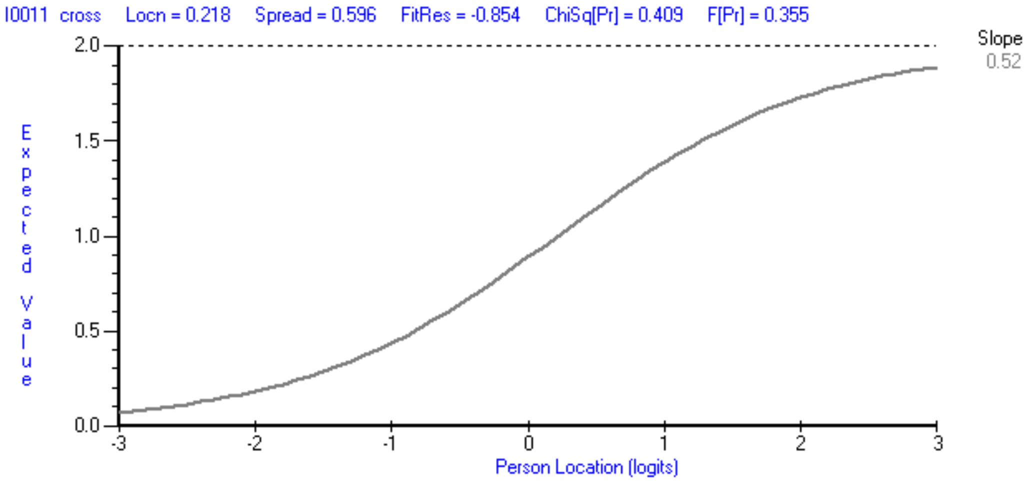

Figure 7.1. Example of an item characteristic curve.

This shows a participant's expected score depending on their location (level of impact) on the logit scale, with a higher level of impact indicated by a higher positive logit location.

Rasch measurement model

The Rasch measurement model was originally used in educational testing, but more recently has been used in the development and validation of PROMs (Batcho et al., 2012; Chien et al., 2014; Shelton et al., 2015). Internal consistency of a scale is usually represented by Cronbach's alpha, however, approaches based on the Rasch model of measurement have also been applied to accurately assess unidimensionality of patient-reported scales (Tennant and Conaghan, 2007). Rasch analysis allows formal testing of a scale against a mathematical model as first proposed by Georg Rasch (Rasch, 1960). Data from questionnaires which are intended to be used to achieve an overall score are tested against the measurement model's expectations. These expectations are based on the probabilistic form of a Guttman scaling and a variety of statistics show how well the responses fit the model (Guttman, 1950; Smith, 2000). Thus, the model expresses the probability of an item that represents a given level of impact (difficulty) being endorsed by people with a given level of experience of that impact (ability), as a logistic function of the difference between the person's ability and the item difficulty (Rasch, 1960). These values are expressed in logits (log odds probability units), which allows the scores achieved by participants to be evaluated as an interval scale, rather than the ordinal scores obtained from the raw data. This transformation allows more meaningful interpretation of the original ordinal data, as all raw scores are non-linear and therefore the values at the margins of the curve cover a wider part of the underlying trait than those at the centre (Figure 7.2) (Tennant and Conaghan, 2007). Thus, change scores can be accurately calculated using the interval scale produced.

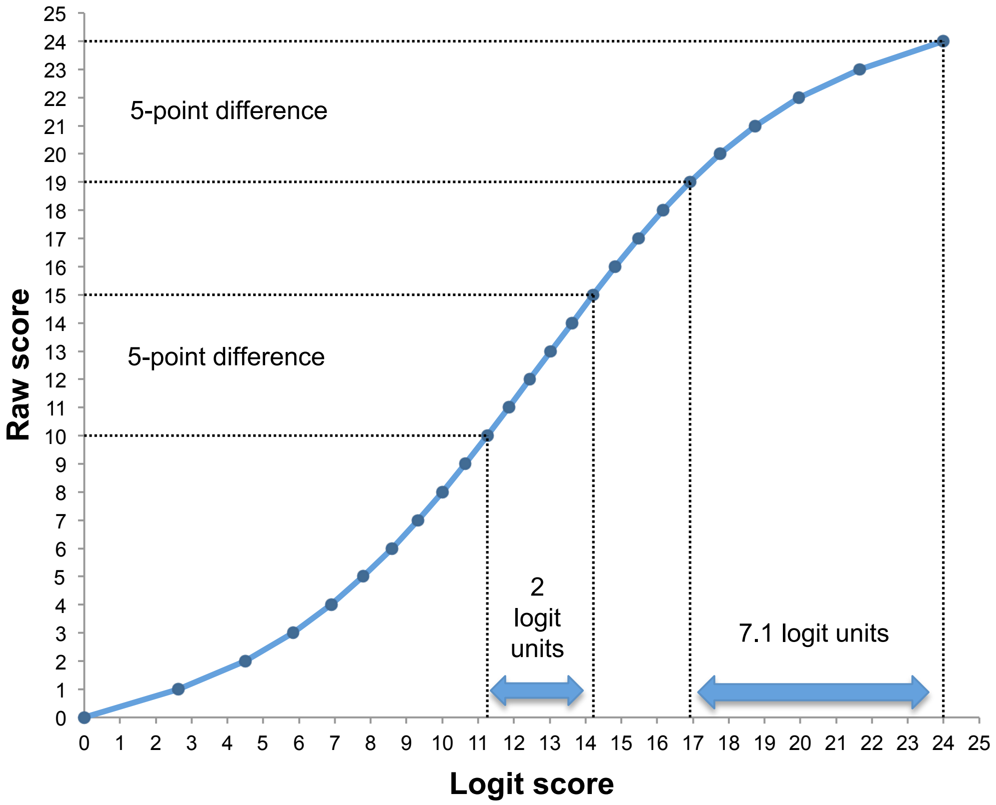

Figure 7.2. Raw score to linear measurement transformation.

Raw scores of CARIES-QC are shown on the y-axis and logit scores on the x-axis. Distortions can be seen in the raw score due to the sigmoid curve, therefore a five-point raw score difference in upper tail is equal to 7.1 logits compared with a five-point difference in the mid-part being equal to 2 logits.

Using Rasch analysis in the development and validation of a measure ensures fit to the model from the outset. Items which are chosen for the final measure should be unidimensional, free from differential item functioning (i.e. the items function in the same way across groups, for example: gender, age) and fit the model expectations (Tennant and Conaghan, 2007). The method for exploring each of these aspects is described below.

- Local independence: To be valid, a scale must demonstrate local independence. That is, responses to previous items should not influence those that follow. This can be examined using residual correlations between items, which should be no more than 0.2 – 0.3 above the average residual correlations (Marais and Andrich, 2008).

- Unidimensionality: The Rasch model requires the questionnaire to measure only one construct. This can be examined using two subsets of items, one subset where items load positively and the other where items load negatively following principal components analysis of the item residuals. Unidimensionality is then analysed using an independent t-test on the estimates derived from the subtests (Smith Jr, 2005; Tennant and Pallant, 2006).

- Invariance: It is important that item difficulty remains stable amongst participants of different groups, e.g. age and gender. This is examined using analysis of variance using the group as the main factor. Where variance is observed, this is termed Differential Item Functioning (DIF). DIF can be uniform (bias is present across the trait) or non-uniform (bias is not consistent across the trait) (Holland and Wainer, 1993; Grimsby, 1998). Items where DIF is demonstrated should be removed from the measure.

7.1.2 Internal consistency

Assessment of internal consistency can be undertaken using factor analysis, CTT or Rasch analysis. Exploratory factor analysis can be used to identify domains within a scale, assess unidimensionality and reduce the number of items (Williams et al., 2010). However, the method for doing this involves some arbitrary and subjective criteria to select the appropriate number of factors. As noted in Section 7.1.1, Rasch analysis incorporates the person ability estimate into the item-parameter assessment and vice versa, thus allowing generalisability beyond the original sample. In contrast, exploratory factor analysis cannot separate the person and item characteristics and therefore is only relevant in the context of the original sample (Cosco et al., 2012). Where CTT is used, Cronbach’s alpha should be calculated using data from at least 30 participants.

7.1.3 Reliability

Reliability should be tested using test-retest reliability expressed as the intraclass correlation coefficient (ICC). The two administrations of the questionnaires should be independent, within an appropriate time interval and with at least 30 participants whose condition has remained stable. The appropriate time interval depends upon the condition under scrutiny. For example, a short period may be desirable for patients in palliative care where deterioration may be rapid, however one to two weeks is usually suitable for most clinical conditions (Mokkink et al., 2010a).

7.1.4 Construct validity

Hypotheses regarding the correlations should be made *a priori* with regard to the direction and magnitude of these. In addition, for convergent validity, the comparator instrument should be adequately described.

7.1.5 Responsiveness

Responsiveness should be tested by assessing correlations between change scores using the instrument and those in the global rating score. Hypotheses should be presented regarding the magnitude of these correlations. The MID can be calculated to aid clinical interpretation.

7.1.6 Interpretability

Information should be provided regarding the demographics of the sample, floor and ceiling effects, mean scores for different groups and the MID to allow the results to be interpreted adequately.

7.1.7 Missing items

In addition to the above factors, the proportion of missing items should be reported along with details regarding how these were dealt with. There are a variety of methods available to deal with missing values. These include insertion of sample or individual mean values, single regression and multiple imputation (Shrive et al., 2006). Multiple imputation is based on a technique proposed by Rubin (1976) and aims to estimate a missing value. The missing data are predicted using existing values, these are imputed multiple times resulting in a data set which can then be analysed separately resulting in multiple analyses from which a final value can calculated (Wayman, 2003). This method is said to reflect variability which would be expected within any survey data (Shrive et al., 2006). Shrive and colleagues (2006) found that while multiple imputation was usually the most accurate method to replicate missing values, individual mean imputation also performs in a similar manner in some circumstances. Individual mean imputation is simpler to perform and may be easier for clinicians to interpret. Individual mean imputation is recommended where participants have answered over half of the questions and where the scale is unidimensional (McIver and Carmines, 1981; de Haes et al., 1996; Fayers et al., 1998). It is important to assess the nature of the missing responses as they may introduce bias or indicate questions which are difficult for participants to understand or which participants are unwilling to answer.

7.1.8 Order effect

The order in which questions are presented may influence the way participants answer future items in the measure (Perreault Jr., 1975). This may occur where participants wish to be consistent and therefore base subsequent answers on what came before (Perreault Jr., 1975). Completion of questionnaire items requires complex cognitive processes involving interpretation and memory (Krosnick, 1999). Therefore, this requires great effort, especially if there are a large number of questions. Participants who are willing to put this effort in to "optimise" their responses and produce high quality responses (Krosnick, 1991). However, where this effort becomes too hard to sustain, participants may become fatigued and distracted and answer less thoughtfully. This may mean that they choose the middle response option or may choose an answer randomly (Krosnick, 1999). This is termed "satisficing" and is more likely to occur where the items are difficult, with participants with lower reading ability and where motivation is low (Krosnick, 1991). As it may take a considerable effort for young children to read through a questionnaire, they may be at risk of satisficing. Therefore, where two questionnaires are compared, the order in which they are answered should be allocated randomly and the effect of this analysed.

7.1.9 Summary

To ensure that CARIES-QC is unidimensional, free of local dependence and DIF, Rasch analysis will be performed. Internal consistency, reliability, construct validity, responsiveness and order effect will then be assessed and reported in detail to ensure interpretability. These steps will be used to ascertain which of the original items should be retained and to validate CARIES-QC.

7.2 Aim

The overall aim of this stage of the research project is to assess the validity, reliability and responsiveness of CARIES-QC. In addition, raw scores will be transformed to an interval level scale which can be used in future studies.

The specific objectives are to:

1. Produce a unidimensional measure which is free from DIF using Rasch analysis.
2. To apply the COSMIN checklist to the analysis to ensure all elements are tested robustly.

7.3 Method

Ethical approval was gained from the South Yorkshire Regional Ethics Committee as previously detailed in Section 6.3.1.

7.3.1 Recruitment

A convenience sample of participants was recruited from new patient clinics in the paediatric dental department of the Charles Clifford Dental Hospital (CCDH) and from Sheffield Salaried Primary Dental Care Services at Firth Park Clinic, Sheffield. Participants were recruited by FG, HR, CD, CB and VW at CCDH and JH at Firth Park Clinic between July 2014 and January 2015. Children and their parents were provided with age-appropriate information sheets. Instructions were given to the families who wished to participate. Parents were advised that they could help the children to read the questions if required but that it was the children's responses which were desired. FG, HR, CD, VW or CB read the questions to children where the parent could not read English. Inclusion and exclusion criteria were similar to those applied during the development process.

Inclusion criteria

- Children aged 5-16 years at recruitment
- Children with active dental caries present
- Children who are able to understand spoken English

Exclusion criteria

- Children with other pre-existing medical conditions (American Society of Anesthesiologists' Physical Status Classification System Grade 3 or greater (American Association of Anesthesiologists, 1974))
- Children with dental conditions other than dental caries (e.g. hypodontia, cleft lip and palate)
- Children with severe learning difficulties who would be unable to participate with the intended activities even with support

7.3.2 Questionnaires

Data were collected at three time points where possible. These were:

- Baseline (T0), for example, at new patient clinic appointment
- Prior to the start of dental treatment to allow test-retest reliability (T1)
- Following dental treatment for dental caries (T2).

Baseline responses (T0)

At T0, participants were asked to complete two questionnaires: CPQ_{11-14}-ISF:16 and CARIES-QC. CPQ_{11-14}-ISF:16 contains 16 items and two global questions (Jokovic et al., 2006). Participants are asked how often (never, once or twice, sometimes, often, everyday or almost everyday) they have experienced the listed impacts during the past three months. As discussed in Section 4.6.1, the questions are scored on a 5-point Likert scale from 0-4, with increasing score indicating increased impact (possible total score thus ranges from 0–64). This version of the short form of the CPQ_{11-14}has been used successfully in children aged 5-8 years (Foster Page et al., 2013a).

CARIES-QC contains 16 items and one global question as described in the previous chapter. The items are scored on a 3-point Likert scale and scored 0-2, with increasing score indicating increased impact (possible total score thus ranges from 0–24). Participants were randomised (research randomizer) to receive either Questionnaire A (CARIES-QC followed by CPQ_{11-14}-ISF:16) or Questionnaire B (CPQ_{11-14}-ISF:16 followed by CARIES-QC) to control for order effect. Clinical and demographic data were collected including: age; gender; ethnicity; postcode (to assign Index of Multiple Deprivation (IMD) score/rank); dmft/DMFT; presence of anterior caries; pain and pulpal involvement. DMFT/dmft were assessed by experienced clinicians using a combination of clinical and radiographic evidence where available. Caries was judged to be present where there was caries into dentine clinically or radiographically. Where a tooth had a restoration and caries, the tooth was judged to be carious rather than restored. A tooth was judged to have pulpal involvement where the pulp was exposed, there were clinical signs of infection (sinus, swelling etc.) and/or where it was judged restoration of the tooth would require root canal therapy either clinically or radiographically. FG discussed the criteria with clinicians who were to collect the data and FG was available during clinical sessions in the first month of data collection to ensure that clinicians applied the criteria correctly.

Prevention visit (T1)

Participants who returned for a prevention visit prior to treatment were asked to complete only CARIES-QC (Questionnaire C) for test-retest reliability. This questionnaire contained a supplementary global rating of change question which asked whether the participant's teeth felt "the same", "better" or "worse" compared with the previous administration. Only those who reported that their oral condition had remained stable were included in analysis of test-retest reliability. At this visit children either read the questions themselves or had a parent or FG read the questions to them.

Follow-up on completion of treatment (T2)

Participants were invited to return following their last treatment visit to complete the final questionnaire (Questionnaire D), which was identical to Questionnaire C. At this visit, data relating to the treatment that the patient had received and the date of the final treatment visit were recorded. At this visit children either read the questions themselves or had a parent or FG read the questions to them.

7.3.3 Data analysis

The RUMM2030 (RUMM Laboratory Pty Ltd, Perth, Australia) software was used for all Rasch analyses, which were based on the unrestricted or partial credit model (Masters, 1982; Andrich et al., 2009). All other analyses were undertaken using SPSS 20 (IBM, New York, United States). The Index of Multiple Deprivation score/rank (2010) was calculated using GeoConvert applied to the participant's postcodes (The UK Data Service).

Missing items

Where greater than two values were missing from CPQ_{11-14}-ISF:16, the individual questionnaire was excluded from further analysis. Where greater than two questions were unanswered in CARIES-QC, the participant was eliminated from further analysis. This threshold was chosen, as it eliminated participants who failed to answer entire pages and was equivalent to approximately 10% of the items. This value (10%) was shown by Shrive and co-workers (2006) to produce a Kappa of 0.88 when individual mean imputation was used compared to the original value. Therefore, where fewer than two missing values were present, the individual mean for that participant was used (i.e. values were filled with the computed mean of the participant's completed items) (Shrive et al., 2006). This was not required for the Rasch analysis, as the calculations take account of missing values and therefore mean values are not substituted.

Item response theory

Rasch analysis was performed for the data pertaining to the initial administration of CARIES-QC. Items demonstrating misfit to the model, local dependence or DIF were removed to create a unidimensional scale. A sample size of at least 150 participants is recommended to give 99% confidence that the estimate is within 0.5 logits (Linacre, 1994).

The category structure of the scale was analysed to ensure that participants had used the response options as expected. An ordered set of response thresholds for each item should be present (Tennant and Conaghan, 2007). That is, the response options ("not at all", "a bit" and "a lot") should discriminate as expected. Where there are disordered categories, this indicates that participants were unable to distinguish between response options adequately. This may be due to having too many response options or that the labeling of the options is confusing and difficult for participants to discriminate between the different levels (e.g. sometimes, often, frequently). Figure 7.3 and Figure 7.4 show examples of ordered and disordered categories.

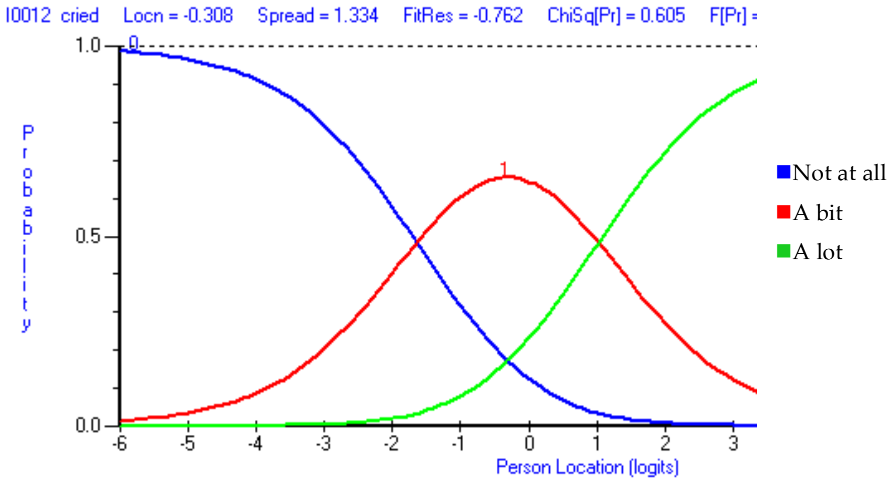

Figure 7.3. Example of ordered response categories.

Each response category ("Not at all", "a bit" or "a lot") systematically has a point along the continuum where it is the most likely response (i.e. the peak).

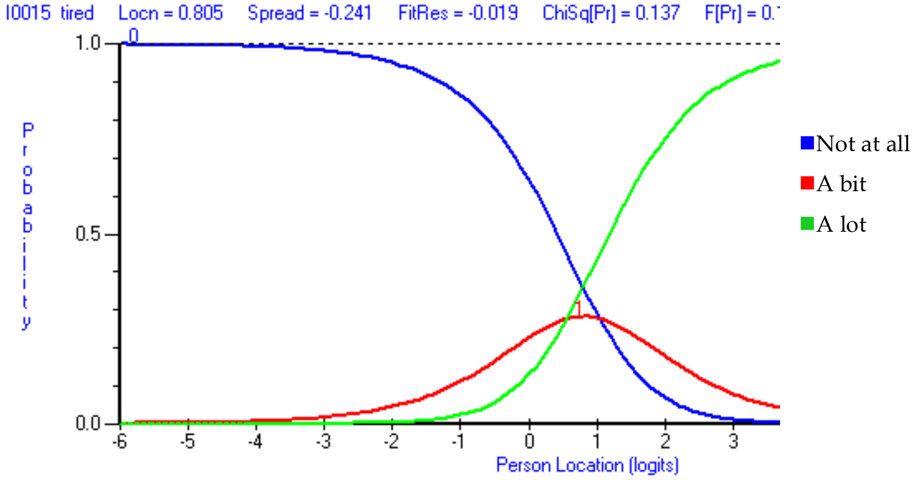

Figure 7.4. Example of disordered response categories.

Responses for the second option ("a bit") is never the most likely to be endorsed at any point along the underlying trait.

To overcome this, adjacent categories may be collapsed to reduce the response options or the question removed. Where disordered thresholds were detected, these questions were removed, as there were only three response options and therefore collapsing options was not appropriate. Local dependency was deemed to be present if residual correlations were greater than 0.2 above the average residual correlation (Kersten et al., 2014). DIF was checked for by age (5–7 years, 8–11 years, 12–16 years), gender, ethnicity (white British and non-white British) and deprivation. For analysis, three equal groups were created using the IMD scores of the sample (lower one-third of scores, middle third of scores, higher third of scores).

If the data fit the Rasch model, each item and person fit residual should be within the range +/- 2.5 and the mean item and person fit statistics should be close to zero with a standard deviation of one (Kersten et al., 2014). Finally, the individual items and summary chi-square interaction statistics should be non-significant (>0.05), although these are subject to Bonferroni adjustment based on the number of items. Strict unidimensionality was then examined using an independent t-test on two subsets of items (those which load positively and those which load negatively) identified using principal component analysis of the items residuals. If the 95% confidence interval of these t-tests includes 5%, then unidimensionality is confirmed.

Reliability can also be assessed using the Person Separation Index (PSI). This is equivalent to Cronbach's alpha, however the logit value is used instead of the raw score (Tennant and Conaghan, 2007). It is interpreted in the same manner, i.e. a value of 0.7 is recommended.

As most OHRQoL measures are ordinal in nature, this precludes simple addition and subtraction of raw scores (Svensson, 2001). As CARIES-QC focuses on attributes which are not directly measurable, such as pain and emotional impacts, the raw score will only be indicative of a rank along the scale (Tennant et al., 2004). In order, to use the raw score to accurately measure change, conversion to an interval level scale is required. This can be achieved by transforming the ordinal score to a logit score (Tennant et al., 2004). Following fit to the model, a transformation from raw score to interval data was undertaken. All further analyses, where appropriate, were based on the scale created from this analysis.

Interpretability

Floor and ceiling effects were defined to be present if more than 15% of participants reported the best or worst possible score (Terwee et al., 2007). The mean, range and standard deviation (SD) of scores were calculated for all subgroups (gender, age, ethnicity and deprivation group). Independent t-tests and one-way ANOVA were used to test for differences between the transformed interval CARIES-QC score (CIS) according to clinical and demographic variables as this data was normally distributed. Mann-Whitney U Tests and Kruskal-Wallis tests were used to identify differences between the clinical and demographic subgroups as the total CPQ_{11-14}-ISF:16 score was not normally distributed.

Internal consistency, reliability and construct validity

In addition to the Rasch analysis, to allow comparison with similar scales, Cronbach's alpha was calculated for both CARIES-QC and CPQ_{11-14}-ISF:16. Cronbach's alpha >0.7 is recommended to indicate a homogenous scale (Terwee et al., 2007).

ICC was calculated for all participants who remained stable between T0 and T1. A value >0.7 is deemed acceptable (Terwee et al., 2007).

A one-way ANOVA was performed to check for order effect (between Questionnaire A and B). Construct validity was tested using the appropriate bivariate correlations between CARIES-QC total score and: CPQ_{11-14}-ISF:16 total score; the presence of pain; pulpal involvement; anterior caries; total number of carious teeth; total caries experience and the global scores of both CARIES-QC and CPQ_{11-14}-ISF:16. It was hypothesised that there would be positive correlations with CPQ_{11-14}-ISF:16 (especially the oral symptoms, functional and emotional wellbeing domains) and the global questions from both CARIES-QC and CPQ_{11-14}-ISF:16. It was also hypothesised that there would also be positive correlations between CARIES-QC total score and clinical data (the total number of carious teeth, the presence of pain and pulpal involvement). These hypotheses were based on the information obtained from the qualitative data described in Chapter Six.

Responsiveness

This was analysed using correlations between the mean change score at T2 and the global change score reported by the participant (-1=worse, 0=same, +1=better). It was hypothesised that change score (difference between T0 and T2) would correlate with the global score, with those feeling there had been an improvement in their condition having a lower total score than those who felt they had stayed the same or felt worse. It was also hypothesised that those who felt that they had improved would have lower mean scores than those who felt their dental condition had deteriorated or remained unchanged following treatment.

7.4 Results

Of those who were approached to participate, only one child declined (response rate 99%). A total of 202 participants were recruited between July 2014 and January 2015. Two (1%) participants did not complete CARIES-QC (the pages containing CARIES-QC were omitted) and therefore were eliminated from further analysis. There were 95 (47.5%) males and 105 (52.5%) females with a mean (range) age of 8.01 (5.01–16.02) years. The majority of children were reported to be white British (65%, n=130). IMD scores ranged from 5.01-75.31 (mean=37.83). The IMD rank scores ranged from 81-29897 (mean = 8015), with 59.5% (n=119) of participants living in the most deprived quintile according to national IMD ranking scores (McLennan et al., 2011). Further demographic information is shown in Table 7.1.

Twenty-two (11%) children were in the permanent dentition, 72 (36%) in the primary dentition and 106 (53%) were in the mixed dentition. The mean (range) dmft was 6.22 (0-16), with two (1%) children having no caries in their primary teeth. The mean (range) DMFT was 1.6 (0-13) with 57 (46%) children having no caries in their permanent teeth. Further details of caries experience are shown in Table 7.2. Anterior caries was present in 41 (20.5%) of the participants. Pulpal involvement was present in 160 (80%) and pain reported in 145 (72.8%) of participants.

Table 7.1. Demographic characteristics of the participants (n=200).

Variable	Proportion	Number
Gender		
Male	47.5%	95
Female	52.5%	105
Ethnicity		
Asian background	15.5%	31
Black background	2.5%	5
Mixed background	4.5%	9
White British background	65.0%	130
Other background	4.5%	9
Unknown background	8.0%	16
Socioeconomic status		
Most deprived	59.5%	119
More deprived	18.5%	37
Average	10.0%	20
Less deprived	6.5%	13
Least deprived	5.5%	11

Table 7.2. Caries experience of included participants (n = 200).

	Minimum	Maximum	Mean (SD)
dmft	0	16	6.24 (3.45)
Number of carious primary teeth	0	14	5.74 (3.40)
Number of missing primary teeth	0	10	0.27 (1.15)
Number of filled primary teeth	0	4	0.22 (0.69)
DMFT	0	13	1.57 (2.18)
Number of carious permanent teeth	0	9	1.38 (1.77)
Number of missing permanent teeth	0	4	0.05 (0.38)
Number of filled permanent teeth	0	7	0.13 (0.74)
Total number of carious teeth	1	14	6.01 (3.27)
Total number of missing teeth	0	10	0.27 (1.12)
Total number of filled teeth	0	7	0.28 (0.89)

SD = standard deviation
dmft = total number of decayed, missing and filled primary teeth
DMFT = total number of decayed, missing and filled permanent teeth

7.4.1 Missing data

At baseline, a total of 11 (5.5%) participants had missing values for CARIES-QC. Three (1.5%) participants had greater than two missing values in CARIES-QC and were eliminated from the following analyses with the exception of the Rasch analysis. Seven participants had one missing value and one had two missing values in CARIES-QC. There was no discernable pattern to these missing values. Missing values for these eight participants were replaced with the participant's overall mean value. One participant omitted the global question, this was not replaced by a mean value and therefore correlations with the global score excluded this participant. There were no missing values at T1 and only one missing value at T2. The missing value at T2 was replaced with the participant's overall mean value.

A total of 8.5% (n=17) of participants had missing values for CPQ_{11-14}-ISF:16. Nine (4.5%) participants had greater than two missing values in CPQ_{11-14}-ISF:16, one had two missing values and seven one missing value. Of these, two had omitted the global question regarding the effect on life overall. These values were not replaced and therefore not included in correlations with total score. Mean participant values were used to replace the missing data, with the exception of the global questions. Those with more than two missing responses generally related to entire pages being omitted. Of those where isolated values were absent, three participants had not answered the question relating to having "sores". No other items had more than two missing values.

It was noted that on both questionnaires, participants had occasionally supplied two responses to a question. Where one of these had not been indicated as an error by the participant, these values were assigned as missing.

7.4.2 Rasch analysis

Two hundred participants were included in the Rasch analysis. One item ("feeling tired") had a disordered threshold and three items demonstrated misfit to the model ("taking medicine", "front teeth looking brown or black" and "being able to see holes"). A residual correlation > 0.13 was seen between the items “being kept awake” and “interfering with schoolwork”. However, removal of either of these items did not improve the overall fit. No DIF was observed. Seven participants did not fit the Rasch model and were therefore removed. This resulted in acceptable fit statistics (Table 7.3). Table 7.4 shows the item fit statistics for the 12 retained items which are ordered from ″easiest″ (“food stuck”) to “most difficult” (“interfering with schoolwork”). The mean person location is -1.12 when the items are centred on zero. This demonstrates that the scale is targeted to a population with slightly more impacts than the participants, which may be due to the number of participants who reported no or low levels of impacts. Figure 7.5 shows the person-item threshold map which shows that participants are distributed in a similar pattern to the items, indicating that the items measure the impacts of caries along the construct from least to most. As the items fit the Rasch model, a transformation from the raw score to interval scaling is shown in Table 7.5.

Table 7.3. CARIES-QC fit to the Rasch model.

Analysis name	Item residual		Person residual		Chi-square		Reliability	Unidimensionality	
	Mean	**SD**	**Mean**	**SD**	**Value (df)**	**P**		**Percentage of tests > 5%**	**95% CI**
Initial analysis	-0.39	1.62	-0.23	1.04	112 (32)	0	0.86	4%	1.96 – 14.04
Remove item “tired”	-0.41	1.66	-0.21	1.03	92 (30)	0	0.85	4.5%	3 – 15.04
Remove items “holes”, “brown” and “medicine”	-0.04	0.90	-0.25	1.00	39 (24)	0.03	0.85	4%	2 – 14.04
Remove 7 participants with misfit	-0.37	0.86	-0.21	0.89	38 (24)	0.04	0.85	4.15%	2.07 – 13.93
Ideal	**0**	**1**	**0**	**1**		**>0.004***	**>0.7**	**<5%**	**LCI <5**

*Bonferroni adjusted for 12 items; SD=standard deviation; df=degrees of freedom; CI=confidence interval; LCI=Lower confidence interval.

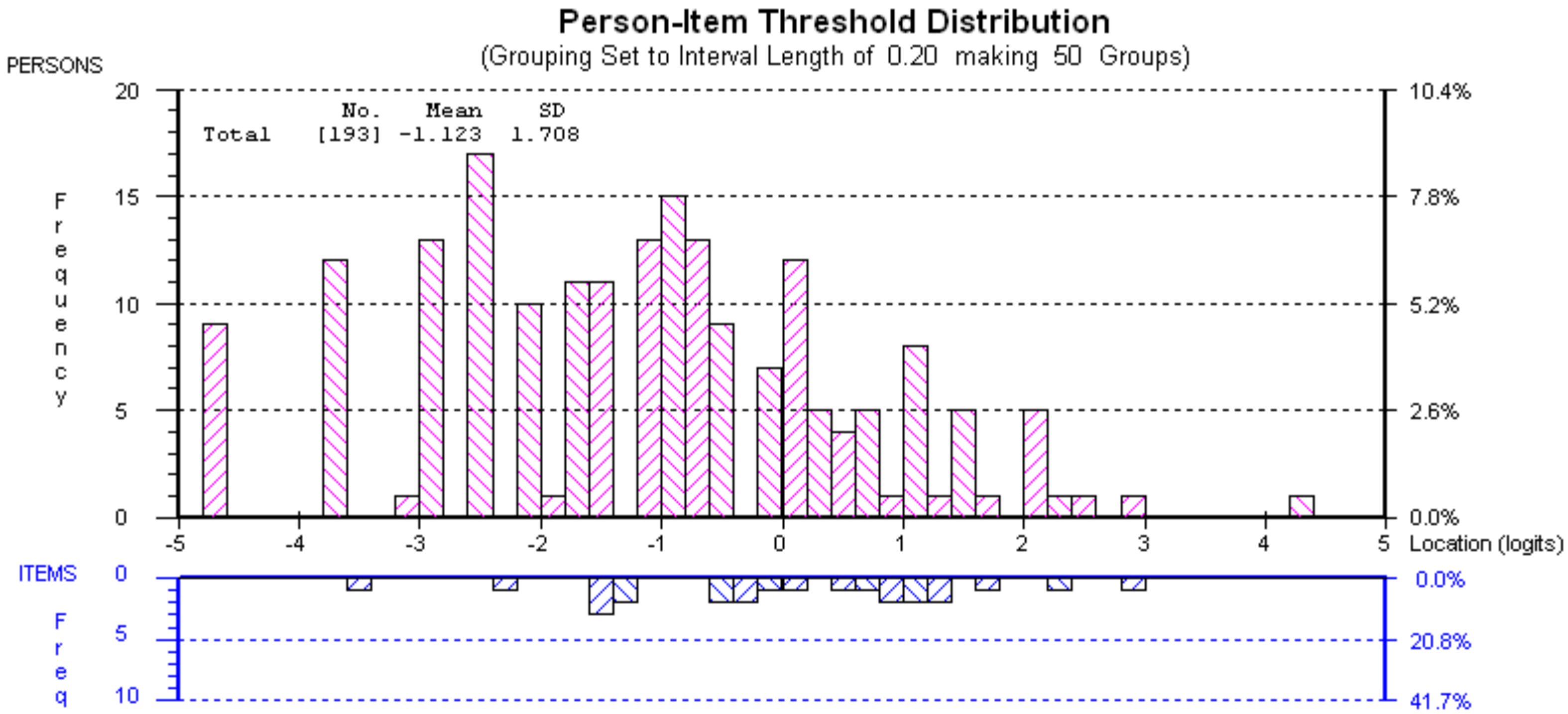

Figure 7.5. Targeting of CARIES-QC.

The upper section of the graph shows the distribution of participants and the lower part the distributions of thresholds (category transitions) of the items. The x-axes display the location (severity of impact) of the participants and the item location (difficulty) of the item thresholds. The y-axes show the frequency of item thresholds and participants.

Table 7.4. Fit of the 12 items of CARIES-QC to the Rasch model in location order.

Item	Location	Standard error	Fit residual	Degrees of freedom	Chi-square	Probability
Food stuck	-1.88	0.14	0.57	162.04	5.62	0.060
Eating on one side	-0.79	0.13	-0.75	161.14	0.61	0.738
Hurts	-0.69	0.14	-0.70	162.95	5.03	0.081
Annoyed	-0.45	0.13	-1.69	162.04	8.42	0.015
Eating carefully	-0.34	0.13	-1.68	164.76	4.44	0.108
Cried	-0.22	0.14	1.17	164.76	4.68	0.097
Difficult to eat	0.16	0.15	0.11	162.04	1.05	0.593
Cross	0.33	0.14	0.08	164.76	2.80	0.247
Eating slowly	0.35	0.14	-0.85	163.85	1.83	0.400
Brushing teeth	0.39	0.14	0.35	163.85	1.42	0.492
Kept awake	1.15	0.16	-0.48	162.95	0.54	0.762
Interfering with schoolwork	2.00	0.19	-0.61	163.85	1.45	0.485
Ideal			**+/- 2.5**			**>0.004***

*Bonferroni adjustment based on 12 items.

Table 7.5. Transformation of raw (ordinal) score to interval score.

Raw score	Interval score	Raw score	Interval score
0	0	**13**	13.03
1	2.63	**14**	13.62
2	4.50	**15**	14.22
3	5.84	**16**	14.84
4	6.90	**17**	15.48
5	7.80	**18**	16.17
6	8.60	**19**	16.92
7	9.32	**20**	17.76
8	10.00	**21**	18.75
9	10.64	**22**	19.96
10	11.26	**23**	21.65
11	11.86	**24**	24.00
12	12.45		

7.4.3 Interpretability

One hundred and ninety-seven participants were included in the analysis of CARIES-QC, three were excluded as they had omitted more than two questions. One hundred and eighty-eight participants were included in the analysis of CPQ$_{11\text{-}14}$-ISF:16 as nine of the remaining participants had omitted more than two questions.

CARIES-QC

At baseline, nine (4.6%) participants scored the lowest possible score (possible range 0–24) and one (0.5%) the highest, which is within an acceptable range for floor and ceiling effects. The mean (range) raw score was 8.08 (0–24) and the CIS mean score was 9.34. Mean CIS for the different subgroups are shown in Table 7.6. Statistically significant higher mean scores were found for children who reported pain (p=0.00) compared those who had not. This was also true for those who had anterior caries (p=0.03) when compared with those without anterior caries and those who had pulpal involvement (p=0.03) when compared with those without pulpal involvement. A statistically significant difference was found between ethnic groups (p=0.03), with those from non-white British backgrounds having higher scores than those from white British backgrounds. This was despite there being no significant difference between the number of carious teeth, the presence of pain or pulpal involvement between subgroups.

At T1 (n=70), the mean (range) raw score was 6.34 (0–20) with three (4.5%) participants scoring the lowest possible score. The CIS mean (range) score was 8.00 (0–17.76). Where participants reported that they had experienced no change (n=30) since T0, the mean (range) CIS was 8.03 (2.63–16.17). This was contrasted with a lower mean (range) CIS of 7.08 (0–17.76) in those who reported that their teeth felt better and a higher mean (range) CIS of 12.14 (8.6–14.84) in those who felt their teeth had deteriorated (n=7).

Table 7.6. Mean, range and standard deviation of transformed interval level CARIES-QC baseline scores.

Participants/subgroups		CARIES-QC interval score		
Participants	**Number (%)**	**Mean score**	**Range**	**SD**
Overall	197	9.34	0 - 24	4.44
Gender				
Female	103 (52.3%)	9.71	0 - 24	4.59
Male	94 (47.7%)	9.05	0 – 18.75	4.28
Age group				
5 – 7 years	114 (57.9%)	9.82	0 - 24	4.51
8 – 11 years	66 (33.5%)	8.91	0 – 18.75	4.55
12 – 16 years	17 (8.6%)	8.90	2.63 – 16.17	3.49
Pain				
Yes	142 (72.1%)	10.49a**	0 - 24	4.09
No	55 (27.9%)	6.57	0 – 17.76	4.10
Pulpal involvement				
Yes	157 (79.7%)	9.86b*	0 - 24	4.26
No	40 (20.3%)	7.56	0 – 17.76	4.72
Anterior caries				
Yes	41 (20.8%)	10.72c*	0 – 24	4.71
No	156 (79.2%)	9.04	0 – 19.96	4.32
Ethnicity				
White British background	130 (66.0%)	8.91	0 – 18.75	4.29
Other background	53 (34.0%)	10.57d*	0 - 24	4.82
Deprivation				
Deprivation group 1 (lowest)	67 (34.0%)	9.20	0 – 24	4.92
Deprivation group 2 (middle)	65 (33.0%)	9.93	0 – 19.96	4.35
Deprivation group 3 (highest)	65 (33.0%)	9.04	0 – 17.76	4.02

SD=standard deviation; a**=children who reported pain had significantly higher mean CARIES-QC interval score than those who did not report pain ($p<0.01$); b*=children who had pulpal involvement had significantly higher mean CARIES-QC interval score than those who had no pulpal involvement (p<0.05); c*=children with anterior caries had significantly higher mean CARIES-QC interval score than those who did not have anterior caries (p<0.05); d*=children from non-white British backgrounds had significantly higher mean CARIES-QC interval score than those from white British backgrounds (p<0.05).

The impact which was most commonly reported by participants at baseline was “food stuck” (n=173, 87.5%) and the least reported impact was “interfering with schoolwork” (n=34, 17.3%). The majority of items were reported by greater than 50% of participants. Further details are displayed in Table 7.7. Interestingly, there were 62 (31.5%) children who answered that they had experienced no pain. However, 53 (26.9%) of these children reported other impacts related to their caries experience with CARIES-QC total scores ranging from 1–11 (mean=2.97). This compared with scores of 1–24 (mean=10.24) for those answering that they either had “a bit” or “a lot” of pain. A statistically significant difference in mean CIS was found between the pain and non-pain subgroups (p=0.00, independent t-test). At baseline 64.5% (n=127/196) reported that their teeth were “a bit” or “a lot” of a problem with respect to the global question.

Table 7.7. Number and proportion of participants responding positively (“a bit” or “a lot”) to each item at baseline (n=197).

Item	Number with impact	Proportion with impact
Food stuck	173	87.5%
Hurts	135	68.5%
Eating on one side	123	62.4%
Cried	121	61.4%
Annoyed	118	59.9%
Eating carefully	111	56.3%
Difficult to eat some foods	111	56.3%
Eating slowly	87	44.2%
Brushing teeth	84	42.6%
Feeling cross	80	40.6%
Kept awake	64	32.5%
Interfering with schoolwork	34	17.3%

CPQ_{11-14}-ISF:16

One hundred and eighty-eight participants were included in the analysis of data derived from CPQ_{11-14}-ISF:16. Three (1.5%) participants scored the lowest possible score (possible range 0–64) and there were no participants who scored the highest score, which is within an acceptable range for floor and ceiling effects. The mean (range) raw score was 15.06 (0–44). There were statistically significant differences between the mean CPQ_{11-14}-ISF:16 score in participants who reported pain versus those who had not reported pain and between those who had pulpal involvement and those who did not have any pulpal involvement ($p<0.05$ Mann-Whitney U test). Mean scores for the different subgroups are shown in Table 7.8.

The most frequently reported impact in CPQ_{11-14}-ISF:16 was "food stuck" (n=167, 88.8%) and least frequently experienced was "difficulty saying words" (n=30, 16%). Further detail regarding the impacts reported using CPQ_{11-14}-ISF:16 is shown in Table 7.9.

Table 7.8. Mean, range and standard deviation CPQ$_{11-14}$-ISF:16 scores.

		CPQ$_{11-14}$-ISF:16 score		
Participants	**Number (%)**	**Mean**	**Range**	**SD**
Overall	188	15.06	0 - 44	9.50
Gender				
Female	99 (52.7%)	14.79	0 - 44	9.72
Male	89 (47.3%)	14.79	2 – 40	9.29
Age group				
5 – 7 years	108 (57.4%)	15.05	2 – 44	9.18
8 – 11 years	63 (33.5%)	14.90	0 – 39	9.94
12 – 16 years	17 (9.0%)	15.76	4 – 38	10.31
Pain				
Yes	136 (72.3%)	16.64a*	0 – 40	9.13
No	52 (27.7%)	10.94	0 – 44	9.28
Pulpal involvement				
Yes	148 (78.7%)	15.82b*	0 – 40	9.36
No	40 (21.3%)	12.25	0 – 44	9.59
Anterior caries				
Yes	40 (21.3%)	15.30	4 – 31	7.7
No	148 (78.7%)	15.00	0 – 44	9.95
Ethnicity (n=175)				
White British background	124 (71.0%)	14.06	0 – 44	8.94
Other background	51 (29.0%)	17.22	0 – 38	10.37
Deprivation				
Deprivation group 1 (lowest)	64 (34.0%)	15.75	0 – 44	11.13
Deprivation group 2 (middle)	61 (32.4%)	15.74	0 – 36	8.65
Deprivation group 3 (highest)	63 (33.5%)	13.71	2 - 38	8.43

SD=standard deviation; a*=children who reported pain had higher mean scores when compared to children who did not report pain ($p<0.05$ Mann Whitney U test);

b*=children who had pulpal involvement had higher mean scores when compared to children who did not have pulpal involvement (p<0.05 Mann Whitney U test).

Table 7.9. Proportion of children who responded positively (once or twice, sometimes, often, everyday or almost everyday) to items in CPQ$_{11\text{-}14}$-ISF:16.

Item	Proportion (number) with impact (n=188)
Food stuck	88.8% (167)
Pain	82.4% (155)
Taken longer to eat	64.9%(122)
Difficult to bite foods	60.1% (113)
Bad breath	56.9% (107)
Sores	54.8% (103)
Felt upset	54.3% (102)
Feeling irritable	53.2% (100)
Pain with hot and cold	48.9%(92)
Felt shy	36.7% (69)
Worried about what others think	36.2% (68)
Argued with family	30.3% (57)
Been asked questions	21.8% (41)
Avoided smiling/laughing	20.2% (38)
Been teased	18.1% (34)
Difficulty saying words	16.0% (30)

7.4.4 Internal consistency, reliability and construct validity of measures

CARIES-QC

Participants (n=5) who had omitted more than two answers from CARIES-QC were excluded from these analyses, therefore a sample of 197 was available for analysis. Cronbach's alpha was found to be 0.9. This did not increase if any items were deleted, as would be expected following the Rasch analysis. Item total correlations ranged from 0.509 (food stuck) to 0.719 (eating carefully).

A total of 65 participants participated in the test-retest analysis at a mean (range) of 29 (3–127) days. Thirty participants reported that their condition had remained stable. Those participants who remained stable had similar characteristics at baseline to the participants at T0 (Table 7.10). Their mean (range) age was 8.21 (5.61-13.72) years and 48.4% were male. The mean (range) dmft was 6.32 (1-13) and mean (range) DMFT was 0.90 (0-4). In terms of clinical presentation 22.6% (n= 7) had anterior caries, 71% (n=22) had pulpal involvement and 67.7% (n=21) had reported pain. The ICC was 0.68, which is slightly lower than the desired value of 0.7 (Terwee et al., 2007).

Table 7.10. Characteristics at presentation for participants involved in each stage.

	Baseline	Test-retest	Follow-up
Mean (range) age in years	8.01 (5.01-16.02)	8.21 (5.61-13.72)	7.9 (5.33-12.56)
Mean (range) dmft	6.22 (0-16)	6.32 (1-13)	6.26 (1-12)
Mean (range) DMFT	1.6 (0-13)	0.90 (0-4)	1.09 (0-4)
Proportion (number) with anterior caries	20.5% (41)	22.6% (7)	19.0% (8)
Proportion (number) with pulpal involvement	80% (160)	71.0% (22)	88.1% (37)
Proportion (number) who had reported pain	72.8% (145)	67.7% (21)	76.2% (32)

No order effect was present (p=0.732). A strong correlation was found between CARIES-QC total and the global question (r=0.734) (p<0.01). Weaker but significant (p<0.01) correlations were found between the total CARIES-QC score and pain (r=0.392), the total number of carious teeth (r=0.188) and with the presence of pulpal involvement (r=0.187). Further details are shown in Table 7.11. No significant correlation was found between CARIES-QC total score and dmft or DMFT. Significant (p < 0.01) correlations were also found between CARIES-QC total score and the total CPQ_{11-14}-ISF:16 score (r=0.733), the global oral health rating from CPQ_{11-14}-ISF:16 (r=0.291) and life overall rating from CPQ_{11-14}-ISF:16 (r=0.392). Statistically significant correlations were found with CARIES-QC total score and all domains of CPQ_{11-14}-ISF:16 with the highest correlation being between the oral symptoms (r=0.646), functional limitations (r=0.665) and emotional wellbeing (r=0.630) domains.

As data from seven participants did not fit the Rasch model, these were removed to ascertain if there was an effect on these estimates of reliability and validity. Removing these participants' data resulted in Cronbach's alpha remaining stable at 0.9. The ICC based on 26 participants was 0.69. Correlations with clinical data, global scores and CPQ_{11-14}-ISF:16 were similar but no further statistically significant correlations were discovered (Table 7.11). It can therefore be seen that while these aberrant cases altered the correlation values, they did not affect the overall psychometric properties.

CPQ_{11-14}-ISF:16

A total of 10 participants had not answered more than two items from CPQ_{11-14}-ISF:16, however one of these had already been eliminated as they had also not answered more than two items from CARIES-QC, therefore a sample of 188 was available for analysis. Cronbach's alpha was found to be 0.84 and this was shown to increase if the item "argued with family" was removed. Inter-item correlations ranged from 0.248 (argued) to 0.620 (irritable). Correlations with clinical data are shown in Table 7.11. There were also statistically significant ($p<0.01$) correlations with the global ratings of oral health ($r=0.397$), global rating of life overall ($r=0.403$) and CARIES-QC global question ($r=0.612$).

Table 7.11. Correlations (Spearman's rho) between CARIES-QC total scores and clinical data and CPQ$_{11-14}$-ISF:16 total scores and clinical data.

Clinical variable	CARIES-QC total score (n = 197)	Transformed interval CARIES-QC total score (n = 197)	CARIES-QC total score without aberrant cases (n = 190)	CPQ$_{11-14}$-ISF:16 total score (n = 188)
Total number of carious teeth	0.188**	0.188**	0.198**	0.147*
Total dmft/DMFT	0.129	0.129	0.140	0.121
Pain	0.392**	0.392**	0.395**	0.312**
Pulpal involvement	0.187**	0.187**	0.188**	0.178*
Anterior caries	0.154*	0.147*	0.154*	0.048

*p<0.05; **p<0.01; dmft = total number of decayed, missing and filled primary teeth; DMFT = total number of decayed, missing and filled permanent teeth.

7.4.5 Responsiveness

Data from 42 participants were available following a comprehensive course of treatment (T2). These participants had a similar profile (Table 7.10), although a higher proportion had pulpal involvement at baseline to those included at T0. Their mean (range) age was 7.90 (5.33-12.56) and 40.5% were male. The mean (range) dmft was 6.26 (1-12) and mean (range) DMFT was 1.09 (0-4). In terms of clinical presentation, 19% (n=8) had anterior caries, 88.1% (n=37) had pulpal involvement and 76.2% (32) had reported pain. All but two of the children had no remaining carious teeth at T2.

Of the children who participated at T2, 33 (78.6%) reported an improvement since baseline, 6 (14.3%) reported no change and 3 (7.1%) reported a deterioration in their oral condition. The mean (range) raw score was 4.17 (0–17), with six participants (14.3%) scoring the lowest possible score. Extraction of teeth under GA (XGA) was the most common treatment provided (n=22, 52.4%), followed by the provision of Hall technique PMCs (provided prior to the GA) in combination with XGA (n=9, 21.4%). A total of 31 (73.8%) participants had some of their treatment provided under GA. Details of the treatment provided are shown in Table 7.12. The mean (range) time between the final treatment session and T2 and between T0 and T2 was 62.6 (0–177) days and 134.7 (48–261) days respectively.

Impacts were reported in relation to all items following treatment. The most frequently reported impact at T2 was "food stuck" (n=26, 61.9%), followed by "hurts", "hard to eat some foods" and "eating carefully" (n=16, 38.1%). Responses of "a bit" or "a lot" to the global CARIES-QC question accounted for 28.6% (n=12) answers. With the exception of the item "interfering with schoolwork", the number of children who reported experience of each impact decreased between baseline and follow-up in those who reported an improvement. For those who reported improvement, the item "annoyed" was associated with the largest reduction in the proportion of participants experiencing it between baseline and follow-up (60.6%), followed by "hurts"

(45.5%) and "eating on one side" (42.5%). Comparisons with baseline data are shown in Table 7.13 and Table 7.14.

Table 7.12. Treatment provided to participants.

Treatment provided	Number	Proportion
GA extractions	21	50.0%
Hall crowns (placed prior to GA) and GA extractions	9	21.4%
IS extractions	3	7.1%
Hall crowns	2	4.8%
Hall crowns with symptomless carious primary teeth left to exfoliate	2	4.8%
LA restoration	1	2.4%
IS extraction and restoration	1	2.4%
IS extraction and Hall crowns	1	2.4%
IS pulpotomy/PMC and Hall crown	1	2.4%
Restorations and extractions under GA	1	2.4%

GA = General anaesthetic; IS = Inhalation sedation with local anaesthesia; LA = Local anaesthesia.

The mean (range) raw score for those who reported an improvement (n= 33) was 2.94 (0–14) compared to 6.67 (2–12) for those who reported no change and 12.67 (9–17) for those who felt they were worse. The mean difference in CIS between baseline and follow-up for those who felt they had improved was minus 4.46 (range=minus 12.45–plus 2.76), thus indicating a MID of 4.46 points. There was a statistically significant difference between the mean score at baseline and follow-up in those who reported an improvement (p=0.00, paired t-test). A strong statistically significant correlation (p=0.01) was found between the global rating of change and the CIS change score (r=0.601, Pearson correlation). Comparisons between the groups can be seen in Table 7.15.

Table 7.13. Number and proportion of participants responding positively ("a bit" or "a lot") to each item in CARIES-QC following treatment at baseline and follow-up.

Item and response	Proportion (number) with impact at baseline (n = 197)	Proportion (number) with impact at follow-up (n = 42)
Food stuck	**87.8% (173)**	**61.9% (26)**
A bit	55.3% (109)	50.0% (21)
A lot	32.5% (64)	11.9% (5)
Hurts	**68.5% (135)**	**38.1% (16)**
A bit	51.3% (101)	35.7% (15)
A lot	17.3% (34)	**2.4% (1)**
Eating on one side	**62.4% (123)**	**33.3% (14)**
A bit	36.5% (72)	16.7% (7)
A lot	25.9% (51)	16.7% (7)
Cried	**61.4% (121)**	**31.0% (13)**
A bit	47.7% (94)	26.2% (11)
A lot	13.7% (27)	4.8% (2)
Annoyed	**59.9% (118)**	**19.0% (8)**
A bit	40.6% (80)	16.7% (7)
A lot	19.3% (38)	**2.4% (1)**
Eating carefully	**56.3% (111)**	**38.1% (16)**
A bit	38.1% (75)	31.0% (13)
A lot	18.3% (36)	7.1% (3)
Difficult to eat some foods	**56.3% (111)**	**38.1% (16)**
A bit	46.7% (92)	35.7% (15)
A lot	9.6% (19)	2.4% (1)
Eating slowly	**44.2% (87)**	**28.6% (12)**
A bit	33.5% (66)	19.0% (8)
A lot	10.7% (21)	9.5% (4)
Brushing teeth	**42.6% (84)**	**19.0% (8)**
A bit	32.0% (63)	19.0% (8)
A lot	10.7% (21)	0
Feeling cross	**40.6% (80)**	**14.3% (6)**
A bit	28.4% (56)	11.9% (5)
A lot	12.2% (24)	2.4% (1)
Kept awake	**32.5% (64)**	**11.9% (5)**
A bit	27.4% (54)	9.5% (4)
A lot	5.1% (10)	2.4% (1)
Interfering with schoolwork	**17.3% (34)**	**11.9% (5)**
A bit	14.7% (29)	7.1% (3)
A lot	2.5% (5)	4.8% (2)
Global question	**64.5% (127)**	**28.6% (12)**
A bit	49.0% (96)	26.2% (11)
A lot	15.7% (31)	2.4% (1)

Table 7.14. Number and proportion of participants who reported an improvement responding positively ("a bit" or "a lot") to each item following treatment at baseline and follow-up (n = 33).

Item	Proportion (number) with impact at baseline	Proportion (number) with impact at follow-up	Difference in proportion (number) between baseline and follow-up
Food stuck	90.9% (30)	63.6% (21)	27.3% (9)
Annoyed	72.7% (24)	12.1% (4)	60.6% (20)
Cried	63.6% (21)	24.2% (8)	39.4% (13)
Hurts	69.7% (23)	24.2% (8)	45.5% (15)
Eating on one side	66.7% (22)	24.2% (8)	42.5% (14)
Difficult to eat some foods	57.6% (19)	30.3% (10)	27.3% (9)
Eating carefully	51.5% (17)	33.3% (11)	18.2% (6)
Brushing teeth	45.5% (15)	9.1% (3)	36.4% (12)
Feeling cross	39.4% (13)	6.1% (2)	33.3% (11)
Eating slowly	39.4% (13)	18.2% (6)	21.2% (7)
Kept awake	21.2% (7)	3.0% (1)	18.2% (6)
Interfering with schoolwork	6.1% (2)	6.1% (2)	0% (0)
Global question	62.5% (20/32)	18.2% (6)	44.3% (14)

Further analysis of those participants who reported that they had improved, but whose scores did not show improvement, revealed that two participants' scores had stayed the same between baseline and follow-up and two participants' scores had increased. The two participants who had increased scores, had undergone their treatment (extractions under GA) less than one month previously, which may account for why they were still experiencing impacts.

Table 7.15. Mean (range) and change scores calculated using CARIES-QC interval scores (n=42).

Reported condition at follow-up	Mean (range) CARIES-QC interval score at baseline	Mean (range) CARIES-QC interval score at follow-up	Mean (range) change score
All follow-up participants (n=42)	9.48 (2.63 – 19.96)	5.99 (0– 5.48)	-3.48 (minus 12.45–4.10)
Improved (n=33)	9.33 (2.63 – 19.96)	4.87 (0–13.62)	-4.46 (minus 12.45–2.76)
Unchanged (n=6)	8.32 (4.5 – 12.45)	8.74 (4.5 – 12.45)	0.42 (minus 1.52–4.10)
Deteriorated (n=3)	13.44 (11.86 -14.84)	12.86 (10.64 – 15.48)	-0.58 (minus 1.22–0.64)

7.5 Discussion

The aim of this chapter was to describe how a unidimensional scale was derived and to evaluate this scale in terms of its reliability, validity and responsiveness. The combination of modern psychometric techniques and classical test theory following the COSMIN guidelines have demonstrated that CARIES-QC has the potential to be a useful measure of the impacts experienced by children with dental caries.

7.5.1 Interpretability

Caries experience

As expected, this population had high levels of caries experience. The NHS DEP 2011/12 found that the mean dmft in 5-year-olds in Sheffield was 1.30 and the mean dmft for Sheffield children who had caries experience was 3.62 (Public Health England, 2012). This is 42% lower than the mean dmft of 6.24 for children in the present study. The mean DMFT in this study was 1.57, compared with 0.90 in Sheffield in the NHS DEP 2008/2009. However, of the children in Sheffield who had caries experience, the mean DMFT was 2.35 which is higher than the mean DMFT of 1.57 in this study (Rooney et al., 2010).

This may reflect the case-mix of children referred to a specialist service. Young children with extensive caries in primary teeth are often referred to specialist services for treatment under GA, whereas it would be expected that the majority of children with caries in permanent teeth would be treated within primary care.

Deprivation and ethnicity

The majority of participants (59.9%) in this study were from the most deprived areas in England according to national statistics (McLennan et al., 2011). This is disproportionately higher than the demography of Sheffield, where around 35% reside in the most deprived areas of England (Public Health England, 2014e). Around 35% of the participants were from black and ethnic minority groups, which slightly exceeds the figure reported for school-aged children in Sheffield (29%) (Public Health England, 2014a). Previous studies have shown that dental caries is unequally distributed with higher levels found in those living in deprived areas and those from ethnic minority groups (Lader et al., 2004; Conway et al., 2007; Marcenes et al., 2013; Steele et al., 2015). This study also found that children from non-white British backgrounds had significantly higher mean CARIES-QC scores than those who were white British despite the fact that there was no significant difference between their clinical status and no questions displaying DIF. However, full regression analysis has not been undertaken and there may be other underlying factors such as social deprivation and the disparity of the sample sizes which may account for this finding. Some studies have found that OHRQoL and HRQoL may be affected by ethnicity and it would be useful to explore this association in the future (Quittner et al., 2010; Nanayakkara et al., 2013; Kumar et al., 2014).

Acceptability and missing values

The acceptability of the measures appeared to be good, with very few missing responses and all of the questionnaires returned (Fitzpatrick et al., 1998; Fayers and Machin, 2006). Acceptability was not formally assessed, although it was discussed with both parents and children informally where possible. Children reported that they enjoyed completing CARIES-QC and parents reported that the child had found it easy to complete. Another measure of acceptability is the time that it takes to complete. This was not assessed in this stage as it was completed away from the clinic. However, the time taken to complete was 2-5 minutes when measured during the development stage. Measures should be short and thus participants are able to be complete them relatively quickly to prevent fatigue. Therefore a completion time of 2-5 minutes would be assumed to be acceptable for children aged 5-16 years (Matza et al., 2004). A final measure of acceptability is that the measure uses language which is familiar to the participants (Patient Reported Outcomes Measurement Group, Unknown). By ensuring that the measure used the language adopted by children and refining this throughout the development stage, this criteria was fulfilled.

In the case of a missing response, the majority related to entire pages being omitted, rather than individual questions. It is difficult to know why pages in the booklet were missed, but as most participants completed the questionnaires while waiting for radiographs to be taken, it is speculated that they were interrupted whilst completing them. One way of reducing participant burden and missing data is through the use of Computer Adaptive Testing (CAT). CAT utilises software which allows the selection of questions which are appropriate to the participant's level of impact by using responses to the previous questions to provide a latent trait estimate (Wainer, 2000). While participants answer different questions, the items are calibrated, therefore allowing comparison between individuals (Haley et al., 2009). The items are calibrated using values obtained through Rasch analysis or other types of IRT. The use of this type of testing in HRQoL has shown that participant burden is reduced compared to the use of a static version of the same questionnaire (Allen et al., 2008; Coster et al., 2008; Chien et al., 2009). Recently, Kids-CAT, based on the generic HRQoL instrument, has shown that seven items were required to be administered to achieve measurement precision of 0.8-0.9 when compared to the static version (Devine et al., 2014). Thus it is possible to reduce the number of items whilst retaining the original measure's precision. As CARIES-QC has only 12 items, there may be little additional benefit to creating a CAT. However, future work could investigate this possibility. If the creation of a CAT is not possible, provision of

CARIES-QC on a tablet or similar electronic device may be advantageous, to increase efficiency and reduce errors (Vinney et al., 2012). This would have been of benefit, as some of the missing data in this study were related to children filling in two responses to some questions (i.e. ticking or circling multiple responses to the same question). While observing children completing the questionnaire, there were occasions when children circled an option from the preceding or following question by accident, as they lost their place while considering their answer. Development of the questionnaire in a computerised format would prevent this type of error.

Although there were few missing responses in CPQ$_{11-14}$-ISF:16, some parents reported that their children had struggled to understand the 5-point Likert scale. In particular, parents of young children and those with autism and learning difficulties felt that the question structure and response options were confusing for their children. The questions regarding "having sores" was the most frequently missed question. A number of children and parents reported that either they did not know what this meant or found it difficult to explain to their children. It would be useful in future for the response options to be analysed to evaluate whether there are any disordered thresholds in CPQ$_{11-14}$-ISF-16 and whether adjacent categories could be collapsed to simplify the scale.

7.5.2 Rasch analysis

The use of Rasch analysis produced a unidimensional set of items by removing items which did not fit the mathematical model. The first of these was the item "tired" which showed a disordered thresholds, implying that participants were unable to distinguish between their teeth causing "a bit" or " a lot" of tiredness. It may be that rewording of this question and its response options would remedy this. However, the aim was to produce a measure where the response options were consistent, to reduce participant burden and therefore changing the response options would be inconsistent with this approach. Thus, it was felt that removal of this item would be the preferred strategy as it had a high residual correlation (implying it was examining the same concept) with the item "kept awake".

The items regarding dental aesthetics were shown to be a poor fit to the model. This perhaps was not an unexpected finding, as two separate themes emerged during the item generation stage. One theme related to pain and its related impacts and the other theme was associated with aesthetics. The results of the Rasch analysis would appear to confirm that these are indeed separate constructs. The impacts related to the aesthetic impact of dental caries are not unimportant, however, the majority of treatment performed is to relieve or prevent pain and infection, not primarily to improve aesthetics. It was also noted during completion of the questionnaire that many participants who had anterior caries, did not endorse the aesthetic items. This may be because of the young age of the sample and therefore many of the children were not bothered by their appearance at this stage.

The other item that was removed related to having to take medicine. This was a question which was observed to have caused difficulty for the participants to complete. In many cases they had to ask their parents if this had occurred. It may be that they had been given analgesia for other conditions and were not able to remember why they had been given it.

The resulting scale showed good unidimensionality with only a few participants appearing to have completed the questionnaire in an unexpected way. These participants had a variety of demographic characteristics and clinical features and it was not clear why their answers were different to what was expected. In any sample it can be anticipated that some participants will answer differently, but as these were only a small proportion (3.5%) of the overall sample, it is not cause for concern. This is further highlighted by the fact that the psychometric properties were not significantly changed when these participants were removed from the analysis. The PSI 0.85 was within an acceptable range. PSI of 0.85 indicates that the measure is suitable for use at the individual level, that the reliability of the fit statistics produced is high and that statistically the measure can discriminate between three groups (in terms of severity of the impact) of patients (Fisher Jr, 1992).

It could be argued that using a statistical method to reduce the items in the measure is contrary to the inherent child-centred nature of this measure. However, it should be noted that Rasch analysis can be used to complement participant-generated information. The use of Rasch analysis allows individual items to be assessed in terms of their discriminative properties and whether they may be redundant as they are asking a similar question to that asked by another item (Tennant and Conaghan, 2007). Thus ensuring that the best items to examine the construct under scrutiny can be retained, and shortening the measure without losing important information. This level of detail cannot be obtained using CTT methods such as item impact analysis (Tennant and Conaghan, 2007). For example, the item regarding "difficulty doing schoolwork" was endorsed by few participants (17.3%), however, in the Rasch analysis it was shown that a participant who experienced this impact was one with a high level of impact and that it was associated with being "kept awake". The item "kept awake" was endorsed by less than one-third of children, but similarly was shown by the Rasch analysis to be a highly discriminative item indicating a high level of impact. Therefore, using CTT these items may have been eliminated but would have reduced the breadth of information obtained from the children's responses.

7.5.3 Internal consistency, reliability and construct validity

CARIES-QC had a high Cronbach's alpha of 0.9 indicating a high level of internal consistency. This would be expected following the creation of a unidimensional scale using Rasch analysis. The test-retest reliability (ICC=0.68) was a little lower than the ideal of >0.7 (Terwee et al., 2007). There may be several explanations for this.

1. *Small sample size (n=30)*

   Although 70 children participated in this stage of the study, only 30 stated that their dental condition had remained stable during this period. This is unsurprising given the sporadic nature of toothache. Many children (n=33) stated that their teeth already felt better at their second appointment, despite not having received any treatment. On some occasions this was attributed to the pain they had experienced subsiding, but others felt that their teeth had improved because they had been brushing them more frequently.

2. *Difficulty in interpretation of the question asking if their teeth felt "the same", "better" or "worse" than before*
   Some of the children found it difficult to answer this question. For children where there had been a definite increase or decrease in symptoms, there appeared to be no difficulty in answering. Whereas for those with lower level symptoms, it seemed difficult for them to formulate an answer. In addition, it appeared that some children felt that they should say they had improved, as they had been trying hard to follow the preventative advice they had been given at their first visit.
3. *Length of time between appointments (mean 29 days)*
   The length of time between appointments may have made it more difficult for children to recall how they had felt at their first visit. The majority of dental studies have used a period of two weeks. However, this was not possible in the present study due to the availability of appointments, and reflects a 'real-life" clinical context.

It may therefore be useful for test-retest reliability to be tested in a larger sample with a more consistent recall period.

Both CARIES-QC and CPQ$_{11-14}$-ISF:16 had a similar proportion of participants who reported impacts related to getting "food stuck", "pain" and "difficulty eating". Construct validity was good with all correlations in the hypothesised direction. Significant correlations were found between CARIES-QC total score and CPQ$_{11-14}$-ISF:16 total score, CARIES-QC global score, the global oral health rating and global life overall rating of CPQ$_{11-14}$-ISF:16, pain, pulpal involvement, total number of carious teeth and presence of anterior caries indicating that CARIES-QC is a valid measure of caries-related impacts.

7.5.4 Comparison of CPQ$_{11-14}$-ISF:16 with previous studies

Cronbach's alpha for CPQ$_{11-14}$-ISF:16 was 0.84, which is comparable to values reported in the original validation study (Jokovic et al., 2006). The mean score was 15.06 with a range of 0–44. This mean score is similar to that found in children with caries in New Zealand and Brazil, although the children in these studies were older than those in the present study (Torres et al., 2009; Foster Page et al., 2013b). The mean score in the present study was higher than found in 5–8-year-olds in another study in New Zealand with children who had similar caries experience to those in this study (Foster Page et al., 2013a). Many of the children included in this study had been referred for treatment specifically because of the impacts they were experiencing and therefore this may account for the higher scores achieved in this population. Correlations with the global rating of oral health (r=0.397) and global rating of life overall (r=0.403), were also similar to those found in the original validation study, although the correlation with oral health was slightly stronger (r=0.397 compared with r=0.21) (Jokovic et al., 2006). Stronger correlations were also found with the rating of life overall than with the global rating of oral health when the measure was used in New Zealand (Foster Page et al., 2008). However, Torres and co-workers found stronger correlations with the overall score and the global oral health rating (Torres et al., 2009). In contrast a study in Thailand did find lower mean CPQ_{11-14}-ISF:16 scores than in the present study (Gururatana et al., 2011). However, although the children had a comparable caries experience in their permanent dentition to those in the present study, children with carious primary teeth were not included. This may account for the difference in mean scores between the two studies. It appears that CPQ_{11-14}-ISF:16 functioned similarly in the present study compared to previous studies. However, the majority of children in these studies were older and therefore it is not possible to draw direct comparisons.

7.5.5 Responsiveness

An initial analysis of responsiveness was performed with a relatively small sample, the majority of whom felt that their dental condition had improved. However, there were significant correlations with the global rating of change score and the CIS total score at follow-up. In addition, the mean CIS for those who reported an improvement was lower than for those who did not report an improvement. This would indicate, that despite the small sample size, scores appear to be responding in the expected direction.

The impact which decreased most from baseline to follow-up was “feeling annoyed” in those who reported an improvement. Reduction in this emotional aspect is an interesting finding, as one might have expected functional impacts to be most improved following treatment. This shows the value of including all aspects that are important to patients and not just focusing on functional aspects. Large reductions were also found in the impacts “crying”, “hurts”, “brushing teeth”, “eating on one side” and “feeling cross”, demonstrating the breadth of impacts which were perceived to have reduced following dental treatment. Not surprisingly, given that the majority of patients had teeth extracted as part of their treatment, there were lower reductions in some of the reported impacts relating to eating which might be expected following multiple extractions. There was also a 44.3% reduction in the proportion of children who responded “a bit” or “a lot” to the global CARIES-QC question where an improvement in oral condition was reported at follow-up.

Future studies investigating different treatment modalities may be able to evaluate which treatments most reduce the impacts experienced from the patient’s perspective.

7.5.6 Strengths and limitations

Strengths

There was an excellent response rate with only one child declining to complete the questionnaire. This study demonstrates the feasibility of administering questionnaires to young children in the clinical setting. Indeed, many participants stated that they had enjoyed completing the questionnaires and were excited about completing them again. The involvement of trained research nurses was invaluable in the recruitment process, as this allowed participants to be recruited during sessions where FG was engaged in other activities. As CARIES-QC is designed to be administered in English, children who could not understand spoken English were not included. However, the availability of a number of trained staff meant that where parents were unable to read English sufficiently, help could be given to ensure the inclusion of as many children from minority ethnic groups as possible. A higher proportion of participants were from minority ethnic groups when compared to the Sheffield population, therefore it would appear that these groups had good representation. However, further cross-cultural adaptation of the measure into different languages would aid inclusion of minority ethnic groups in future studies.

Discussion with children when they completed the questionnaires at T1 and T2, allowed some insights to be gained into why children thought their teeth were "better", "the same" or "worse". This is important for future research, as many children felt that their teeth had improved, despite no improvement in symptoms and where no treatment was provided, because they had been trying hard to adhere to the dietary and oral hygiene advice provided at their initial visit. Therefore, as was found in this study, it may be difficult to recruit adequate numbers of participants who report that their dental condition is unchanged in order to assess test-retest reliability in this type of environment.

Limitations

The limitations of this part of the study were that the vast majority of participants were recruited from a secondary referral centre. Only two children were recruited from primary care in this sample. This was due to recruitment starting within primary care after the process had been established in the hospital environment. Recruitment within the hospital was higher than expected with only one child refusing to participate. Due to waiting list pressures, several additional new patient clinics were available, thus increasing the number of potential participants within the recruitment period. Due to ethical constraints it was not possible to recruit further patients once the target had been achieved. Additionally, recruitment in primary care proved challenging as there were few patients who attended who met the inclusion criteria.

Recruitment from the secondary care environment resulted in the engagement of mostly younger children with high levels of untreated dental caries as this reflects the referral pattern to this unit. Despite this, there were still children who reported no impacts resulting from their dental caries and there was a wide range of scores on both CARIES-QC and CPQ_{11-14}-ISF:16. Therefore it is likely that the participants reflected the range of impacts experienced by children with dental caries. However, further testing in other areas and settings is required to confirm its validity.

Rasch analysis revealed that the participants had slightly fewer impacts than where the scale was centred. However, the range of questions and participants was similar, indicating that the scale measures the impact of dental caries. This can be expected as many children reported few impacts, which is consistent with clinical observations that some children report few impacts even when they have extensive disease. However, further testing in populations with lower caries levels is required to test the suitability of the scale in these populations.

CPQ_{11-14}-ISF16 was not used at follow-up which does not allow the responsiveness of the two measures to be compared. This is a limitation as it cannot be assumed that CARIES-QC is more responsive than CPQ_{11-14}-ISF:16 following treatment for dental caries. The reason CPQ_{11-14}-ISF:16 was not used at follow-up was twofold. First, CPQ_{11-14}-ISF:16 has not been shown to be unidimensional and therefore it was not known whether accurate change scores could be calculated for this measure, therefore it would not have been possible to directly compare the two measures in this way. Additionally, a number of parents had reported that their children had not been able to understand or found it difficult to complete CPQ_{11-14}-ISF:16 and that they would not wish to complete this a second time. As it was not possible to directly compare the two measures, it was felt to be preferable to maximise retention of participants at follow-up by including only one measure. Future studies with defined follow-up periods and treatment schedules may be able to compare the raw scores of the two measures in future, although it would be recommended that efforts were made to ensure that accurate change score could be calculated for CPQ_{11-14}-ISF:16.

As discussed in Section 7.5.5**Error! Reference source not found.** only a few participants returned for their review visit at T2. Attempts were made to enable further participants to return at this stage but many of the participants were not contactable. The children who did return were representative of the original sample and the majority had extractions under GA. It would be expected that children who had such treatments may experience reductions in the caries-related impacts they had previously experienced. However, it is not clear whether they may then experience different impacts as a result of their treatment (for example, multiple extractions). Children who had less invasive treatments (Hall crowns), also experienced reductions in the number of impacts they experienced. Therefore it would appear that any reduction in impacts was not limited solely to those who had extraction of carious teeth. It should be noted that these finding relate only to a small number of children and were only intended to give an initial assessment of whether the measure appeared to responding in the expected direction. Future studies are required to investigate the measure's responsiveness in a more robust fashion and in populations with lower levels of caries experience.

7.5.7 Conclusions

In conclusion, CARIES-QC proved to have acceptable psychometric properties and appears to be acceptable to participants. It had strong correlations with $CPQ_{11\text{-}14}$-ISF:16 but had stronger correlations than $CPQ_{11\text{-}14}$-ISF:16 with clinical indicators and the global ratings, indicating the value of a disease-specific measure of OHRQoL. Initial testing of responsiveness was promising, but future studies with larger samples and more consistent treatment regimens are required to confirm this finding and to accurately calculate the MID. Further testing in larger and more diverse populations would be beneficial as would the development of a computer-based or online version.

7.6 Publications arising from the work presented in this chapter

Published abstract

- Gilchrist F, Rodd HD, Deery C, Marshman Z. Development of CARIES-QC: a caries-specific measure of oral health-related quality of life. Accepted for presentation at IAPD, Glasgow 2015.

Chapter Eight

Discussion

8.1 Outline

This thesis has presented three interlinked studies which contributed to the development of CARIES-QC, a child-centred, caries-specific measure of OHRQoL. This chapter will consider the key findings, important aspects of the study design and the strengths and limitations of these. This will be followed by discussion of the implications of the research for clinical care and oral health policy. Finally, recommendations for future research will be proposed.

8.2 Overview

The narrative review described in **Error! Reference source not found.** confirmed that dental caries was a disease of public health significance. A gap in the literature was identified regarding the description of the impacts of caries from the child's perspective. Various impacts had been reported following use of different child self-report measures of OHRQoL. However, there had been no comprehensive review of the properties of these measures and it was often unclear whether they had been validated for the purpose for which they were being used. In addition, the measures were generic and there was no existing caries-specific measure which would be suitable for use in clinical trials evaluating treatment for dental caries. Therefore the three interlinked studies aimed to address these issues. The next section will discuss the key findings from each part of the study.

8.3 Key findings

The aim of this research was to develop a child-centred caries-specific measure of OHRQoL which has been achieved by fulfilling the following objectives:

1. *Perform a review of the most commonly used self-report measures of OHRQoL for children against existing criteria, to assess strengths and weaknesses which will aid the development of the new measure.*

   This objective was met by performing the systematic review of child self-report measures of OHRQoL in Chapter 5. This study identified deficiencies in the development and evaluation of existing measures of OHRQoL. The COSMIN checklist was found to be a useful tool and was used as a framework to develop CARIES-QC. In addition, Rasch analysis was chosen to ensure that CARIES-QC had been evaluated using modern psychometric techniques, which was lacking in the evaluation of existing measures.

2. *Explore, through qualitative methods, the impacts of dental caries on children.*

   This objective was met in the item generation stage described in Chapter 6. Children described a wide variety of impacts, principally related to pain. The impacts discovered during this exploration were used to develop a caries-specific measure of OHRQoL, which used the language and descriptions adopted by children themselves.

3. *Involve children in the design and content of the measure to ensure that it is meaningful and relevant to them.* This objective was met by involving children throughout the questionnaire development process and making changes to the structure and wording of the measure following their suggestions.

4. *Test the measure for validity, reliability and responsiveness using modern psychometric techniques and existing criteria.*

   This objective was fulfilled by using the COSMIN checklist as the framework to validate the measure and Rasch analysis to produce a measure which is unidimensional and free of DIF. CARIES-QC was found to be valid, reliable and responsive.

Further details of the key findings for each part of the study will now be discussed.

8.3.1 Systematic review

The systematic review used criteria proposed by Locker and Allen (2007) and the COSMIN initiative to assess the three most commonly used self-report measures of OHRQoL for children (Terwee et al., 2007; Mokkink et al., 2010a). These measures had been widely used, but there had been little critical evaluation of them and no refinement of the measures in their original language.

This is the first study to evaluate OHRQoL measures using the COSMIN checklist. The COSMIN initiative recommends that the checklist is used in the development of measures and in systematic reviews of health measurement instruments to standardise reporting of measurement properties (Mokkink et al., 2010a). The COSMIN checklist has been used successfully in systematic reviews of other health measurement instruments (Schellingerhout et al., 2012; Chow et al., 2013; Park et al., 2013; Pusic et al., 2013; Weldam et al., 2013; Haywood et al., 2014). In the present study, the COSMIN checklist, in combination with the measurement property quality criteria, allowed the psychometric properties of the three measures to be objectively assessed. In addition, the evidence for each measure could be synthesised allowing the strength and weaknesses of each to be identified (Table 5.3).

Both the COSMIN checklist and the measurement criteria had high intra- and inter-examiner reproducibility (weighted Kappa >0.9) in the present study. The level of agreement between reviewers was higher than has been reported elsewhere (Mokkink et al., 2010b). This may be due to the limited number of reviewers involved (FG and ZM) and the use of a training and calibration exercise prior to the start of the study as recommended by Mokkink and co-workers (2010b). As the COSMIN checklist and quality criteria proved useful tools, it was decided to use them as a framework for the development of the new measure, thereby ensuring that the psychometric properties were adequately evaluated.

The majority of studies in the review reported cross-sectional studies in various populations and the three included measures had been translated into several different languages. The review found that it was difficult to compare studies due to omissions in the data presented and the variety of non-validated versions of the measures which were used. In particular, there were omissions relating to: the proportion of missing items and how these were handled; how the scores for different subgroups were presented, and the prevalence of the condition under scrutiny. In addition, some of the measures had been used longitudinally, without being validated for this purpose.

The measures included in the systematic review were all developed prior to the existence of the criteria used to assess them, which may account for some of the deficiencies identified. There was a lack of involvement of children in the early stages of development of the measures which may affect their face and content validity. Only one study was identified which had evaluated face and content validity and this found that children with malocclusion in the UK found some of the questions in CPQ_{11-14}-ISF:16 irrelevant and the response format was not based on how they would describe the impacts of malocclusion (Marshman et al., 2010). Further evaluation of face and content validity of these measures is required to ensure that they are relevant to the populations in which they are used. Furthermore, modern psychometric techniques were seldom used to confirm the unidimensionality of the scales. Techniques such as Rasch analysis are able to estimate this and provide information enabling evaluation of the suitability of the response format. However, despite these deficiencies the majority of measurement properties were adequate and therefore there is value in using existing measures in appropriate populations. The availability of more modern psychometric techniques does offer scope for these to be further refined to improve their measurement properties.

8.3.2 Development of the measure

Children were capable and willing to discuss the impacts they had experienced. Previous studies have found that young children are able to report on aspects related to their general health (Otley et al., 2002; Young et al., 2004; Detmar et al., 2006; Morris et al., 2007; Carlton, 2013b). Not surprisingly, the overarching theme identified was related to pain. However, there were many impacts which were related to pain and children discussed mechanisms that they used to avoid pain, such as altering their eating patterns and avoiding certain foods. It became obvious during the testing of the resultant measure that parents often were not aware of some of these difficulties that children reported, reinforcing the importance of directly seeking children's views rather than relying on parental reports. Long-term avoidance of certain foods may have other general health effects in growing children, such as have been reported elsewhere (Acs et al., 1999; Clarke et al., 2006; Monse et al., 2012; Schroth et al., 2013). Children generally discussed the impacts they had experienced in terms of their severity, rather than the frequency with which they occurred. This is an important finding as some of the existing measures of OHRQoL (CPQ and COHIP) rely on a frequency-based response format which does not appear to reflect how children describe the impacts they experience. However, severity-based response formats have been used by others who have utilised children's terminology to develop response formats (Stevens, 2009; Carlton, 2013a).

Conducting the analysis of face and content validity gave an insight into how children read and answer questions. The analysis of face validity was essential as it revealed problems with the way some of the questions were worded. In particular, children's perception of their oral health was interesting, as it appeared to have little to do with the condition of their teeth or the symptoms they had experienced. This is an important finding, as it may explain why previous studies have found relatively low correlations with global oral health ratings and total CPQ scores in children with caries (Jokovic et al., 2002; Jokovic et al., 2004; Robinson et al., 2005; Goursand et al., 2008; Aguilar-Diaz and Irigoyen-Camacho, 2011; Gururatana et al., 2011; Kolawole et al., 2011). In the present study there was a strong (r = 0.734) correlation between the global oral health rating in CARIES-QC and the total CARIES-QC score. This strength of correlation between CARIES-QC total score and the global oral health question in CPQ$_{11\text{-}14}$-ISF:16 was not seen, demonstrating that although they both seek to address the same construct, they do not appear to ask the same question.

Recruitment to the interview stage was challenging and there was a low response rate (27%). There may be a number of reasons for this including: families forgetting to post reply slips due to the amount of information received at their initial dental appointment; family commitments including extracurricular activities and parent work schedules or that some families would prefer not to have a researcher visit their home. Families were offered a choice of venue, however, as many of the children did not have follow-up appointments at the hospital, this option was unlikely to be popular as it would have required time to be taken off school and for the family to travel to the hospital. The addition of a telephone call to follow-up the invitation, did result in an increased response rate, indicating that this method is likely to result in increased recruitment when compared to simply supplying a reply slip. The recruitment rate for this part of the study is similar to other qualitative studies which have been carried out in Sheffield with children requiring tooth extraction under GA and with children with cleft lip and palate (Hall et al., 2012; Rodd et al., 2013). Despite these difficulties, children from different ethnicities, socioeconomic status and clinical presentations were recruited. These difficulties were not encountered in the other parts of this stage of the investigation, the majority of which took place during routine clinic visits, thus reducing the burden for the participants and their families.

8.3.3 Ethical considerations

Conducting qualitative studies such as these with children requires consideration of a number of ethical issues, such as power imbalance, assent, confidentiality, as discussed in Section 4.3. Efforts were made to reduce the power imbalance by ensuring that interviews took place where children were comfortable and emphasising that they could stop the interview at any time. The interviewer (FG) is a paediatric dentist, and therefore has a wealth of experience in assessing children's body language and adapting to it. This experience was used to assess if children were showing signs of distress or if they wanted to stop the process but did not verbally express this. None of the children showed any distress, but observation of their body language or changes in the way they answered questions did allow the interviewer to change the subject if they appeared to be becoming bored with a particular line of enquiry or to close the interview if they were becoming fatigued. Children were asked to give their assent to participate in the focus groups and interviews following discussion with the interviewer. No safeguarding concerns were raised in any of the interviews, however, normal procedures would have been followed if any had arisen. To reduce preconceptions of the child's dental condition and allow the children's reports to be taken at face value, the case-notes were not examined by FG until after data analysis had taken place. Steps were taken to ensure the interviewer's safety, however, the interviewer was made to feel very welcome in each of the homes that were visited and no personal safety issues arose.

8.3.4 Evaluation of CARIES-QC

Children as young as five years of age were able to complete the questionnaires, with their parents or one of the research team helping them to read the questions. This concurs with previous observations that five-year-old children are able to reliably report their HRQoL given the opportunity to do so using an age-appropriate instrument (Varni et al., 2007).

The children who participated in the evaluation stage of the study had higher caries experience, than those reported for the general population in Sheffield (Public Health England, 2012), which was to be expected from a secondary care case mix. However, the same trend was not seen for the caries experience of children in the permanent dentition. This may be because young children are more likely to be referred for treatment of caries under GA, while older children may be more able to accept treatment in primary care.

This study found that children from ethnic minority groups had higher mean CARIES-QC total scores than their white-British counterparts, despite their being no significant difference in their caries experience. While it is not possible to draw any conclusions from these data, others have found similar findings using child self-report measures (Meeske et al., 2007; Wallander et al., 2012). However, Wallander and Colleagues (2012) reported that when family context and socioeconomic status were controlled these differences were reduced, concurring with the view that investigating ethnicity alone may fail to consider the relevance of social class and low income as contributing factors (Bedi and Uppal, 1995; Watt and Sheiham, 1999).

Construct validity was tested using correlations with the total CARIES-QC score and clinical data, the global rating of CARIES-QC and CPQ_{11-14}-ISF:16 total score. CARIES-QC had stronger correlations with clinical data than CPQ_{11-14}-ISF:16 indicating that it may be more sensitive to assessing those impacts specifically associated with dental caries. CPQ_{11-14}-ISF:16 was designed to be generic and be able to evaluate the impacts of a variety of oral conditions, thus it may be that some of the included items it contains are irrelevant to children with caries.

Rasch analysis allowed a unidimensional measure to be produced and identified questions which did not fit the mathematical model. Indeed, the question relating to taking medicine, was one which many children seemed to seek confirmation from their parents. Children appeared to be unsure if they had taken medicine for problems with their teeth and it may be that children are often given analgesia for other problems and therefore cannot remember why they were given it. This question was subsequently removed following Rasch analysis, as it had poor fit statistics, indicating that indeed, it did not work as intended.

A three-point Likert scale was adopted as the response format for CARIES-QC. This was chosen as others have suggested that young children tend to choose extremes and other measures for young children have also used a three-point scale (Varni et al., 2001; Carlton, 2013c). Furthermore, young children have also expressed a preference for three-point scales, while older children prefer more options (Detmar et al., 2006). With the exception of one question (“feeling tired”), Rasch analysis demonstrated that children were able to distinguish between the three different response options. It was also in keeping with the ethos to minimise participant burden and to allow children with low literacy levels to participate. There is some debate about the number of response options which should be used, with some suggesting sensitivity increases with increasing number of options, whilst others argue that reliability is increased with fewer response options (Weng, 2004). It has been found that the number of response options may have less of an effect where the items are homogenous, such as in a unidimensional scale (Weng, 2004). Despite the use of a three-point scale, CARIES-QC was able to differentiate between participants with differing clinical presentations and to evaluate changes following treatment.

Thirty-three of the 42 participants who completed CARIES-QC at follow-up felt that their oral condition had improved following their treatment, with only three feeling it had got worse. There was a large reduction in some of the impacts experienced following treatment (in those who felt there had been an improvement), notably those relating related to pain and emotions.

8.4 Strengths

8.4.1 Systematic review

This is the first study to report a comprehensive review of child self-report measures of OHRQoL against existing criteria. The original versions of the measures were found to have some limitations, as discussed in Section 8.3.1. The majority of measurement properties were found to be adequate, although refinement of these measures using modern psychometric techniques is recommended. The present study demonstrated the feasibility of using the COSMIN checklist and quality criteria and recommends its use as a framework for reporting future development or validation studies to ensure consistency.

8.4.2 Development

This is the first study to investigate the impacts of dental caries from the child's perspective using the theoretical framework of the new sociology of childhood. Children as young as five years were able to describe the impacts they had experienced. While younger children provided less descriptive answers than their older counterparts, it was clear that they were capable of discussing the impacts they had encountered.

Children were involved throughout the development process and their views were taken on board and modifications made based on their suggestions. Although this study was designed by adults, children's involvement could be classified as lying somewhere between "Assigned but informed" and "Consulted and informed" on Hart's Ladder of Involvement (Table 4.1) as children's views were acknowledged at all points (Hart, 1992). However, the level of understanding of their overall contribution will have varied depending on age and developmental stage, with some of the older children being more able to understand how their views were being used to design and alter the questionnaire. At the conclusion of data collection, all children who participated will be sent a report detailing the findings and how each stage of the investigation contributed to the final measure. It is hoped that this report will aid their understanding of how their views influenced the development of the measure.

During the design process, children were seen to put a great effort into reading every word in a statement and hence it became obvious that the number of words should be kept to a minimum to reduce participant burden and to allow those with low literacy levels to participate. This approach was used throughout the study to maximise children's involvement. Previous studies have used an item reduction questionnaire, to choose the final item pool (Juniper, 1997a; Jokovic et al., 2002). However, in the present study the method was simplified by allowing children to choose cards with the impacts they had experienced and then choosing which of these had most bothered them or had occurred most frequently. This method was chosen as it was felt that it would allow children with different reading abilities and whose first language was not English to participate. A similar strategy was employed during item reduction of the Canadian Hemophilia Outcomes-Kids Life Assessment Tool (Young et al., 2004).

8.4.3 Evaluation

Recruitment to the evaluation stage of the investigation was excellent and an adequate sample size was obtained to perform an initial evaluation of CARIES-QC. Children were enthusiastic participants and many reported enjoying completing the questionnaire, indeed some felt sad that they could not complete it at each visit. There were very few isolated missing responses in CARIES-QC, suggesting that the children who completed it understood the questions and the response format. A higher proportion of isolated missing responses was found in $CPQ_{11\text{-}14}$-ISF:16, which concurred with some parents' admittance that their children had found it difficult to understand the response format of this measure. In addition, parents whose first language was not English, also required help to read and explain $CPQ_{11\text{-}14}$-ISF:16 to their children.

This is the first study to use Rasch analysis to develop a measure of OHRQoL for children. This technique showed that the two items related to aesthetics ("being able to see holes" and "front teeth looking brown or black") did not fit the mathematical model. This finding concurred with that of the qualitative part of the study, where two themes (pain and aesthetics) were identified. The use of Rasch analysis also demonstrated that the response format did not seem to work for the item "tired" although there were clear boundaries between the thresholds for each of the other items. Use of Rasch analysis has produced a unidimensional scale which is free from DIF. The production of an interval scale means that change scores can be accurately calculated if CARIES-QC is used in future clinical trials.

8.5 Limitations

8.5.1 Systematic review

The systematic review included articles which had used CPQ, COHIP or C-OIDP as these were the most commonly used self-report measures. Only studies which were published in English were included, therefore some relevant studies may have been excluded. However, two of the three measures (CPQ and COHIP) were developed in English. Therefore, as the quality criteria and COSMIN checklist were only applied to the original versions of the measure, it is unlikely that studies reporting validation of the original versions of these measures were omitted. Since the systematic review was performed, SOHO-5 has been used more frequently, however, the majority of published studies have used the Brazilian version of this measure and therefore would not have been included in the analysis using the COSMIN checklist and the quality criteria (Abanto et al., 2013b; Abanto et al., 2013a). The COSMIN checklist was developed after the included measures were developed and therefore it is not surprising that the included measures may not have met all the criteria, in what is a rapidly developing field. The aim was not to criticise the existing measures, as they were developed with the best techniques available at the time, but rather to assess objectively their deficiencies. This will allow these measures to be refined in future and the critique provides information regarding their limitations. Using up-to-date criteria also informed the development of CARIES-QC to ensure that it was developed robustly using modern techniques.

8.5.2 Development of the measure

Recruitment to the in-depth interviews was difficult, but as the children were purposively sampled to include a range of ethnicities and clinical presentations this did not result in under-representation of specific groups. However, it should be noted that the group who participated were self-selecting and therefore the experiences of those who did not participate may differ. In addition, as only English speaking families could be recruited to this stage, the views of children with parents who did not speak English were not included. This group is likely to have included families of recent immigrants, such as families of Roma descent (an increasing population in Sheffield), who may have high levels of dental caries (Sheffield City Council, 2012). Although the interviews and focus groups did not include any Roma children, they were included in the other stages of the investigation where informed consent could be obtained from their parents.

This stage of the investigation aimed to retain the language children used in the resultant measure. Therefore, while the measure reflects the language children use in Sheffield, it remains to be seen whether this language is appropriate outwith this geographical location. Attempts were made to ensure local colloquialisms were not included and certainly it would appear that the language the children used to describe pain was similar to that found throughout the UK and in the US (Mills, 1989; Stanford et al., 2005; Franck et al., 2010). The CHU9D, which was also developed in Sheffield, incorporates similar words in its response format and has been used successfully in other parts of the UK and in Australia and New Zealand (Stevens and Ratcliffe, 2012; Canaway and Frew, 2013; Page et al., 2014). Therefore it is hoped that CARIES-QC would also be understood elsewhere, however, testing in other populations is required to evaluate its properties in other English speaking countries.

CARIES-QC is designed to be used with children over the age range of 5-16 years, as the impacts that children described were similar regardless of age. It had initially been hypothesised that at least two versions would be required to cover the wide age range, but as the impacts were similar it was decided to have one measure allowing it to be used longitudinally in trials. The items were worded using the language used by the youngest children to ensure that it could be understood by all. The design stage included a range of children aged 6-15 years, who commented on the wording and layout. The oldest children did not feel that the language or layout was too "babyish" and the youngest children were able to read the questions and understood the response format. These opinions were reiterated during the testing of face and content validity. Indeed, there were some 14-16-year-olds who found reading the items challenging during those stages. Further qualitative enquiry with children in the permanent dentition is required to assess whether the aesthetic items would be useful addition as this may be something which is more important to children in this age group.

8.5.3 Evaluation of CARIES-QC

CARIES-QC had good construct validity and unidimensionality in this population. However, it is acknowledged that only two children were recruited from primary care, and the participants had a high caries prevalence and extensive disease. Although the majority of participants had extensive disease (80% had pulpal involvement), a number reported that they had not experienced pain and few other impacts. It should be noted that the findings may be different in a population with less extensive disease. Children with lower levels of caries may have very low scores on CARIES-QC and therefore it may be difficult to detect differences following less invasive treatment modalities. Therefore, it would be useful for the measure to be tested in populations with lower levels and less extensive disease to confirm its validity, reliability and responsiveness in those populations. However, it should be recognised that the measure was primarily designed to evaluate change in those who require interventions for dental caries and not to discriminate between children with different levels of disease. Therefore, it is likely that those who require treatment, will have some impacts which may improve following an intervention.

The ICC for test-retest reliability was slightly lower than ideal, and this may be due to the length of time between the two administrations of the measure, the low number of children who felt that their oral condition had remained stable or the fact that use of a three-point response scale may reduce test-retest reliability (Weng, 2004). Testing of this element in a larger population with a shorter and more consistent time period between administrations is required. Responsiveness was only tested in a very small proportion of the participants (n=42) to give an indication of whether it responded as expected. However, the initial results are promising with those who felt they had improved having lower mean scores than those who did not report an improvement. An estimation of the MID was made for this population, however, additional testing in a larger population is required to further test the responsiveness of the measure. Larger populations would also allow the calculation of effect sizes and the standardised response mean, which would aid the clinical interpretation and calculation of future sample sizes (Guyatt et al., 1987a).

The following sections will discuss the implications of the research for clinical care, policy and future research.

8.6 Implications for clinical care

The findings in these studies have several implications for clinical care of children with caries. Firstly, understanding the language children use can help clinicians to ask better questions when taking histories from children. For example, there may be little point in asking if a child has had toothache, as the results of the qualitative study demonstrated that many young children do not know or use this term. The most commonly used word to describe pain, was "hurts" and so adopting this terminology may help children to express their experiences. Whilst others have reported that children may be able to choose a number of different words from a list to describe pain (Wilkie et al., 1990), in this study some of the children used terms such as "prodding" to describe the pain, whilst others just described it as "hurting". As the aim was not to produce a pain measure, the word "hurts" was chosen for CARIES-QC, rather than attempting to describe the quality of the pain. In addition, some children did not report pain, but had experienced other impacts of which their parents were not always aware. Discussing aspects such as difficulties eating may help clinicians to more accurately assess a child's need for treatment, as while they might not report pain, caries may be having other effects on their daily lives, which may improve following treatment. In addition, the routine use of CARIES-QC in the clinical environment may help clinicians better understand the impacts that young children are experiencing, allowing treatment to be tailored to reducing these.

Secondly, participants reported an overall reduction in impacts following treatment and the majority of children reported that they felt the condition of their teeth had improved following treatment. Although this was finding was in a limited sample, it does demonstrate the benefit of treating dental caries. Others have cast doubt upon the benefit of treatment for dental caries in the primary dentition when analysing only pain as an outcome (Levine et al., 2002; Tickle et al., 2002). Impacts on eating improved less than other impacts, which is not surprising given the large number who had tooth extractions as part of their treatment. While extraction of some teeth may be the only feasible option due to the extent of the disease, future studies may be able to ascertain which treatments reduce impacts most from the child's perspective.

8.7 Implications for policy

The results of this investigation clearly show that caries has multiple effects on children's lives, which appear to improve following treatment. Children's oral health should be a national priority and policies developed to help prevent development of this disease, to minimise the impacts experienced by children. Caries shares common risk factors with obesity and Type II diabetes, therefore public policy should adopt a common risk factor approach to reducing the incidence of these conditions, which is more cost-effective than disease-specific approaches (Sheiham and Watt, 2000). Indeed, the WHO has issued guidance recommending the reduction of free sugars to below 10% of total energy intake, due to the association of free sugar intake and non-communicable diseases, such as dental caries, diabetes and obesity (World Health Organization, 2015a). The majority of the children in the evaluation study lived in the most deprived areas, according to the IMD score. Population-based interventions, including school toothbrushing and fluoride varnish schemes, such as Childsmile have been shown to reduce the effect of social inequalities on the prevalence of caries and should be given further consideration in England (McMahon et al., 2011).

The new prototype dental contracts in England have a focus on prevention with remuneration for this activity (Department of Health, 2015). It is essential that any dental contract reform adequately remunerates general dental practitioners for providing preventive interventions and allows them to perform clinically effective treatments for their patients. The evaluation study showed that 80% of children who participated had caries which extended to the pulp. Earlier treatment with non-invasive techniques such as the Hall technique, may have arrested the progression of caries and prevented many of the impacts the children had suffered. In addition, commissioners must ensure that families can access appropriate dental services, as this study shows that young children in particular may require specialist services for treatment under general anaesthesia. There were reductions in the majority of impacts following treatment, however, this reduction was less apparent for impacts related to eating, which is not surprising given that most of the children had undergone tooth extraction. Future studies are required to identify which treatments are associated with the most positive outcomes, but perhaps restoration of teeth would improve impacts associated with eating when compared to tooth extraction.

Part of the new prototype dental contract aims to develop PROMs and Patient-reported Experience Measures (PREMs) (Department of Health, 2015). CARIES-QC could be considered for inclusion in these new contract monitoring arrangements. However, if not found to be suitable, this study demonstrates that young children are competent in discussing their conditions and should be involved in the development of any future PROMs or PREMs for this purpose.

8.8 Implications for future research

This study has produced a valid caries-specific measure of OHRQoL. CARIES-QC had higher correlations with clinical data and global ratings of oral health than CPQ_{11-14}-ISF:16 indicating the potential benefits of using a disease-specific measure in studies investigating the effectiveness of interventions for dental caries. Future studies investigating the effect of different interventions for caries in children would benefit from the inclusion of CARIES-QC to assess effectiveness from the child's perspective. Further testing of CARIES-QC in large populations would allow responsiveness and test-retest reliability to be assessed and would permit calculation of the MID, effect sizes and the standardised response mean which would aid future sample size calculation.

There are limitations to using only a disease-specific measure in clinical trials, as this means that any reductions in impact cannot be compared with reductions seen in other conditions. The limitations of this can be mitigated by also including a generic measure or a health utility measure which allows comparison with other conditions. However, as funding for dentistry in the UK is separate from that of general health conditions, the use of a disease-specific measure would provide valuable information to commissioners regarding which dental treatments to support.

Whilst using CARIES-QC may provide information regarding the effectiveness of different interventions from the child's perspective, it does not allow economic evaluation to be performed. Preference-based measures (PBMs) of HRQoL, such as the CHU9D, permit cost utility analysis to be performed by producing a single summary measure, usually the quality adjusted life year (QALY) (Stevens, 2010). The QALY can then be used to assess the cost-effectiveness of different interventions. To date, one study has investigated the use of CHU9D as an outcome measure for child dental health (Page et al., 2014). Although CHU9D showed promise, it was less sensitive to caries as a variable than $CPQ_{11\text{-}14}$-ISF:16. The development of a condition-specific preference-based measure based on CARIES-QC may be more sensitive to changes associated with caries than a generic health measure such as CHU9D. Condition-specific PBMs have been successfully developed from existing HRQoL measures (Brazier et al., 2008; Yang et al., 2011). It has been reported that condition-specific PBMs may not be comparable with their generic counterparts, as they may not be sensitive to side-effects which are not covered by the condition-specific measure due to the narrow focus (Versteegh et al., 2012). However, this is less relevant for caries-related interventions which are less likely to have side-effects outwith the oral cavity, when compared with drug interventions.

Researchers using existing measures of OHRQoL should be aware of their deficiencies and ensure that they are used for the purposes for which they were developed and validated. As yet, they have not been validated appropriately to assess their evaluative properties and therefore caution should be used when interpreting change scores, especially as their unidimensionality has not been confirmed. In addition, correlations with CPQ total scores and the CPQ global ratings of oral health should be interpreted with caution, as it appears that children may not rate their oral health based on the condition of their teeth, but rather on broader aspects of general health. This may have implications for the construct validity of this measure.

In conclusion, the studies contained within this thesis have added knowledge about the impact of caries on children's everyday lives. Additionally they have also demonstrated that children are able to discuss the impact of caries and contribute to the development of a measure to investigate this. The resultant measure, CARIES-QC, has been evaluated using a mixture of qualitative and quantitative methods using the criteria suggested in the COSMIN checklist and been found to be valid and reliable. The use of Rasch analysis has ensured that CARIES-QC is unidimensional and can accurately be used to evaluate change. Future studies are required to further evaluate its responsiveness.

Chapter Nine

Conclusions and recommendations

The studies presented within this thesis aimed to produce a child-centred caries-specific measure of OHRQoL. The objectives were to:

1. Perform a review of the most commonly used self-report measures of OHRQoL for children against existing criteria, to assess strengths and weaknesses which will aid the development of the new measure
2. Explore, through qualitative methods, the impacts of dental caries on children
3. Involve children in the design and content of the measure to ensure that it is meaningful and relevant to them
4. Test the measure for validity, reliability and responsiveness using modern psychometric techniques and existing criteria

These objectives were fulfilled in the three interlinked studies using a variety of novel methods. The research presented in this thesis provides additional knowledge regarding existing self-report measures of OHRQoL, the impact of caries on children's everyday lives and their ability to report these. The main outcome of these studies is CARIES-QC, a caries-specific measure of OHRQoL for children which was developed using child-centred methodologies. This chapter summarises the findings and recommendations arising from the studies.

9.1 Summary of findings

- Studies which used existing child self-report measures of OHRQoL were difficult to compare as vital information affecting their generalisability and interpretability were omitted.

- Existing child self-report measures of OHRQoL involved children only at the latter stages of their development which may compromise their face and content validity. The use of modern psychometric techniques to develop and validate these measures was lacking.
- The COSMIN checklist and measurement quality criteria proposed by Terwee and co-workers (2007) proved a useful and reliable tool to objectively review the quality of studies reporting validation of the measures and their associated measurement properties.
- Children reported a number of impacts of dental caries. These were mainly related to the pain they experienced and its associated effects. Children also discussed how they adapted their eating habits to minimise the pain they experienced.
- Children mainly discussed the impacts they experienced in terms of their severity rather than their frequency.
- Involving children in the design and face and content validity of the measure ensured that the response format made sense to them and that questions were worded appropriately.
- The use of Rasch analysis allowed items which did not fit the model to be removed and demonstrated that the response format was being used as expected. The resultant 12-item measure is unidimensional and free of DIF. The conversion to an interval scale allows accurate calculation of change scores.
- CARIES-QC has good construct validity and significant correlations were found with global ratings of change which indicate it can be used to evaluate change.
- The majority of children reported improvement following treatment. Reduction was greatest for emotional impacts and pain. Impacts on eating were reduced but to a lesser extent. Potential reasons for this were discussed.

9.2 Recommendations for clinical care

- Clinicians may be better able to understand the impacts children experience by using the terminology that children use. For example, some children in this study were not aware of the term "toothache".
- Pain should not be the only symptom assessed, as it is clear that other impacts may be present even when children do not report pain. Treatment may reduce these other impacts, therefore treatment should not be restricted to only those in pain.
- Children generally reported an improvement in their oral condition following treatment, indicating the benefits of treating dental caries in both primary and permanent teeth. Future research may be able to determine which treatments are most effective at reducing impacts.

9.3 Implications for policy

- It is clear that caries has multiple impacts on children's lives. Prevention of caries using a common risk factor approach may be an appropriate way to reduce these impacts.
- The majority of children in the evaluation study came from the most deprived areas of England. Population-based prevention schemes may help to reduce social inequalities associated with dental caries.
- PROMs and PREMs need to be used in general dental practice. CARIES-QC should be considered for inclusion but if it is not found to be suitable, then children should be included in the development of these to ensure that they are relevant to young patients.

9.4 Recommendations for future research

- Future clinical trials investigating interventions for dental caries should include CARIES-QC to assess the effectiveness from the child's perspective and also permit further analysis of the measure's responsiveness.
- Consideration should be given to the conversion of CARIES-QC into a condition-specific PBM to allow cost-utility analysis of different interventions.

- Future studies reporting the results of studies using OHRQoL should ensure that information allowing them to be interpreted and generalised is included.
- Refinement of existing measures using Rasch analysis may result in improvement of their measurement properties and allow shorter versions to be developed.
- Future qualitative enquiry of the impacts of other dental conditions may promote greater understanding of what the important issues are for children.

In conclusion, the studies within this thesis have demonstrated the wide-reaching impact of caries on children's lives and have allowed the development of a valid child-centred caries-specific measure. Future studies are required to evaluate which interventions for dental caries are most effective from the child's perspective.

References

Abanto, J., Tsakos, G., Paiva, S.M., et al. 2013a. Cross-cultural adaptation and psychometric properties of the Brazilian version of the scale of oral health outcomes for 5-year-old children (SOHO-5). *Health and Quality of Life Outcomes.* **11,** 16.

Abanto, J., Tsakos, G., Ardenghi, T.M., et al. 2013b. Responsiveness to change for the Brazilian Scale of Oral Health Outcomes for 5-year-old children (SOHO-5). *Health and Quality of Life Outcomes.* **11,** 137.

Abanto, J., Tsakos, G., Paiva, S.M., et al. 2014. Impact of dental caries and trauma on quality of life among 5- to 6-year-old children: perceptions of parents and children. *Community Dentistry and Oral Epidemiology.* **42,** 385-394.

Abu-Saad, H. 1984. Assessing children's responses to pain. *Pain.* **19,** 163-171.

Acs, G., Shulman, R., Ng, M.W. & Chussid, S. 1999. The effect of dental rehabilitation on the body weight of children with early childhood caries. *Pediatric Dentistry.* **21,** 109-113.

Adulyanon, S., Vourapukjaru, J. & Sheiham, A. 1996. Oral impacts affecting daily performance in a low dental disease Thai population. *Community Dentistry and Oral Epidemiology.* **24,** 385-389.

Agou, S., Locker, D., Streiner, D.L. & Tompson, B. 2008. Impact of self-esteem on the oral-health-related quality of life of children with malocclusion. *Am J Orthod Dentofacial Orthop.* **134,** 484-489.

Agou, S., Locker, D., Muirhead, V., Tompson, B. & Streiner, D.L. 2011. Does psychological well-being influence oral-health-related quality of life reports in children receiving orthodontic treatment? *American Journal of Orthodontics and Dentofacial Orthopedics.* **139,** 369-377.

Aguilar-Diaz, F.-d.C. & Irigoyen-Camacho, M.-E. 2011. Validation of the CPQ(8-10ESP) in Mexican School children in urban areas. *Med Oral Patol Oral Cir Bucal.* **16,** E430-E435.

Ahn, Y.S., Kim, H.Y., Hong, S.M., et al. 2012. Validation of a Korean version of the Child Oral Health Impact Profile (COHIP) among 8- to 15-year-old school children. *International Journal of Paediatric Dentistry.* **22,** 292-301.

Ahovuo-Saloranta, A., Forss, H., Walsh, T., et al. 2013. Sealants for preventing dental decay in the permanent teeth. *The Cochrane Database of Systematic Reviews.* **3,** CD001830.

Alderson, P. 1995. *Listening to Children: Children and Social Research Ethics* London, Barnardos.

Allen, D.D., Ni, P. & Haley, S.M. 2008. Efficiency and sensitivity of multidimensional computerized adaptive testing of pediatric physical functioning. *Disability and Rehabilitation.* **30,** 479-484.

Alm, A., Wendt, L.K., Koch, G. & Birkhed, D. 2007. Prevalence of approximal caries in posterior teeth in 15-year-old Swedish teenagers in relation to their caries experience at 3 years of age. *Caries Research.* **41,** 392-398.

Alonso, Y. 2004. The biopsychosocial model in medical research: the evolution of the health concept over the last two decades. *Patient Education and Counseling.* **53,** 239-244.

Amato, J.N., Barbosa, T.S., Kobayashi, F.Y. & Gaviao, M.B. 2014. Changes in the oral-health-related quality of life of Brazilian children after an educational preventive programme: an 1-month longitudinal evaluation. *International Journal of Dental Hygiene.* **12,** 226-233.

American Association of Anesthesiologists. 1974. ASA Physical Classification System. American Association of Anesthesiologists. Available: http://www.asahq.org/resources/clinical-information/asa-physical-status-classification-system [Accessed 12th April 2015].

American Association of Pediatric Dentistry. 2008. Guideline on Pediatric Restorative Dentistry. AAPD. Available: http://www.aapd.org/media/Policies_Guidelines/G_Restorative.pdf [Accessed 15th April 2015].

Andrich, D., Lyne, A., Sheridon, B. & Luo, G. 2009. RUMM 2030. Perth: RUMM Laboratory.

Antunes, L.A., Andrade, M.R., Leao, A.T., Maia, L.C. & Luiz, R.R. 2013. Systematic review: change in the quality of life of children and adolescents younger than 14 years old after oral health interventions: a systematic review. *Pediatric Dentistry.* **35,** 37-42.

Asgari, I., Ahmady, A.E., Broder, H., Eslamipour, F. & Wilson-Genderson, M. 2013. Assessing the oral health-related quality of life in Iranian adolescents: validity of the Persian version of the Child Oral Health Impact Profile (COHIP). *Oral Health and Preventive Dentistry.* **11,** 147-154.

Asimakopoulou, K. & Newton, J.T. 2015. The contributions of behaviour change science towards dental public health practice: a new paradigm. *Community Dentistry and Oral Epidemiology.* **43,** 2-8.

Baker, S.R., Mat, A. & Robinson, P.G. 2010. What psychosocial factors influence adolescents' oral health? *Journal of Dental Research.* **89,** 1230-1235.

Barbosa, T.S., Tureli, M.C.M. & Gaviao, M.B.D. 2009. Validity and reliability of the Child Perceptions Questionnaires applied in Brazilian children. *BMC Oral Health.* **9,** 13.

Barbosa, T.S., Leme, M.S., Castelo, P.M. & Gaviao, M.B.D. 2011. Evaluating oral health-related quality of life measure for children and preadolescents with temporomandibular disorder. *Health Qual Life Outcomes.* **9,** 32.

Basavaraj, P., Sunil, M.K., Nagarajappa, R., Ashish, S. & Ramesh, G. 2013. Correlation Between Oral Health and Child-OIDP Index in 12- and 15-Year-Old Children From Modinagar, India. *Asia-Pacific Journal of Public Health.* **26,** 390-400.

Bastos, J.L., Peres, M.A., Peres, K.G., Araujo, C.L. & Menezes, A.M. 2008. Toothache prevalence and associated factors: a life course study from birth to age 12 yr. *European Journal of Oral Sciences.* **116,** 458-466.

Batcho, C.S., Tennant, A. & Thonnard, J.L. 2012. ACTIVLIM-Stroke: a crosscultural Rasch-built scale of activity limitations in patients with stroke. *Stroke.* **43,** 815-823.

Beales, J.G., Keen, J.H. & Holt, P.J. 1983. The child's perception of the disease and the experience of pain in juvenile chronic arthritis. *Journal of Rheumatology.* **10,** 61-65.

Beatty, P. & Willis, G. 2007. Research synthesis: The practice of of cognitive interviewing. *Public Opinion Quarterly.* **71,** 287-311.

Bedi, R. & Uppal, R.D.K. 1995. The oral health of minority ethnic communities in the United Kingdom. *British Dental Journal.* **179,** 421-425.

Bekes, K., John, M.T., Schaller, H.-G. & Hirsch, C. 2011. The German version of the Child Perceptions Questionnaire on oral health-related quality of life (CPQ-G11-14) Population-based norm values. *Journal of Orofacial Orthopedics.* **72,** 223-233.

Bendo, C.B., Paiva, S.M., Torres, C.S., et al. 2010. Association between treated/untreated traumatic dental injuries and impact on quality of life of Brazilian schoolchildren. *Health and Quality of Life Outcomes.* **8,** 114.

Benjamin, R.M. 2010. Oral health: the silent epidemic. *Public Health Reports.* **125,** 158-159.

Benson, P., O'Brien, C. & Marshman, Z. 2008. Agreement between mothers and children with malocclusion in rating children's oral health-related quality of life. *American Journal of Orthodontics and Dentofacial Orthopedics.* **137,** 631-638.

Beresford, B. 1997. *Personal Accounts: Involving Disabled Children in Research* London, The Stationery Office.

Bernabe, E., Tsakos, G. & Sheiham, A. 2007a. Intensity and extent of oral impacts on daily performances by type of self-perceived oral problems. *European Journal of Oral Sciences.* **115,** 111-116.

Bernabe, E., Flores-Mir, C. & Sheiham, A. 2007b. Prevalence, intensity and extent of Oral Impacts on Daily Performances associated with self-perceived malocclusion in 11-12-year-old children. *BMC Oral Health.* **7,** 6.

Bernabe, E., Sheiham, A. & Tsakos, G. 2008. A comprehensive evaluation of the validity of Child-OIDP: further evidence from Peru. *Community Dentistry & Oral Epidemiology.* **36,** 317-325.

Bernabe, E., Krisdapong, S., Sheiham, A. & Tsakos, G. 2009. Comparison of the discriminative ability of the generic and condition-specific forms of the Child-OIDP index: a study on children with different types of normative dental treatment needs. *Community Dentistry and Oral Epidemiology.* **37,** 155-162.

Bernays, S., Seeley, J., Rhodes, T. & Mupambireyi, Z. 2015. What am I 'living' with? Growing up with HIV in Uganda and Zimbabwe. *Sociolology of Health and Illness.* **37,** 270-283.

Bianco, A., Fortunato, L., Nobile, C.G.A. & Pavia, M. 2010. Prevalence and determinants of oral impacts on daily performance: results from a survey among school children in Italy. *European Journal of Public Health.* **20,** 595-600.

Bichenbach, J.E., Chatterji, S., Badley, E.M. & Ustun, T.B. 1999. Models of disablement, universalism and the international classification of impairments, disabilities and handicaps. *Social Science and Medicine.* **48,** 1173-1187.

Bille, B.S. 1962. Migraine in school children. A study of the incidence and short-term prognosis, and a clinical, psychological and electroencephalographic comparison between children with migraine and matched controls. *Acta Paediatrica. Supplement.* **136,** 1-151.

Bodecker, C.F. 1939. The Modified Dental Caries Index. *The Journal of the American Dental Association.* **26,** 1453-1460.

Borgers, N., de Leeuw, E. & Hox, J. 2000. Children as respondants in survey research: Cognitive development and response quality. *Bulletin de Methodologie Sociologiqe.* **66,** 60-75.

Bos, A. & Prahl, C. 2010. Oral health-related quality of life in Dutch children with cleft lip and/or palate. *Angle Orthodontist.* **81,** 865-871.

Bos, A., Hoogstraten, J. & Zentner, A. 2010. Perceptions of Dutch orthodontic patients and their parents on oral health-related quality of life. *Angle Orthodontist.* **80,** 367-372.

Bradley, C., Todd, C., Gorton, T., et al. 1999. The development of an individualized questionnaire measure of perceived impact of diabetes on quality of life: the ADDQoL. *Quality of Life Research.* **8,** 79-91.

Brazier, J., Czoski-Murray, C., Roberts, J., et al. 2008. Estimation of a preference-based index from a condition-specific measure: the King's Health Questionnaire. *Medical Decision Making.* **28,** 113-126.

Bridgman, C.M., Ashby, D. & Holloway, P.J. 1999. An investigation of the effects on children of tooth extraction under general anaesthesia in general dental practice. *British Dental Journal.* **186,** 245-247.

British Society of Paediatric Dentistry. 2015. Written evidence submitted by the British Society of Paediatric Dentistry in response to the House of Common's Select Committee inquiry into children's oral health in England. BSPD. Available: http://www.bspd.co.uk/LinkClick.aspx?fileticket=mLwjzp2AvrE%3d&tabid=40 [Accessed 22nd April 2015].

Brod, M., Tesler, L.E. & Christensen, T.L. 2009. Qualitative research and content validity: developing best practices based on science and experience. *Quality of Life Research.* **18,** 1263-1278.

Broder, H.L. 2007. Children's oral health-related quality of life. *Community Dentistry & Oral Epidemiology.* **35 Suppl 1,** 5-7.

Broder, H.L., McGrath, C. & Cisneros, G.J. 2007. Questionnaire development: Face validity and item impact testing of the Child Oral Health Impact Profile. *Community Dentistry and Oral Epidemiology.* **35,** 8-19.

Broder, H.L. & Wilson-Genderson, M. 2007. Reliability and convergent and discriminant validity of the Child Oral Health Impact Profile (COHIP Child's version). *Community Dentistry and Oral Epidemiology.* **35 Suppl 1,** 20-31.

Broder, H.L., Wilson-Genderson, M. & Sischo, L. 2012. Reliability and validity testing for the Child Oral Health Impact Profile-Reduced (COHIP-SF 19). *Journal of Public Health Dentistry.* **72,** 302-312.

Brosens, V., Ghijselings, I., Lemiere, J., et al. 2014. Changes in oral health-related quality of life reports in children during orthodontic treatment and the possible role of self-esteem: a follow-up study. *European Journal of Orthodontics.* **36,** 186-191.

Brown, A. & Al-Khayal, Z. 2006. Validity and reliability of the Arabic translation of the child oral-health-related quality of life questionnaire (CPQ11-14) in Saudi Arabia. *International Journal of Paediatric Dentistry.* **16,** 405-411.

Bullinger, M., Anderson, R., Cella, D. & Aaronson, N. 1993. Developing and evaluating cross-cultural instruments from minimum requirements to optimal models. *Quality of Life Research.* **2,** 451-459.

Calis, E.M., Geels, L.M., Prahl-Andersen, B. & Zentner, A. 2009. Oral health-related quality of life and dental esthetics in Amsterdam schoolchildren. *Journal of Dentistry for Children.* **76,** 130-135.

Campbell, J. 1975. Illness Is a Point of View: The Development of Children's Concepts of Illness. *Child Development.* 92-100.

Canaway, A.G. & Frew, E.J. 2013. Measuring preference-based quality of life in children aged 6-7 years: a comparison of the performance of the CHU-9D and EQ-5D-Y--the WAVES pilot study. *Quality of Life Research.* **22,** 173-183.

Cano, S.J. & Hobart, J.C. 2011. The problem with health measurement. *Patient Prefer Adherence.* **5,** 279-290.

Carlton, J. 2013a. Developing the draft descriptive system for the child amblyopia treatment questionnaire (CAT-Qol): a mixed methods study. *Health and Quality of Life Outcomes.* **11,** 174.

Carlton, J. 2013b. Identifying potential themes for the Child Amblyopia Treatment Questionnaire. *Optometry and Vision Science.* **90,** 867-873.

Carlton, J. 2013c. Refinement of the Child Amblyopia Treatment Questionnaire (CAT- QoL) using Rasch analysis. *HEDS Discussion Paper* [Online], Available: http://www.shef.ac.uk/polopoly_fs/1.321886!/file/13.13.pdf.

Castro, R.A.L., Cortes, M.I.S., Leao, A.T., et al. 2008. Child-OIDP index in Brazil: cross-cultural adaptation and validation. *Health & Quality of Life Outcomes.* **6,** 68.

Castro, R.d.A.L., Portela, M.C., Leao, A.T. & de Vasconcellos, M.T.L. 2011. Oral health-related quality of life of 11-and 12-year-old public school children in Rio de Janeiro. *Community Dentistry and Oral Epidemiology.* **39,** 336-344.

Chadwick, B.L., Gash, C. & Stewart, K. 2007. Preformed metal crowns: views of a group of dental practitioners in North Wales. *Primary Dental Care.* **14,** 140-144.

Chadwick, B.L. & Evans, D.J. 2007. Restoration of class II cavities in primary molar teeth with conventional and resin modified glass ionomer cements: a systematic review of the literature. *European Archives of Paediatric Dentistry.* **8,** 14-21.

Cheretakis, C., Locker, D., Dror, Y. & Glogauer, M. 2007. Oral health-related quality of life of children with neutropenia. *Special Care in Dentistry.* **27,** 6-11.

Chesters, R.K., Huntington, E., Burchell, C.K. & Stephen, K.W. 1992. Effect of oral care habits on caries in adolescents. *Caries Research.* **26,** 299-304.

Chien, C.W., Rodger, S. & Copley, J. 2014. Development and psychometric evaluation of a new measure for children's participation in hand-use life situations. *Archives of Physical Medicine and Rehabilitation.*

Chien, T.W., Wu, H.M., Wang, W.C., Castillo, R.V. & Chou, W. 2009. Reduction in patient burdens with graphical computerized adaptive testing on the ADL scale: tool development and simulation. *Health and Quality of Life Outcomes.* **7,** 39.

Children's Commissioner. 2013. Participation Strategy. Ensuring the voice of children and young people in the work of the Office of the Children's Commissioner. London: Office of the Children's Commissioner. Available: http://www.childrenscommissioner.gov.uk/content/publications/content_675 [Accessed 3rd March 2015].

Chow, M.Y., Morrow, A.M., Cooper Robbins, S.C. & Leask, J. 2013. Condition-specific quality of life questionnaires for caregivers of children with pediatric conditions: a systematic review. *Quality of Life Research.* **22,** 2183-2200.

Christensen, P. & James, A. 2000. *Research with Children: Perspectives and Practices.* London, Falmer Press.

Christie, M.J., French, D., Sowden, A. & West, A. 1993. Development of child-centered disease-specific questionnaires for living with asthma. *Psychosomatic Medicine.* **55,** 541-548.

Clarke, M., Locker, D., Berall, G., et al. 2006. Malnourishment in a population of young children with severe early childhood caries. *Pediatr Dent.* **28,** 254-259.

Cocks, A. 2006. The ethical maze: Finding an inclusive path towards gaining children's agreement to research participation. *Childhood.* **13,** 247-266.

Conway, D.I., Quarrell, I., McCall, D.R., et al. 2007. Dental caries in 5-year-old children attending multi-ethnic schools in Greater Glasgow--the impact of ethnic background and levels of deprivation. *Community Dental Health.* **24,** 161-165.

Cortes-Martinicorena, F.-J., Rosel-Gallardo, E., Artazcoz-Oses, J., Bravo, M. & Tsakos, G. 2010. Adaptation and validation for Spain of the Child-Oral Impact on Daily Performance (C-OIDP) for use with adolescents. *Medicina Oral Patologia Oral y Cirugia Bucal.* **15,** e106-111.

Cosco, T.D., Doyle, F., Ward, M. & McGee, H. 2012. Latent structure of the Hospital Anxiety And Depression Scale: a 10-year systematic review. *Journal of Psychosomatic Research.* **72,** 180-184.

Costa, A., Ferreira, M.C., Serra-Negra, J.M., Pordeus, I.A. & Paiva, S.M. 2011. Impact of wearing fixed orthodontic appliances on oral health-related quality of life among Brazilian children. *Journal of Orthodontics.* **38,** 275-281.

Coster, W.J., Haley, S.M., Ni, P., Dumas, H.M. & Fragala-Pinkham, M.A. 2008. Assessing self-care and social function using a computer adaptive testing version of the pediatric evaluation of disability inventory. *Archives of Physical Medicine and Rehabilitation.* **89,** 622-629.

Crow, C.S. 1997. Children's Pain Perspectives Inventory (CPPI): developmental assessment. *Pain.* **72,** 33-40.

Cunnion, D.T., Spiro, A., 3rd, Jones, J.A., et al. 2010. Pediatric oral health-related quality of life improvement after treatment of early childhood caries: a prospective multisite study. *Journal of Dentistry for Children.* **77,** 4-11.

Cushing, A.M., Sheiham, A. & Maizels, J. 1986. Developing socio-dental indicators--the social impact of dental disease. *Community Dental Health.* **3,** 3-17.

Davis, J. 1998. Understanding the meanings of children: a reflexive process. *Children and Society.* **12,** 325-335.

de Haes, J.C.J.M., Olschewski, M., Fayers, P.M., et al. 1996. The Rotterdam Symptom Checklist (RSCL): A Manual,. Gronigen: Northern Centre for Healthcare Research. Available: https://www.umcg.nl/SiteCollectionDocuments/research/institutes/SHARE/assessment_tools/handleiding_rscl2edruk.pdf [Accessed 20th April 2015].

de Oliveira, C.M., Sheiham, A., Tsakos, G. & O'Brien, K.D. 2008. Oral health-related quality of life and the IOTN index as predictors of children's perceived needs and acceptance for orthodontic treatment. *British Dental Journal.* **204,** 1-5; discussion 384-385.

de Paula, J.S., Sarracini, K.L., Meneghim, M.C., et al. 2015. Longitudinal evaluation of the impact of dental caries treatment on oral health-related quality of life among schoolchildren. *European Journal of Oral Sciences.* **123,** 173-178.

Department of Health. 2004. National Service Framework for Children, Young People and Maternity Services: Key Issues for Primary Care. London: DH. Available: https://www.gov.uk/government/publications/national-service-framework-children-young-people-and-maternity-services [Accessed 20th April 2015].

Department of Health. 2007. You're Welcome quality criteria: making health services young people friendly. London: DH. Available: https://www.gov.uk/government/publications/quality-criteria-for-young-people-friendly-health-services [Accessed 20th April 2015].

Department of Health. 2010. Achieving Equity and Excellence for Children. London: DH. Available: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_119449 [Accessed 20th April 2015].

Department of Health. 2014. Reference costs 2013 to 2014. London: DH. Available: https://www.gov.uk/government/publications/nhs-reference-costs-2013-to-2014 [Accessed 26th February 2015].

Department of Health. 2015. Dental Contract Reform: Prototypes. Overview Document. London: DH. Available:

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/395384/Reform_Document.pdf [Accessed 13th April 2015].

Detmar, S.B., Bruil, J., Ravens-Sieberer, U., et al. 2006. The use of focus groups in the development of the KIDSCREEN HRQL questionnaire. *Quality of Life Research.* **15,** 1345-1353.

Devine, J., Otto, C., Rose, M., et al. 2014. A new computerized adaptive test advancing the measurement of health-related quality of life (HRQoL) in children: the Kids-CAT. *Quality of Life Research.* **24,** 871-874.

Do, L.G. & Spencer, A. 2007. Oral health-related quality of life of children by dental caries and fluorosis experience. *Journal of Public Health Dentistry.* **67,** 132-139.

Do, L.G. & Spencer, A.J. 2008. Evaluation of oral health-related quality of life questionnaires in a general child population. *Community Dent Health.* **25,** 205-210.

Do, L.G. 2012. Distribution of caries in children: variations between and in populations. *Journal of Dental Research.* **91,** 536-543.

Doherty, E., Yanni, G., Conroy, R.M. & Bresnihan, B. 1993. A comparison of child and parent ratings of disability and pain in juvenile chronic arthritis. *Journal of Rheumatology.* **20,** 1563-1566.

Dumitrache, M.A., Comes, C., Teodorescu, E., et al. 2009. Life quality related to oral health of schoolchildren from Bucharest. *Revista Romana De Bioetica.* **7,** 169-178.

Dunlow, N., Phillips, C. & Broder, H.L. 2007. Concurrent validity of the COHIP. *Community Dentistry & Oral Epidemiology.* **35,** 41-49.

Dye, B.A., Tan, S., Smith, V., et al. 2007. Trends in oral health status: United States, 1988-1994 and 1999-2004. *Vital and Health Statistics. Series 11: Data from the National Health Survey.* 1-92.

Edelen, M.O. & Reeve, B.B. 2007. Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. *Quality of Life Research.* **16 Suppl 1,** 5-18.

Edit Central. Available: http://www.editcentral.com/gwt1/EditCentral.html [Accessed 29th January 2015].

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

Eiser, C. & Varni, J.W. 2013. Health-related quality of life and symptom reporting: similarities and differences between children and their parents. *European Journal of Pediatrics.* **172,** 1299-1304.

Eiser, C. & Kopel, S. 2013. *Perceptions of Health & Illness* New York, Routledge.

Embretson, S. & Reise, S. 2013. *Item Response Theory for Psychologists.* New Jersey, Lawrence Erbaulm Associates.

Engel, G.L. 1977. The need for a new medical model: a challenge for biomedicine. *Science.* **196,** 129-136.

Evans, A.M., Scutter, S.D. & 2004. Prevalence of "growing pains" in young children. *Journal of Pediatrics.* **145,** 255-258.

Fakhruddin, K.S., Lawrence, H.P., Kenny, D.J. & Locker, D. 2008. Impact of treated and untreated dental injuries on the quality of life of Ontario school children. *Dental Traumatology.* **24,** 309-313.

Fayers, P. & Machin, D. 2006. *Quality of Life: The Assessment, Analysis and Interpretation of Patient-Reported Outcomes* Chichester, England, Wiley.

Fayers, P.M., Curran, D. & Machin, D. 1998. Incomplete quality of life data in randomized trials: missing items. *Statistics in Medicine.* **17,** 679-696.

Feitosa, S., Colares, V. & Pinkham, J. 2005. The psychosocial effects of severe caries in 4-year-old children in Recife, Pernambuco, Brazil. *Cadernos de Saúde Publica.* **21,** 1550-1556.

Fergus, C.E. 2010. Caries prevalence and experience of 12-year old children in Montserrat. *West Indian Medical Journal.* **59,** 573-577.

Ferreira, M.C., Goursand, D., Bendo, C.B., et al. 2012. Agreement between adolescents' and their mothers' reports of oral health-related quality of life. *Brazilian Oral Research.* **26,** 112-118.

Fisher Jr, W. 1992. Reliability, separation, strata statistics. *Rasch Measurement Transactions* [Online],**6**. Available: http://www.rasch.org/rmt/rmt63i.htm [Accessed 15th January 2015].

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

Foster Page, L., Thomson, W. & Locker, D. 2010. Assessing the responsiveness of the CPQ11-14 in New Zealand adolescents. *Social Science in Dentistry.* **1,** 48-53.

Foster Page, L.A., Thomson, W.M., Jokovic, A. & Locker, D. 2005. Validation of the Child Perceptions Questionnaire (CPQ 11-14). *Journal of Dental Research.* **84,** 649-652.

Foster Page, L.A., Thomson, W.M., Jokovic, A. & Locker, D. 2008. Epidemiological evaluation of short-form versions of the Child Perception Questionnaire. *European Journal of Oral Sciences.* **116,** 538-544.

Foster Page, L.A. & Thomson, W.M. 2012. Caries prevalence, severity, and 3-year increment, and their impact upon New Zealand adolescents' oral-health-related quality of life. *Journal of Public Health Dentistry.* **72,** 287-294.

Foster Page, L.A., Boyd, D. & Thomson, W.M. 2013a. Do we need more than one Child Perceptions Questionnaire for children and adolescents? *BMC Oral Health.* **13,** 26.

Foster Page, L.A., Thomson, W.M., Ukra, A. & Farella, M. 2013b. Factors influencing adolescents' oral health-related quality of life (OHRQoL). *International Journal of Paediatric Dentistry.* **23,** 415-423.

Franck, L., Noble, G. & Liossi, C. 2010. Translating the tears: parents' use of behavioural cues to detect pain in normally developing young children with everyday minor illnesses or injuries. *Child: Care, Health and Development.* **36,** 895-904.

Freeman, R., Whelton, H. & Gibson, B. 2010. Toothbrushing rules: power dynamics and toothbrushing in children. *Social Science and Dentistry.* **1,** 37-47.

French, D.J., Christie, M.J. & Sowden, A.J. 1994. The reproducibility of the Childhood Asthma Questionnaires: measures of quality of life for children with asthma aged 4-16 years. *Quality of Life Research.* **3,** 215-224.

Gaffney, A. & Dunne, E.A. 1986. Developmental aspects of children's definitions of pain. *Pain.* **26,** 105-117.

Geels, L.M., Hoogstraten, J. & Prahl-Andersen, B. 2008a. Confirmative factor analysis of the dimensions of the Child Oral Health Impact Profile (Dutch version). *European Journal of Oral Sciences.* **116,** 148-152.

Geels, L.M., Kieffer, J.M., Hoogstraten, J. & Prahl-Andersen, B. 2008b. Oral health-related quality of life of children with craniofacial conditions. *Cleft Palate-Craniofacial Journal.* **45,** 461-467.

Gherunpong, S., Tsakos, G. & Sheiham, A. 2004a. Developing and evaluating an oral health-related quality of life index for children; the CHILD-OIDP. *Community Dental Health.* **21,** 161-169.

Gherunpong, S., Tsakos, G. & Sheiham, A. 2004b. The prevalence and severity of oral impacts on daily performances in Thai primary school children. *Health and Quality of Life Outcomes.* **2,** 57.

Gherunpong, S., Tsakos, G. & Sheiham, A. 2006a. A socio-dental approach to assessing children's orthodontic needs. *European Journal of Orthodontics.* **28,** 393-399.

Gherunpong, S., Sheiham, A. & Tsakos, G. 2006b. A sociodental approach to assessing children's oral health needs: integrating an oral health-related quality of life (OHRQoL) measure into oral health sevice planning. *Bulletin of the World Health Organization.* **84,** 36-42.

Gill, T.M. & Feinstein, A.R. 1994. A critical appraisal of the quality of quality-of-life measurements. *JAMA.* **272,** 619-626.

Glasgow Centre for the Child and Society. undated. Code of Practice on Research Ethics: Responsibilities to Research Participants. Glasgow: Glasgow Centre for the Child and Society. Available: http://www.strath.ac.uk/humanities/socialwork/glasgowcentreforthechildandsociety/publications/ [Accessed 26th March 2015].

Goodwin, M., Sanders, C., Davies, G., Walsh, T. & Pretty, I.A. 2015. Issues arising following a referral and subsequent wait for extraction under general anaesthetic: impact on children. *BMC Oral Health.* **15,** 3.

Goursand, D., Paiva, S.M., Zarzar, P.M., et al. 2008. Cross-cultural adaptation of the Child Perceptions Questionnaire 11-14 (CPQ11-14) for the Brazilian Portuguese language. *Health & Quality of Life Outcomes.* **6,** 2.

Greaves, C.J., Sheppard, K.E., Abraham, C., et al. 2011. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. *BMC Public Health.* **11,** 119.

Griffiths, A.M., Nicholas, D., Smith, C., et al. 1999. Development of a quality-of-life index for pediatric inflammatory bowel disease: dealing with differences related to age and IBD type. *Journal of Pediatric Gastroenterology and Nutrition.* **28,** S46-52.

Grimsby, G. 1998. Useful reporting of DIF. *Rasch Measurement Transactions.* **12,** 651.

Gururatana, O., Baker, S. & Robinson, P.G. 2011. Psychometric properties of long and short forms of the Child Perceptions Questionnaire (CPQ11-14) in a Thai population. *Community Dental Health.* **28,** 232-237.

Guttman, L. 1950. *The basis for Scalogram analysis.* Princeton, Princeton University Press.

Guyatt, G., Walter, S. & Norman, G. 1987a. Measuring change over time: assessing the usefulness of evaluative instruments. *Journal of Chronic Diseases.* **40,** 171-178.

Guyatt, G.H., Berman, L.B., Townsend, M. & Taylor, D.W. 1985. Should study subjects see their previous responses? *Journal of Chronic Diseases.* **38,** 1003-1007.

Guyatt, G.H., Bombardier, C. & Tugwell, P.X. 1986. Measuring disease-specific quality of life in clinical trials. *CMAJ: Canadian Medical Association Journal.* **134,** 889-895.

Guyatt, G.H., Berman, L.B., Townsend, M., Pugsley, S.O. & Chambers, L.W. 1987b. A measure of quality of life for clinical trials in chronic lung disease. *Thorax.* **42,** 773-778.

Guyatt, G.H., Veldhuyzen Van Zanten, S.J., Feeny, D.H. & Patrick, D.L. 1989. Measuring quality of life in clinical trials: a taxonomy and review. *CMAJ: Canadian Medical Association Journal.* **140,** 1441-1448.

Guyatt, G.H., Feeny, D.H. & Patrick, D.L. 1993. Measuring health-related quality of life. *Annals of Internal Medicine.* **118,** 622-629.

Guyatt, G.H. & Cook, D.J. 1994. Health status, quality of life, and the individual. *JAMA.* **272,** 630-631.

Haley, S.M., Fragala-Pinkham, M.A., Dumas, H.M., et al. 2009. Evaluation of an item bank for a computerized adaptive test of activity in children with cerebral palsy. *Physical Therapy.* **89,** 589-600.

Hall, M., Gibson, B., James, A. & Rodd, H.D. 2012. Children's experiences of participation in the cleft lip and palate care pathway. *International Journal of Paediatric Dentistry.* **22,** 442-450.

Harbeck, C. & Peterson, L. 1992. Elephants dancing in my head: a developmental approach to children's concepts of specific pains. *Child Development.* **63,** 138-149.

Harden, J., Scott, S., Beckett-Milburn, K. & Jackson, S. 2000. Can't talk, won't talk: methodological issues in researching children. *Sociological Research Online* [Online], **5**. Available: http://www.socresonline.org.uk/5/2/harden.html [Accessed 8th January 2013].

Harman, K., Lindsay, S., Adewami, A. & Smith, P. 2005. An investigation of language used by children to describe discomfort expected and experienced during dental treatment. *International Journal of Paediatric Dentistry.* **15,** 319-326.

Hart, R. 1992. Children's Participation: From Tokenism to Citizenship. Florence, Italy: UNICEF International Children's Development Centre. Available: http://www.unicef-irc.org/publications/pdf/childrens_participation.pdf [Accessed 3rd March 2015].

Hayden, C., Bowler, J.O., Chambers, S., et al. 2013. Obesity and dental caries in children: a systematic review and meta-analysis. *Community Dentistry and Oral Epidemiology.* **41,** 289-308.

Haywood, K.L., Collin, S.M. & Crawley, E. 2014. Assessing severity of illness and outcomes of treatment in children with Chronic Fatigue Syndrome/Myalgic Encephalomyelitis (CFS/ME): a systematic review of patient-reported outcome measures (PROMs). *Child: Care, Health and Development.* **40,** 806-824.

Helseth, S. & Slettebo, A. 2004. Research involving children: some ethical issues. *Nursing Ethics.* **11,** 298-308.

Herrman, J. 2006. Children's and Young Adolescents' Voices: Perceptions of the Costs and Rewards of Diabetes and Its Treatment. *Journal of Pediatric Nursing.* **21,** 211-221.

Holland, P. & Wainer, H. 1993. *Differential Item Functioning* NJ; Hillsdale, Lawrence Erlbaum.

Holland, S., Renold, E., Ross, N. & Hillman, A. 2008. The everyday lives of children in care: using a sociological perspective to inform social work practice. ESRC National Centre for Research Methods. Available: http://eprints.ncrm.ac.uk/466/ [Accessed 21st January 2015].

Hooley, M., Skouteris, H., Boganin, C., Satur, J. & Kilpatrick, N. 2012. Body mass index and dental caries in children and adolescents: a systematic review of literature published 2004 to 2011. *Systematic Reviews.* **1,** 57.

Hoppe, M., Wells, E., Morrison, D., Gillmore, M. & Wilsdon, A. 1995. Using focus groups to discuss sensitive topics with children. *Evaluation Review.* 102-114.

Hosey, M.T., Macpherson, L.M., Adair, P., et al. 2006. Dental anxiety, distress at induction and postoperative morbidity in children undergoing tooth extraction using general anaesthesia. *British Dental Journal.* **200,** 39-43; discussion 27; quiz 50.

Hospital Episode Statistics. 2015. Admitted Patient Care, England - 2013-14. Health and Social Care Information Centre. Available: http://www.hscic.gov.uk/searchcatalogue?productid=17192&q=title%3a%22Hospital+Episode+Statistics%2c+Admitted+patient+care+-+England%22&sort=Relevance&size=10&page=1 - top [Accessed 26th February 2015].

Humphris, G., Freeman, R., Gibson, B., Simpson, K. & Whelton, H. 2005. Oral health-related quality of life for 8-10-year-old children: an assessment of a new measure. *Community Dentistry and Oral Epidemiology.* **33,** 326-332.

IASP Task Force on Taxonomy 1994. *Part III: Pain Terms, A Current List with Definitions and Notes on Usage* Seattle, IASP Press.

Innes, N.P., Evans, D.J. & Stirrups, D.R. 2007. The Hall Technique; a randomized controlled clinical trial of a novel method of managing carious primary molars in general dental practice: acceptability of the technique and outcomes at 23 months. *BMC Oral Health.* **7,** 18.

Innes, N.P., Evans, D.J. & Stirrups, D.R. 2011. Sealing caries in primary molars: randomized control trial, 5-year results. *Journal of Dental Research.* **90,** 1405-1410.

Innes, N.P., Clarkson, J.E., Speed, C., et al. 2013. The FiCTION dental trial protocol - filling children's teeth: indicated or not? *BMC Oral Health.* **13,** 25.

International Caries Detection and Assessment System Coordinating Committee. 2011. Rationale and Evidence for the International Caries Detection and Assessment System (ICDAS II). International Caries Detection and Assessment System Coordinating Committee. Available: https://www.icdas.org/uploads/Rationale and Evidence ICDAS II September 11-1.pdf [Accessed 5th October 2015].

Iranzo-Cortes, J.E., Montiel-Company, J.M. & Almerich-Silla, J.M. 2013. Caries diagnosis: agreement between WHO and ICDAS II criteria in epidemiological surveys. *Community Dental Health.* **30,** 108-111.

Jackson, S.L., Vann, W.F., Jr., Kotch, J.B., Pahel, B.T. & Lee, J.Y. 2011. Impact of poor oral health on children's school attendance and performance. *American Journal of Public Health.* **101,** 1900-1906.

James, A., Jenks, C. & Prout, A. 1998. *Theorizing Childhood* Cambridge, Polity Press.

Jankauskiene, B. & Narbutaite, J. 2010. Changes in oral health-related quality of life among children following dental treatment under general anaesthesia. A systematic review. *Stomatologija.* **12,** 60-64.

Jerrett, M. & Evans, K. 1986. Children's pain vocabulary. *Journal of Advanced Nursing.* **11,** 403-408.

Jokovic, A., Locker, D., Stephens, M., et al. 2002. Validity and reliability of a questionnaire for measuring child oral-health-related quality of life. *Journal of Dental Research.* **81,** 459-463.

Jokovic, A., Locker, D., Stephens, M. & Guyatt, G. 2003a. Agreement between mothers and children aged 11-14 years in rating child oral health-related quality of life. *Community Dentistry and Oral Epidemiology.* **31,** 335-343.

Jokovic, A., Locker, D., Stephens, M., et al. 2003b. Measuring parental perceptions of child oral health-related quality of life. *Journal of Public Health Dentistry.* **63,** 67-72.

Jokovic, A., Locker, D., Tompson, B. & Guyatt, G. 2004. Questionnaire for measuring oral health-related quality of life in eight- to ten-year-old children. *Pediatric Dentistry.* **26,** 512-518.

Jokovic, A., Locker, D. & Guyatt, G. 2005. What do children's global ratings of oral health and well-being measure? *Community Dentistry and Oral Epidemiology.* **33,** 205-211.

Jokovic, A., Locker, D. & Guyatt, G. 2006. Short forms of the Child Perceptions Questionnaire for 11-14-year-old children (CPQ11-14): development and initial evaluation. *Health and Quality of Life Outcomes.* **4,** 4.

Juniper, E.F. 1997a. How important is quality of life in pediatric asthma? *Pediatric Pulmonology. Supplement.* **15,** 17-21.

Juniper, E.F., Guyatt, G.H., Feeny, D.H., Griffith, L.E. & Ferrie, P.J. 1997. Minimum skills required by children to complete health-related quality of life instruments for asthma: comparison of measurement properties. *European Respiratory Journal.* **10,** 2285-2294.

Juniper, E.F. 1997b. Quality of life in adults and children with asthma and rhinitis. *Allergy.* **52,** 971-977.

Kadkhoda, S., Nedjat, S. & Shirazi, M. 2011. Comparison of oral-health-related quality of life during treatment with headgear and functional appliances. *International Journal of Paediatric Dentistry.* **21,** 369-373.

Kassebaum, N.J., Bernabe, E., Dahiya, M., et al. 2015. Global Burden of Untreated Caries: A Systematic Review and Metaregression. *Journal of Dental Research.*

Kersten, P., White, P.J. & Tennant, A. 2014. Is the pain visual analogue scale linear and responsive to change? An exploration using Rasch analysis. *PloS One.* **9,** e99485.

Kindelan, S.A., Day, P., Nichol, R., et al. 2008. UK National Clinical Guidelines in Paediatric Dentistry: stainless steel preformed crowns for primary molars. *International Journal of Paediatric Dentistry.* **18 Suppl 1,** 20-28.

Kirk, S. 2007. Methodological and ethical issues in conducting qualitative research with children and young people: a literature review. *International Journal of Nursing Studies.* **44,** 1250-1260.

Knoema. 2015. *World GDP Ranking 2015 | Data and Charts* [Online]. Knoema. Available: http://knoema.com/nwnfkne/world-gdp-ranking-2015-data-and-charts [Accessed 5th October 2015].

Kohli, R., Levy, S., Kummet, C.M., Dawson, D.V. & Stanford, C.M. 2011. Comparison of perceptions of oral health-related quality of life in adolescents affected with ectodermal dysplasias relative to caregivers. *Special Care in Dentistry.* **31,** 88-94.

Kok, Y.V., Mageson, P., Harradine, N.W.T. & Sprod, A.J. 2004. Comparing a quality of life measure and the Aesthetic Component of the Index of Orthodontic Treatment Need (IOTN) in assessing orthodontic treatment need and concern. *Journal of Orthodontics.* **31,** 312-318; discussion 300-311.

Kolawole, K.A., Otuyemi, O.D. & Oluwadaisi, A.M. 2011. Assessment of oral health-related quality of life in Nigerian children using the Child Perceptions Questionnaire (CPQ 11-14). *European Journal of Paediatric Dentistry.* **12,** 55-59.

Koller, D., Khan, N. & Barrett, S. 2015. Pediatric perspectives on diabetes self-care: a process of achieving acceptance. *Qualitative Health Research.* **25,** 264-275.

Koposova, N., Widstrom, E., Eisemann, M., Koposov, R. & Eriksen, H.M. 2010. Oral health and quality of life in Norwegian and Russian school children: A pilot study. *Stomatologija.* **12,** 10-16.

Koposova, N., Eriksen, H.M., Widstram, E., et al. 2012. Oral health-related quality of life among 12-year-olds in Northern Norway and North-West Russia. *Oral Health and Dental Management.* **11,** 206-214.

Kressin, N.R. (ed.) 1997. *The Oral Health Related Quality of Life Measure (OHRQOL),* Chapel Hill: University of North Carolina, Dental Ecology.

Krisdapong, S., Sheiham, A. & Tsakos, G. 2009. Oral health-related quality of life of 12- and 15-year-old Thai children: findings from a national survey. *Community Dentistry and Oral Epidemiology.* **37,** 509-517.

Krisdapong, S., Prasertsom, P., Rattanarangsima, K., Sheiham, A. & Tsakos, G. 2012a. The impacts of gingivitis and calculus on Thai children's quality of life. *Journal of Clinical Periodontology.* **39,** 834-843.

Krisdapong, S., Sheiham, A. & Tsakos, G. 2012b. Impacts of recurrent aphthous stomatitis on quality of life of 12- and 15-year-old Thai children. *Quality of Life Research.* **21,** 71-76.

Krisdapong, S., Prasertsom, P., Rattanarangsima, K. & Sheiham, A. 2012c. Relationships between oral diseases and impacts on Thai schoolchildren's quality of life: evidence from a Thai national oral health survey of 12- and 15-year-olds. *Community Dentistry and Oral Epidemiology.* **40,** 550-559.

Krisdapong, S., Prasertsom, P., Rattanarangsima, K., Adulyanon, S. & Sheiham, A. 2012d. Using associations between oral diseases and oral health-related quality of life in a nationally representative sample to propose oral health goals for 12-year-old children in Thailand. *International Dental Journal.* **62,** 320-330.

Krisdapong, S., Prasertsom, P., Rattanarangsima, K. & Sheiham, A. 2013a. Impacts on quality of life related to dental caries in a national representative sample of Thai 12- and 15-year-olds. *Caries Research.* **47,** 9-17.

Krisdapong, S., Prasertsom, P., Rattanarangsima, K. & Sheiham, A. 2013b. School absence due to toothache associated with sociodemographic factors, dental caries status, and oral health-related quality of life in 12- and 15-year-old Thai children. *Journal of Public Health Dentistry.* **73,** 321-328.

Kröner-Herwig, B., Heinrich, M. & Morris, L. 2007. Headache in German children and adolescents: a population-based epidemiological study. *Cephalalgia.* **27,** 519-527.

Krosnick, J. 1991. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology.* **5,** 213-236.

Krosnick, J.A. 1999. Survey research. *Annual Review of Psychology.* **50,** 537-567.

Krueger , R. & Casey, M. 2000. *Focus Groups. A Practical Guide for Applied Research.* Thousand Oaks, CA, SAGE publications.

Kruger, E., Dyson, K. & Tennant, M. 2006. Hospitalization of Western Australian children for oral health related conditions: a 5-8 year follow-up. *Australian Dental Journal.* **51,** 231-236.

Kumar, S., Goyal, A., Tadakamadla, J., et al. 2011. Oral health related quality of life among children with parents and those with no parents. *Community Dental Health.* **28,** 227-231.

Kumar, S., Kroon, J. & Lalloo, R. 2014. A systematic review of the impact of parental socio-economic status and home environment characteristics on children's oral health related quality of life. *Health and Quality of Life Outcomes.* **12,** 41.

Kwakkenbos, L., van den Hoogen, F.H., Custers, J., et al. 2012. Validity of the Fear of Progression Questionnaire-Short Form in patients with systemic sclerosis. *Arthritis Care & Research.* **64,** 930-934.

Lader, D., Chadwick, B., Chestnutt, I.G., et al. 2004. Children's Dental Health in the United Kingdom 2003. Office for National Statistics. Available:

http://www.ons.gov.uk/.../health.../dental-health/dental-health...children/cdh-su... [Accessed 20th April 2015].

Laing, E., Cunningham, S.J., Jones, S., Moles, D. & Gill, D. 2010. Psychosocial impact of hypodontia in children. *American Journal of Orthodontics and Dentofacial Orthopedics.* **137,** 35-41.

Lau, A.W.H., Wong, M.C.M., Lam, K.F. & McGrath, C. 2009. Confirmatory factor analysis on the health domains of the Child Perceptions Questionnaire. *Community Dentistry and Oral Epidemiology.* **37,** 163-170.

Leidy, N.K. & Vernon, M. 2008. Perspectives on patient-reported outcomes : content validity and qualitative research in a changing clinical trial environment. *Pharmacoeconomics.* **26,** 363-370.

Levine, R.S., Pitts, N.B. & Nugent, Z.J. 2002. The fate of 1,587 unrestored carious deciduous teeth: a retrospective general dental practice based study from northern England. *British Dental Journal.* **193,** 99-103.

Levine, R.S., Nugent, Z.J. & Pitts, N.B. 2003. Pain prediction for preventive non-operative management of dentinal caries in primary teeth in general dental practice. *British Dental Journal.* **195,** 202-206; discussion 197.

Lewis, C. & Stout, J. 2010. Toothache in US children. *Archives of Pediatrics and Adolescent Medicine.* **164,** 1059-1063.

Li, L.W., Wong, H.M., Peng, S.M. & McGrath, C.P. 2015. Anthropometric measurements and dental caries in children: a systematic review of longitudinal studies. *Advances in Nutrition.* **6,** 52-63.

Linacre, J.M. 1994. Sample size and item calibration [or person measure] stability. *Rasch Measurement Transactions* [Online], **7**. Available: http://www.rasch.org/rmt/rmt74m.htm [Accessed 12th Januay 2015].

Liossi, C., Noble, G. & Franck, L.S. 2012. How parents make sense of their young children's expressions of everyday pain: A qualitative analysis. *European Journal of Pain (London, England).*

Locker, D. 1992. The burden of oral disorders in a population of older adults. *Community Dental Health.* **9,** 109-124.

Locker, D. & Slade, G. 1994. Association between clinical and subjective indicators of oral health status in an older adult population. *Gerodontology.* **11,** 108-114.

Locker, D., Clarke, M. & Payne, B. 2000. Self-perceived oral health status, psychological well-being and life satisfaction in an older adult population. *Journal of Dental Research.* **79,** 970-975.

Locker, D., Jokovic, A. & Tompson, B. 2005. Health-related quality of life of children aged 11 to 14 years with orofacial conditions. *Cleft Palate-Craniofacial Journal.* **42,** 260-266.

Locker, D. 2007. Disparities in oral health-related quality of life in a population of Canadian children. *Community Dentistry & Oral Epidemiology.* **35,** 348-356.

Locker, D., Jokovic, A., Tompson, B. & Prakash, P. 2007. Is the Child Perceptions Questionnaire for 11-14 year olds sensitive to clinical and self-perceived variations in orthodontic status? *Community Dentistry & Oral Epidemiology.* **35,** 179-185.

Locker, D. & Allen, F. 2007. What do measures of 'oral health-related quality of life' measure? *Community Dentistry and Oral Epidemiology.* **35,** 401-411.

Locker, D., Jokovic, A., Prakash, P. & Tompson, B. 2010. Oral health-related quality of life of children with oligodontia. *International Journal of Paediatric Dentistry.* **20,** 8-14.

Lohr, K.N., Aaronson, N.K., Alonso, J., et al. 1996. Evaluating quality-of-life and health status instruments: development of scientific review criteria. *Clinical Therapeutics.* **18,** 979-992.

Lohr, K.N. 2002. Assessing health status and quality-of-life instruments: attributes and review criteria. *Quality of Life Research.* **11,** 193-205.

Long, R. 2015. *Briefing Paper. School meals and nutritional standards* [Online]. London: The House of Commons Library. Available: http://www.parliament.uk/briefing-papers/SN04195.pdf [Accessed 8th October 2015].

Low, W., Tan, S. & Schwartz, S. 1999. The effect of severe caries on the quality of life in young children. *Pediatric Dentistry.* **21,** 325-326.

Ludwig, K.H., Fontana, M., Vinson, L.A., Platt, J.A. & Dean, J.A. 2014. The success of stainless steel crowns placed with the Hall technique: a retrospective study. *Journal of the American Dental Association.* **145,** 1248-1253.

Luoto, A., Lahti, S., Nevanpera, T., Tolvanen, M. & Locker, D. 2009. Oral-health-related quality of life among children with and without dental fear. *International Journal of Paediatric Dentistry.* **19,** 115-120.

Macpherson, L., Ball, G., Conway, I., et al. 2012. Report of the 2012 detailed National Dental Inspection Programme of Primary 1 children and the basic inspection of Primary 1 and Primary 7 children. ISD Scotland. Available: http://www.isdscotland.org/Health-Topics/Dental-Care/Publications/2012-11-27/2012-11-27-DentalNDIP-Report.pdf?41798037291 [Accessed 25th February 2015].

Macpherson, L.M., Ball, G.E., Brewster, L., et al. 2010. Childsmile: the national child oral health improvement programme in Scotland. Part 1: Establishment and development. *British Dental Journal.* **209,** 73-78.

Macpherson, L.M., Anopa, Y., Conway, D.I. & McMahon, A.D. 2013. National supervised toothbrushing program and dental decay in Scotland. *Journal of Dental Research.* **92,** 109-113.

Majewski, R.F., Snyder, C.W. & Bernat, J.E. 1988. Dental emergencies presenting to a children's hospital. *ASDC Journal of Dentistry for Children.* **55,** 339-342.

Malam, S., Clegg, S., Kirwan, S. & McGinigal, S. 2009. Comprehension and use of UK nutrition signpost labelling schemes. London: Prepared for the Food Standards Agency by British Market Research Bureau. Available: http://webarchive.nationalarchives.gov.uk/20131104005023/http://www.food.gov.uk/multimedia/pdfs/pmpreport.pdf [Accessed 8th October 2015].

Malliaropoulos, N., Korakakis, V., Christodoulou, D., et al. 2014. Development and validation of a questionnaire (FASH--Functional Assessment Scale for Acute Hamstring Injuries): to measure the severity and impact of symptoms on function and sports ability in patients with acute hamstring injuries. *British Journal of Sports Medicine.* **48,** 1607-1612.

Mallinson, S. 2002. Listening to respondents: a qualitative assessment of the Short-Form 36 Health Status Questionnaire. *Social Science and Medicine.* **54,** 11-21.

Marais, I. & Andrich, D. 2008. Effects of varying magnitude and patterns of response dependence in the unidimensional Rasch model. *Journal of Applied Measurement.* **9,** 105-124.

Marcenes, W., Muirhead, V.E., Murray, S., et al. 2013. Ethnic disparities in the oral health of three- to four-year-old children in East London. *British Dental Journal.* **215,** E4.

Marinho, V.C., Higgins, J.P., Sheiham, A. & Logan, S. 2003. Fluoride toothpastes for preventing dental caries in children and adolescents. *The Cochrane Database of Systematic Reviews.* CD002278.

Marinho, V.C., Worthington, H.V., Walsh, T. & Clarkson, J.E. 2013. Fluoride varnishes for preventing dental caries in children and adolescents. *The Cochrane Database of Systematic Reviews.* **7,** CD002279.

Marshman, Z., Rodd, H., Stern, M., et al. 2005. An evaluation of the Child Perceptions Questionnaire in the UK. *Community Dental Health.* **22,** 151-155.

Marshman, Z., Gibson, B.J., Owens, J., et al. 2007. Seen but not heard: a systematic review of the place of the child in 21st-century dental research. *International Journal of Paediatric Dentistry.* **17,** 320-327.

Marshman, Z., Baker, S.R., Bradbury, J., Hall, M.J. & Rodd, H.D. 2009. The psychosocial impact of oral conditions during transition to secondary education. *European Journal of Paediatric Dentistry.* **10,** 176-180.

Marshman, Z., Gibson, B.J. & Benson, P.E. 2010. Is the short-form Child Perceptions Questionnaire meaningful and relevant to children with malocclusion in the UK? *Journal of Orthodontics.* **37,** 29-36.

Marshman, Z., Innes, N., Deery, C., et al. 2012. The management of dental caries in primary teeth - involving service providers and users in the design of a trial. *Trials.* **13,** 143.

Martins, M.T., Ferreira, F.M., Oliveira, A.C., et al. 2009. Preliminary validation of the Brazilian version of the Child Perceptions Questionnaire 8-10. *Eur J Paediatr Dent.* **10,** 135-140.

Martins-Junior, P.A., Marques, L.S. & Ramos-Jorge, M.L. 2012a. Malocclusion: social, functional and emotional influence on children. *Journal of Clinical Pediatric Dentistry.* **37,** 103-108.

Martins-Junior, P.A., Oliveira, M., Marques, L.S. & Ramos-Jorge, M.L. 2012b. Untreated dental caries: impact on quality of life of children of low socioeconomic status. *Pediatric Dentistry.* **34,** 49-52.

Mashoto, K.O., Astrom, A.N., David, J. & Masalu, J.R. 2009. Dental pain, oral impacts and perceived need for dental treatment in Tanzanian school students: a cross-sectional study. *Health & Quality of Life Outcomes.* **7,** 73.

Mashoto, K.O., Astrom, A.N., Skeie, M.S. & Masalu, J.R. 2010. Changes in the quality of life of Tanzanian school children after treatment interventions using the Child-OIDP. *European Journal of Oral Sciences.* **118,** 626-634.

Massarente, D.B., Domaneschi, C., Marques, H.H.S., et al. 2011. Oral health-related quality of life of paediatric patients with AIDS. *BMC Oral Health.* **11,** 2.

Masters, G. 1982. Rasch model for partial credit scoring. *Psychometrika.* **47,** 149-174.

Matza, L.S., Swensen, A.R., Flood, E.M., Secnik, K. & Leidy, N.K. 2004. Assessment of health-related quality of life in children: a review of conceptual, methodological, and regulatory issues. *Value in Health.* **7,** 79-92.

Mbawalla, H.S., Masalu, J.R. & Astrom, A.N. 2010. Socio-demographic and behavioural correlates of oral hygiene status and oral health related quality of life, the Limpopo - Arusha school health project (LASH): A cross-sectional study. *BMC Pediatrics.* **10,** 87.

Mbawalla, H.S., Mtaya, M., Masalu, J.R., Brudvik, P. & Astrom, A.N. 2011. Discriminative ability of the generic and condition-specific Child-Oral Impacts on Daily Performances (Child-OIDP) by the Limpopo-Arusha School Health (LASH) Project: A cross-sectional study. *BMC Pediatrics.* **11,** 45.

McColl, E. 2005. *Developing questionnaires* New York, Oxford University Press.

McDonagh, M.S., Whiting, P.F., Wilson, P.M., et al. 2000. Systematic review of water fluoridation. *BMJ.* **321,** 855-859.

McGrath, C., Pang, H.N., Lo, E.C.M., et al. 2008. Translation and evaluation of a Chinese version of the Child Oral Health-related Quality of Life measure. *International Journal of Paediatric Dentistry.* **18,** 267-274.

McIver, J.P. & Carmines, E.G. 1981. *Unidimensional Scaling.* London, Sage Publications Ltd.

McLaughlin, H. 2006. Involving Young Service Users as Co-Researchers: Possibilities, benefits and Costs. *British Journal of Social Work.* **36,** 1395-1410.

McLaughlin, H. 2015. Involving children and young people in policy, practice and research. An introduction. London: National Children's Bureau. Available: http://www.ncb.org.uk/media/1176052/participation_essays_final.pdf [Accessed 2nd March 2015].

McLennan, D., Barnes, H., Noble, M., et al. 2011. The English Indices of Deprivation 2010. *In:* Government, D.o.C.a.L. (ed.). London: Office of National Statistics.

McMahon, A.D., Blair, Y., McCall, D.R. & Macpherson, L.M. 2011. Reductions in dental decay in 3-year old children in Greater Glasgow and Clyde: repeated population inspection studies over four years. *BMC Oral Health.* **11,** 29.

Medical Research Council. 2006. Water fluoridation and health. Working Party Report. London: MRC. Available: http://www.mrc.ac.uk/news-events/publications/water-fluoridation-and-health/ [Accessed 15th April 2015].

Meeske, K.A., Patel, S.K., Palmer, S.N., Nelson, M.B. & Parow, A.M. 2007. Factors associated with health-related quality of life in pediatric cancer survivors. *Pediatric Blood & Cancer.* **49,** 298-305.

Melzack, R. & Torgerson, W.S. 1971. On the language of pain. *Anesthesiology.* **34,** 50-59.

Mills, N.M. 1989. Pain behaviors in infants and toddlers. *Journal of Pain and Symptom Management.* **4,** 184-190.

Mokkink, L.B., Terwee, C.B., Patrick, D.L., et al. 2010a. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research.* **19,** 539-549.

Mokkink, L.B., Terwee, C.B., Gibbons, E., et al. 2010b. Inter-rater agreement and reliability of the COSMIN (COnsensus-based Standards for the selection of health status Measurement Instruments) checklist. *BMC Medical Research Methodology.* **10,** 82.

Monse, B., Duijster, D., Sheiham, A., et al. 2012. The effects of extraction of pulpally involved primary teeth on weight, height and BMI in underweight Filipino children. A cluster randomized clinical trial. *BMC Public Health.* **12,** 725.

Morgan, M., Monaghan, N. & Johnson, I. 2013. Picture of oral health 2012. Dental epidemiological survey of 5 year olds 2011-12. Cardiff: Cardiff University/Public Health Wales. Available: http://www.cardiff.ac.uk/__data/assets/pdf_file/0004/48712/five-year-olds-2011-12-summary.pdf [Accessed 25th February 2015].

Morris, C., Liabo, K., Wright, P. & Fitzpatrick, R. 2007. Development of the Oxford ankle foot questionnaire: finding out how children are affected by foot and ankle problems. *Child: Care, Health and Development.* **33,** 559-568.

Morris, C., Doll, H.A., Wainwright, A., Theologis, T. & Fitzpatrick, R. 2008. The Oxford ankle foot questionnaire for children: scaling, reliability and validity. *Journal of Bone and Joint Surgery (British Volume).* **90,** 1451-1456.

Morrow, V. & Richards, M. 1996. The ethics of social research with children: an overview. *Children and Society.* **10,** 90-105.

Morrow, V. 2008. Ethical dilemmas in research with children and young people about their social environment. *Children's Geographies.* **6,** 49-61.

Moynihan, P.J. & Kelly, S.A. 2014. Effect on caries of restricting sugars intake: systematic review to inform WHO guidelines. *Journal of Dental Research.* **93,** 8-18.

Mtaya, M., Astrom, A.N. & Tsakos, G. 2007. Applicability of an abbreviated version of the Child-OIDP inventory among primary schoolchildren in Tanzania. *Health & Quality of Life Outcomes.* **5,** 40.

Mtaya, M., Astrom, A.N. & Brudvik, P. 2008. Malocclusion, psycho-social impacts and treatment need: A cross-sectional study of Tanzanian primary school-children. *BMC Oral Health.* **8,** 14.

Naidoo, S., Chikte, U.M. & Sheiham, A. 2001. Prevalence and impact of dental pain in 8-10-year-olds in the western Cape. *SADJ.* **56,** 521-523.

Naidoo, S., Sheiham, A. & Tsakos, G. 2013. The relation between oral impacts on daily performances and perceived clinical oral conditions in primary school children in the Ugu District, Kwazulu Natal, South Africa. *SADJ.* **68,** 214-218.

Nalliah, R.P., Allareddy, V., Elangovan, S., Karimbux, N. & Allareddy, V. 2010. Hospital based emergency department visits attributed to dental caries in the United States in 2006. *Journal of Evidence-Based Dental Practice.* **10,** 212-222.

Nanayakkara, V., Renzaho, A., Oldenburg, B. & Ekanayake, L. 2013. Ethnic and socio-economic disparities in oral health outcomes and quality of life among Sri Lankan preschoolers: a cross-sectional study. *International Journal for Equity in Health.* **12,** 89.

Natapoff, J.N. 1978. Children's views of health: a developmental study. *American Journal of Public Health.* **68,** 995-1000.

National Institute for Health and Care Excellence. 2014. Oral health: approaches for local authorities and their partners to improve the oral health of their communities. NICE. Available: http://www.nice.org.uk/guidance/ph55/resources/guidance-oral-health-approaches-for-local-authorities-and-their-partners-to-improve-the-oral-health-of-their-communities-pdf [Accessed 25th March 2015].

NHS Confederation. 2012. Children and young people's health and wellbeing in changing times. Shaping the future and improving outcomes. London: NHS Confederation. Available: http://www.nhsconfed.org/~/media/Confederation/Files/Publications/Documents/Children-and-young-peoples-health-in-changing-times.pdf [Accessed 16th April 2015].

Nurelhuda, N.M., Ahmed, M.F., Trovik, T.A. & Astrom, A.N. 2010. Evaluation of oral health-related quality of life among Sudanese schoolchildren using Child-OIDP inventory. *Health and Quality of Life Outcomes.* **8,** 152.

Nuttall, N.M., Steele, J.G., Evans, D., et al. 2006. The reported impact of oral condition on children in the United Kingdom, 2003. *British Dental Journal.* **200,** 551-555.

O'Brien, C., Benson, P.E. & Marshman, Z. 2007. Evaluation of a quality of life measure for children with malocclusion. *Journal of Orthodontics.* **34,** 185-193; discussion 176.

O'Brien, K., Wright, J.L., Conboy, F., Macfarlane, T. & Mandall, N. 2006. The child perception questionnaire is valid for malocclusions in the United Kingdom. *Am J Orthod Dentofacial Orthop.* **129,** 536-540.

Otley, A., Smith, C., Nicholas, D., et al. 2002. The IMPACT questionnaire: a valid measure of health-related quality of life in pediatric inflammatory bowel disease. *Journal of Pediatric Gastroenterology and Nutrition.* **35,** 557-563.

Page, L.A., Thomson, W.M., Marshman, Z. & Stevens, K.J. 2014. The potential of the Child Health Utility 9D Index as an outcome measure for child dental health. *BMC Oral Health.* **14,** 90.

Page, L.A.F., Thomson, W.M., Mohamed, A.R. & Traebert, J. 2011. Performance and cross-cultural comparison of the short-form version of the CPQ11-14 in New Zealand, Brunei and Brazil. *Health and Quality of Life Outcomes.* **9,** 40.

Park, T., Reilly-Spong, M. & Gross, C.R. 2013. Mindfulness: a systematic review of instruments to measure an emergent patient-reported outcome (PRO). *Quality of Life Research.* **22,** 2639-2659.

Parliament. 1989. Children Act. London: HMSO. Available: http://www.legislation.gov.uk/ukpga/1989/41/contents [Accessed 20th April 2015].

Parliament. 1995. Children (Nothern Ireland) Order. London: HMSO. Available: http://www.legislation.gov.uk/nisi/1995/755/contents/made [Accessed 20th April 2015].

Patel, R.R., Tootla, R. & Inglehart, M.R. 2007. Does oral health affect self perceptions, parental ratings and video-based assessments of children's smiles? *Community Dentistry and Oral Epidemiology.* **35,** 44-52.

Patient Reported Outcomes Measurement Group. Unknown. *Selection Criteria* [Online]. University of Oxford. Available: http://phi.uhce.ox.ac.uk/inst_selcrit.php [Accessed 29th September 2015].

Pau, A., Baxevanos, K.G. & Croucher, R. 2007. Family structure is associated with oral pain in 12-year-old Greek schoolchildren. *International Journal of Paediatric Dentistry.* **17,** 345-351.

Pau, A., Khan, S.S., Babar, M.G. & Croucher, R. 2008. Dental pain and care-seeking in 11-14-yr-old adolescents in a low-income country. *European Journal of Oral Sciences.* **116,** 451-457.

Paula, J.S., Torres, L.H., Ambrosano, G.M. & Mialhe, F.L. 2012a. Association between oral health-related quality of life and atraumatic restorative treatment in school children: an exploratory study. *Indian Journal of Dental Research.* **23,** 738-741.

Paula, J.S., Leite, I.C., Almeida, A.B., et al. 2012b. The influence of oral health conditions, socioeconomic status and home environment factors on schoolchildren's self-perception of quality of life. *Health and Quality of Life Outcomes.* **10,** 6.

Paula, J.S., Meneghim, M.C., Pereira, A.C. & Mialhe, F.L. 2015. Oral health, socio-economic and home environmental factors associated with general and oral-health related quality of life and convergent validity of two instruments. *BMC Oral Health.* **15,** 26.

Perreault Jr., D. 1975. Controlling Order-Effect Bias. *The Public Opinion Quarterly.* **39,** 544-551.

Petersen, P.E. 2003. The World Oral Health Report 2003: continuous improvement of oral health in the 21st century--the approach of the WHO Global Oral Health Programme. *Community Dentistry and Oral Epidemiology.* **31 Suppl 1,** 3-23.

Piaget, J. & Inhelder, B. 1969. *The psychology of the child.* New York, Basic Books.

Piovesan, C., Ferreira Antunes, J.L., Guedes, R.S. & Ardenghi, T.M. 2010. Impact of socioeconomic and clinical factors on child oral health-related quality of life (COHRQoL). *Quality of Life Research.* **19,** 1359-1366.

Piovesan, C., Abella, C. & Ardenghi, T.M. 2011. Child oral health-related quality of life and socioeconomic factors associated with traumatic dental injuries in schoolchildren. *Oral Health and Preventive Dentistry.* **9,** 405-411.

Piovesan, C., Antunes, J.L., Mendes, F.M., Guedes, R.S. & Ardenghi, T.M. 2012. Influence of children's oral health-related quality of life on school performance and school absenteeism. *Journal of Public Health Dentistry.* **72,** 156-163.

Pitts, N.B., Chadwick, B. & Anderson, T. 2015. Children's Dental Health Survey 2013. Report 2: Dental Disease and Damage in Children. England, Wales and Northern Ireland. HSCIC. Available: http://www.hscic.gov.uk/catalogue/PUB17137/CDHS2013-Report2-Dental-Disease.pdf [Accessed 25th March 2015].

Porritt, J.M., Rodd, H.D. & Ruth Baker, S. 2011. Quality of life impacts following childhood dento-alveolar trauma. *Dental Traumatology.* **27,** 2-9.

Price, V.E., Klaassen, R.J., Bolton-Maggs, P.H., et al. 2009. Measuring disease-specific quality of life in rare populations: a practical approach to cross-cultural translation. *Health and Quality of Life Outcomes.* **7,** 92.

Prout, A. & James, A. 1998. A new paradigm for the sociology of childhood: Provenance, provision and problems. *In:* James, A. & Prout, A. (eds.) *Constructing and Reconstructing Childhood.* London: Falmer Press.

Prutkin, J.M. & Feinstein, A.R. 2002. Quality-of-life measurements: origin and pathogenesis. *Yale Journal of Biology and Medicine.* **75,** 79-93.

Public Health England. 2012. National Dental Epidemiology Programme for England: oral health survey of five-year-old children 2012. A report on the prevalence and severity of dental decay. London: Public Health England. Available: http://www.nwph.net/dentalhealth/survey-results5.aspx?id=1 [Accessed 5th February 2015].

Public Health England. 2013. Dental public health epidemiology programme. Oral health survey of three-year-old children 2013. A report on the prevalence and severity of dental decay. London: Public Health England. Available: http://www.nwph.net/dentalhealth/reports/DPHEP for England OH Survey 3yr 2013 Report.pdf [Accessed 19th February 2015].

Public Health England. 2014a. Child Health Profile. Sheffield March 2014. Public Health England. Available: http://www.chimat.org.uk/resource/view.aspx?RID=192229 [Accessed 19th March 2015].

Public Health England. 2014b. Delivering better oral health: an evidence-based toolkit for prevention. Third edition. London: Public Health England. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/367563/DBOHv32014OCTMainDocument_3.pdf [Accessed 25th February 2015].

Public Health England. 2014c. Dental Health Profile. Sheffield. PHE. Available: http://www.nwph.net/dentalhealth/5yearoldprofiles/Yorkshire and the Humber/Sheffield LA Dental Profile 5yr 2012.pdf [Accessed 5th October 2015].

Public Health England. 2014d. Local authorities improving oral health: commissioning better oral health for children and young people. An evidence-informed toolkit for local authorities. London: Public Health England. Available:

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/321503/CBOHMaindocumentJUNE2014.pdf [Accessed 25th March 2015].

Public Health England. 2014e. Sheffield Unitary Authority. Health Profile 2014. Public Health England. Available: http://www.apho.org.uk/resource/view.aspx?RID=142149 [Accessed 5th February 2015].

Public Health England. 2015. South Yorkshire and Bassetlaw oral health needs assessment 2015. London: PHE. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/463067/South_Yorks___Bassetlaw_oral_health_needs_assessment.pdf [Accessed 12th October 2015].

Punch, S. 2002. Research with children-the same or different from research with adults? *Childhood.* **9,** 321-341.

Pusic, A.L., Cemal, Y., Albornoz, C., et al. 2013. Quality of life among breast cancer patients with lymphedema: a systematic review of patient-reported outcome instruments and outcomes. *Journal of Cancer Survivorship.* **7,** 83-92.

Quittner, A.L., Schechter, M.S., Rasouliyan, L., et al. 2010. Impact of socioeconomic status, race, and ethnicity on quality of life in patients with cystic fibrosis in the United States. *Chest.* **137,** 642-650.

Rae, A. 2011. Deprivation in Sheffield. Sheffield: University of Sheffield. Available: https://www.shef.ac.uk/polopoly_fs/1.165648!/file/ajr_sheffield_deprivation_nov_2011.pdf [Accessed 5th October 2015].

Ramos-Jorge, M.L., Vieira-Andrade, R.G., Martins-Junior, P.A., et al. 2012. Level of agreement between self-administered and interviewer-administered CPQ(8-10) and CPQ(11-14). *Community Dentistry and Oral Epidemiology.* **40,** 201-209.

Randall, R.C. 2002. Preformed metal crowns for primary and permanent molar teeth: review of the literature. *Pediatric Dentistry.* **24,** 489-500.

Rasch, G. 1960. *Proababilistic models for some intellegence and attainment tests.* Chicago, University of Chicago Press.

Ratnayake, N. & Ekanayake, L. 2005. Prevalence and impact of oral pain in 8-year-old children in Sri Lanka. *International Journal of Paediatric Dentistry.* **15,** 105-112.

Ravaghi, V., Ardakan, M.M.M., Shahriari, S., Mokhtari, N. & Underwood, M. 2011. Comparison of the COHIP and OHIP- 14 as measures of the oral health-related quality of life of adolescents. *Community Dental Health.* **28,** 82-88.

Raymundo de Andrade, L.H., de Souza Rocha, B., Castro, G.F. & Ribeiro de Souza, I.P. 2011. Impact of oral problems on daily activities of HIV-infected children. *European Journal of Paediatric Dentistry.* **12,** 75-80.

research randomizer. Available: randomizer.org [Accessed 6th January 2015].

Riley, A.W., Forrest, C.B., Rebok, G.W., et al. 2004. The Child Report Form of the CHIP-Child Edition: reliability and validity. *Medical Care.* **42,** 221-231.

Ritchie, J. & Spencer, L. 1994. Qualitative data analysis for applied policy research. *In:* Bryman, A. & Burgess, R. (eds.) *Analyzing Qualitative Data.* New York: Routledge.

Ritchie, J., Spencer, L. & O''Connor, W. 2003. *Carrying out qualitative analysis* Thousand Oaks: CA, SAGE.

Robinson, P.G., Carr, A.J. & Higginson, I.J. 2003. How to choose a quality of life measure. *In:* Robinson, P.G., Carr, A.J. & Higginson, I.J. (eds.) *Quality of Life.* London: BMJ Books.

Robinson, P.G., Nalweyiso, N., Busingye, J. & Whitworth, J. 2005. Subjective impacts of dental caries and fluorosis in rural Ugandan children. *Community Dental Health.* **22,** 231-236.

Rodd, H.D. & Blankenstein, R. 1995. Iron-deficiency anaemia in young children requiring dental extractions under general anaesthesia. *International Journal of Paediatric Dentistry.* **5,** 9-14.

Rodd, H.D., Hall, M., Deery, C., et al. 2013. Video diaries to capture children's participation in the dental GA pathway. *European Archives of Paediatric Dentistry.* **14,** 325-330.

Rooney, E., Davies, G., Neville, J., et al. 2010. NHS Dental Epidemiology Programme for England. Oral Health Survey of 12 year old Children 2008 / 2009. Summary of caries prevalence and severity results. The Dental Observatory. The North West Public Health Observatory. Available: http://www.nwph.net/dentalhealth/survey-results-12.aspx [Accessed 5th February 2015].

Rosel, E., Tsakos, G., Bernabe, E., Sheiham, A. & Bravo, M. 2010. Assessing the level of agreement between the self- and interview-administered Child-OIDP. *Community Dentistry and Oral Epidemiology.* **38,** 340-347.

Ross, D.M. & Ross, S.A. 1984. Childhood pain: the school-aged child's viewpoint. *Pain.* **20,** 179-191.

Royal College of Paediatrics and Child Health: Ethics Advisory Committee 2000. Guidelines for the ethical conduct of medical research involving children. *Archives of Disease in Childhood.* **82,** 177-182.

Rubin, D. 1976. Inference and missing data. *Biometrika.* **63,** 581–592.

Rudestam, K., Brown, P., Zarcadoolas, C. & Mansell, C. 2004. Children's asthma experience and the importance of place. *Health (London).* **8,** 423-444.

Sandelowski, M. 1995. Sample size in qualitative research. *Research in Nursing and Health.* **18,** 179-183.

Sartain, S.A., Clarke, C.L. & Heyman, R. 2000. Hearing the voices of children with chronic illness. *Journal of Advanced Nursing.* **32,** 913-921.

Savedra, M., Gibbons, P., Tesler, M., Ward, J. & Wegner, C. 1982. How do children describe pain? A tentative assessment. *Pain.* **14,** 95-104.

Savedra, M.C., Tesler, M.D., Holzemer, W.L. & Brokaw, P. 1995. A strategy to assess the temporal dimension of pain in children and adolescents. *Nursing Research.* **44,** 272-276.

Schellingerhout, J.M., Heymans, M.W., Verhagen, A.P., et al. 2011. Measurement properties of translated versions of neck-specific questionnaires: a systematic review. *BMC Medical Research Methodology.* **11,** 87.

Schellingerhout, J.M., Verhagen, A.P., Heymans, M.W., et al. 2012. Measurement properties of disease-specific questionnaires in patients with neck pain: a systematic review. *Quality of Life Research.* **21,** 659-670.

Schneider, E.M. & LoBiondo-Wood, G. 1992. Perceptions of procedural pain: parents, nurses, and children. *Children's Health Care.* **21,** 157-162.

Schroth, R.J., Levi, J., Kliewer, E., Friel, J. & Moffatt, M.E. 2013. Association between iron status, iron deficiency anaemia, and severe early childhood caries: a case-control study. *BMC Pediatrics.* **13,** 22.

Schwendicke, F., Dorfer, C.E., Schlattmann, P., et al. 2015. Socioeconomic inequality and caries: a systematic review and meta-analysis. *Journal of Dental Research.* **94,** 10-18.

Scott, J. & Huskisson, E.C. 1979. Accuracy of subjective measurements made with or without previous scores: an important source of error in serial measurement of subjective states. *Annals of the Rheumatic Diseases.* **38,** 558-559.

Scottish Intercollegiate Guidelines Network. 2014. Dental interventions to prevent caries in children.(SIGN publication no. 138). Edinburgh: SIGN. Available: http://www.sign.ac.uk [Accessed 16th February 2015].

Scottish Parliament. 2014. Children and Young People (Scotland) Act. London: TSO. Available: http://www.legislation.gov.uk/ukpga/1995/36 [Accessed 9th January 2013].

Scwartz, N. & Sudman, S. 1996. *Answering questions: Methodology for determining cognitive and communicative processes in survey methodology* San Francisco, Jossey-Bass.

Sheffield City Council. 2010. Indices of Deprivation 2010 — Sheffield. Available: https://www.sheffield.gov.uk/your-city-council/sheffield-profile/deprivation-statistics.html [Accessed 5th February 2015].

Sheffield City Council. 2012. Community Knowledge Profiles. Sheffield City Council. Available: https://www.sheffield.gov.uk/your-city-council/sheffield-profile/community-knowledge-profiles.html [Accessed 8th April 2015].

Sheffield Fairness Commission. 2013. Making Sheffield Fair. Sheffield: Sheffield Fairness Committee. Available: http://www.sth.nhs.uk/clientfiles/File/Enclosure J3 - Fairness Commission Report.pdf [Accessed 5th October 2015].

Sheiham, A., Maizels, J.E. & Cushing, A.M. 1982. The concept of need in dental care. *International Dental Journal.* **32,** 265-270.

Sheiham, A. & Spencer, A.J. 1997. *Health needs assessment.* Bath, Wright.

Sheiham, A. & Watt, R.G. 2000. The common risk factor approach: a rational basis for promoting oral health. *Community Dentistry and Oral Epidemiology.* **28,** 399-406.

Sheiham, A. & Watt, R. 2003. *Oral health promotion and policy.* Oxford, Oxford University Press.

Sheiham, A. 2006. Dental caries affects body weight, growth and quality of life in pre-school children. *British Dental Journal.* **201,** 625-626.

Shelton, A.T., Houghton, N.Y., Morris, D.O., et al. 2015. The development and validation of a psychological questionnaire for patients undergoing orthognathic treatment. *Orthodontics & Craniofacial Research.* **18,** 51-64.

Shepherd, M.A., Nadanovsky, P. & Sheiham, A. 1999. The prevalence and impact of dental pain in 8-year-old school children in Harrow, England. *British Dental Journal.* **187,** 38-41.

Shrive, F.M., Stuart, H., Quan, H. & Ghali, W.A. 2006. Dealing with missing data in a multi-question depression scale: a comparison of imputation methods. *BMC Medical Research Methodology.* **6,** 57.

Sinclair, R. & Franklin, A. 2000. Young People's Participation. Quality Protects Research Briefings, no 3. London: Department of Health. Available: https://lemosandcrane.co.uk/resources/DoH%20-%20Young%20Peoples%20participation.pdf [Accessed 5th January 2015]

Smith Jr, E.V. 2005. Effect of item redundancy on Rasch item and person estimates. *Journal of Applied Measurement.* **6,** 147-163.

Smith, L. & Callery, P. 2005. Children's accounts of their preoperative information needs. *Journal of Clinical Nursing.* **14,** 230-238.

Smith, R.M. 2000. Fit analysis in latent trait measurement models. *Journal of Applied Measurement.* **1,** 199-218.

Solans, M., Pane, S., Estrada, M.D., et al. 2008. Health-related quality of life measurement in children and adolescents: a systematic review of generic and disease-specific instruments. *Value in Health.* **11,** 742-764.

Spencer, A.J., Slade, G.D. & Davies, M. 1996. Water fluoridation in Australia. *Community Dental Health.* **13 Suppl 2,** 27-37.

Stanford, E.A., Chambers, C.T. & Craig, K.D. 2005. A normative analysis of the development of pain-related vocabulary in children. *Pain.* **114,** 278-284.

Steele, J., White, D., Rolland, S. & Fuller, E. 2015. Children's Dental Health Survey 2013. Report 4: The Burden of Dental Disease in Children England, Wales and Northern Ireland. London: HSCIC. Available: http://www.hscic.gov.uk/catalogue/PUB17137/CDHS2013-Report4-Burden-of-Dental-Disease.pdf [Accessed 19th March 2015].

Stephenson, J., Chadwick, B.L., Playle, R.A. & Treasure, E.T. 2010. A competing risk survival analysis model to assess the efficacy of filling carious primary teeth. *Caries Research.* **44,** 285-293.

Stevens, K. 2009. Developing a descriptive system for a new preference-based measure of health-related quality of life for children. *Quality of Life Research.* **18,** 1105-1113.

Stevens, K. & Ratcliffe, J. 2012. Measuring and valuing health benefits for economic evaluation in adolescence: an assessment of the practicality and validity of the child health utility 9D in the Australian adolescent population. *Value in Health.* **15,** 1092-1099.

Stevens, K.J. 2010. Working with children to develop dimensions for a preference-based, generic, pediatric, health-related quality-of-life measure. *Qualitative Health Research.* **20,** 340-351.

Stinson, J.N., Sung, L., Gupta, A., et al. 2012. Disease self-management needs of adolescents with cancer: perspectives of adolescents with cancer and their parents and healthcare providers. *Journal of Cancer Survivorship.* **6,** 278-286.

Sundblad, G.M., Saartok, T. & Engstrom, L.M. 2006. Child-parent agreement on reports of disease, injury and pain. *BMC Public Health.* **6,** 276.

Svensson, E. 2001. Guidelines to statistical evaluation of data from rating scales and questionnaires. *Journal of Rehabilitation Medicine.* **33,** 47-48.

Taylor, K.R., Kiyak, A., Huang, G.J., et al. 2009. Effects of malocclusion and its treatment on the quality of life of adolescents. *Am J Orthod Dentofacial Orthop.* **136,** 382-392.

Tennant, A., McKenna, S.P. & Hagell, P. 2004. Application of Rasch analysis in the development and application of quality of life instruments. *Value in Health.* **7 Suppl 1,** S22-26.

Tennant, A. & Pallant, J. 2006. Unidimensionality counts (A tale of two Smiths?). *Rasch Measurement Transactions.* **20,** 1048-1051.

Tennant, A. & Conaghan, P.G. 2007. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis and Rheumatism.* **57,** 1358-1362.

Terwee, C.B., Bot, S.D., de Boer, M.R., et al. 2007. Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology.* **60,** 34-42.

Terwee, C.B., Mokkink, L.B., Knol, D.L., et al. 2012. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Quality of Life Research.* **21,** 651-657.

The Equality Trust. 2015. *The Scale of Economic Inequality in the UK* [Online]. The Equality Trust. Available: https://www.equalitytrust.org.uk/scale-economic-inequality-uk [Accessed 5th October 2015].

The UK Data Service. *GeoConvert* [Online]. University of Essex and University of Manchester. Available: http://geoconvert.mimas.ac.uk [Accessed 3rd February 2015].

Thorborg, K., Holmich, P., Christensen, R., Petersen, J. & Roos, E.M. 2011. The Copenhagen Hip and Groin Outcome Score (HAGOS): development and validation according to the COSMIN checklist. *British Journal of Sports Medicine.* **45,** 478-491.

Threlfall, A.G., Pilkington, L., Milsom, K.M., Blinkhorn, A.S. & Tickle, M. 2005. General dental practitioners' views on the use of stainless steel crowns to restore primary molars. *British Dental Journal.* **199,** 453-455; discussion 441.

Tickle, M., Milsom, K. & Kennedy, A. 1999. Is it better to leave or restore carious deciduous molar teeth? A preliminary study. *Primary Dental Care.* **6,** 127-131.

Tickle, M., Milsom, K., King, D., Kearney-Mitchell, P. & Blinkhorn, A. 2002. The fate of the carious primary teeth of children who regularly attend the general dental service. *British Dental Journal.* **192,** 219-223.

Toole, R.J., Lindsay, S.J., Johnstone, S. & Smith, P. 2000. An investigation of language used by children to describe discomfort during dental pulp-testing. *International Journal of Paediatric Dentistry.* **10,** 221-228.

Torres, C.S., Paiva, S.M., Vale, M.P., et al. 2009. Psychometric properties of the Brazilian version of the Child Perceptions Questionnaire (CPQ11-14) - short forms. *Health & Quality of Life Outcomes.* **7,** 43.

Tourangeau, R. & Rasinski, K. 1988. Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin.* **103,** 299-314.

Traebert, J., de Lacerda, J.T., Thomson, W.M., Page, L.F. & Locker, D. 2010a. Differential item functioning in a Brazilian-Portuguese version of the Child Perceptions Questionnaire (CPQ(11-14)). *Community Dentistry and Oral Epidemiology.* **38,** 129-135.

Traebert, J., Foster Page, L.A., Thomson, W.M. & Locker, D. 2010b. Differential item functioning related to ethnicity in an oral health-related quality of life measure. *International Journal of Paediatric Dentistry.* **20,** 435-441.

Traebert, J., Lacerda, J.T., Foster Page, L.A., Thomson, W.M. & Bortoluzzi, M.C. 2012. Impact of traumatic dental injuries on the quality of life of schoolchildren. *Dental Traumatology.* **28,** 423-428.

Tsakos, G., Gherunpong, S. & Sheiham, A. 2006. Can oral health-related quality of life measures substitute for normative needs assessments in 11 to 12-year-old children? *Journal of Public Health Dentistry.* **66,** 263-268.

Tsakos, G., Bernabe, E., O'Brien, K., Sheiham, A. & de Oliveira, C. 2008. Comparison of the self-administered and interviewer-administered modes of the child-OIDP. *Health & Quality of Life Outcomes.* **6,** 40.

Tsakos, G., Blair, Y.I., Yusuf, H., et al. 2012. Developing a new self-reported scale of oral health outcomes for 5-year-old children (SOHO-5). *Health and Quality of Life Outcomes.* **10,** 62.

Tsakos, G., Allen, P.F. & Steele, J.G. 2013. What has oral health related quality of life ever done for us? *Community Dental Health.* **30,** 66-67.

Tsakos, G., Hill, K., Chadwick, B. & Anderson, T. 2015. Children's Dental Health Survey 2013. Report 1: Attitudes, Behaviours and Children's Dental Health. England, Wales and Northern Ireland, 2013. HSCIC. Available: http://www.hscic.gov.uk/catalogue/PUB17137/CDHS2013-Report1-Attitudes-and-Behaviours.pdf [Accessed 26th march 2015].

Tubert-Jeannin, S., Pegon-Machat, E., Gremeau-Richard, C., Lecuyer, M.-M. & Tsakos, G. 2005. Validation of a French version of the Child-OIDP index. *European Journal of Oral Sciences.* **113,** 355-362.

Turner, R.R., Quittner, A.L., Parasuraman, B.M., et al. 2007. Patient-reported outcomes: instrument development and selection issues. *Value in Health.* **10 Suppl 2,** S86-93.

Turton, B.J., Thomson, W.M., Foster Page, L.A., Saub, R. & Ishak, A.R. 2015. Responsiveness of the Child Perceptions Questionnaire11-14 for Cambodian children undergoing basic dental care. *International Journal of Paediatric Dentistry.* **25,** 2-8.

U.S. Department of Health and Human Services. 2000. Oral health in America: a report of the Surgeon General. Rockville, MD: National Institute of Dental and Craniofacial Research. Available: silk.nih.gov/public/hck1ocv.@http://www.surgeon.fullrpt.pdf [Accessed 20th April 2015].

U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER) & Center for Devices and Radiological Health (CDRH). 2009. Guidance for Industry.

Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Food and Drug Administration. Available: http://www.fda.gov/downloads/Drugs/Guidances/UCM193282.pdf [Accessed 20th April 2015].

United Nations. 1989. United Nations Convention on the Rights of the Child. Geneva: Available: http://www.unicef.org.uk/Documents/Publication-pdfs/UNCRC_PRESS200910web.pdf [Accessed 20th April 2015].

United Nations Development Programme. 2013. *Human Development Reports* [Online]. United Nations Development Programme. Available: http://hdr.undp.org/en/content/table-1-human-development-index-and-its-components [Accessed 5th October 2015].

Vacik, H.W., Nagy, M.C. & Jessee, P.O. 2001. Children's understanding of illness: students' assessments. *Journal of Pediatric Nursing.* **16,** 429-437.

van Blerk, L. & Ansell, N. 2007. Participatory feedback and dissemination *with* and *for* children*:* Reflections from research with young migrants in southern Africa. *Children's Geographies.* **5,** 313-324.

Vargas-Ferreira, F., Piovesan, C., Praetzel, J.R., et al. 2010. Tooth erosion with low severity does not impact child oral health-related quality of life. *Caries Research.* **44,** 531-539.

Varni, J.W., Seid, M. & Kurtin, P.S. 2001. PedsQL 4.0: reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. *Medical Care.* **39,** 800-812.

Varni, J.W., Limbers, C.A. & Burwinkle, T.M. 2007. How young can children reliably and validly self-report their health-related quality of life?: an analysis of 8,591 children across age subgroups with the PedsQL 4.0 Generic Core Scales. *Health and Quality of Life Outcomes.* **5,** 1.

Vermaire, J.H., van Loveren, C., Brouwer, W.B. & Krol, M. 2014. Value for money: economic evaluation of two different caries prevention programmes compared with standard care in a randomized controlled trial. *Caries Research.* **48,** 244-253.

Versteegh, M.M., Leunis, A., Uyl-de Groot, C.A. & Stolk, E.A. 2012. Condition-specific preference-based measures: benefit or burden? *Value in Health.* **15,** 504-513.

Vinney, L.A., Grade, J.D. & Connor, N.P. 2012. Feasibility of using a handheld electronic device for the collection of patient reported outcomes data from children. *Journal of Communication Disorders.* **45,** 12-19.

Wainer, H. 2000. *Computerized Adaptive Testing: A Primer* Mahwah, NJ, Lawrence Erlbaum Associates.

Wallander, J.L., Fradkin, C., Chien, A.T., et al. 2012. Racial/ethnic disparities in health-related quality of life and health in children are largely mediated by family contextual differences. *Academic Pediatrics.* **12,** 532-538.

Walsh, M.E. & Bibace, R. 1991. Children's conceptions of AIDS: a developmental analysis. *Journal of Pediatric Psychology.* **16,** 273-285.

Walsh, T., Worthington, H.V., Glenny, A.M., et al. 2010. Fluoride toothpastes of different concentrations for preventing dental caries in children and adolescents. *The Cochrane Database of Systematic Reviews.* CD007868.

Watt, R. & Sheiham, A. 1999. Inequalities in oral health: a review of the evidence and recommendations for action. *British Dental Journal.* 6-12.

Watt, R.G. 2007. From victim blaming to upstream action: tackling the social determinants of oral health inequalities. *Community Dentistry and Oral Epidemiology.* **35,** 1-11.

Wayman, J. 2003. Multiple Imputation For Missing Data: What Is It And How Can I Use It? Maryland: John Hopkins University. Available: http://www.csos.jhu.edu/contact/staff/jwayman_pub/wayman_multimp_aera2003.pdf [Accessed 3rd February 2015].

Weldam, S.W., Schuurmans, M.J., Liu, R. & Lammers, J.W. 2013. Evaluation of Quality of Life instruments for use in COPD care and research: a systematic review. *International Journal of Nursing Studies.* **50,** 688-707.

Welsh Assembly Government 2008. Designed to Smile – A National Child Oral Health Improvement Programme Promoting Better Oral Health and Delivering a Fluoride Supplementation Programme. Cardiff: Welsh Assembly Government.

Welsh Oral Health Information Unit & Public Health Wales. 2010. An atlas of oral health in Wales. Volume I. Survey of rising 12 year olds 2008/09. Welsh Oral Health Information Unit/ Public Health Wales Available: http://www.cardiff.ac.uk/__data/assets/pdf_file/0011/48719/twelve-year-olds-2008-09-full-report.pdf [Accessed 25th February 2015].

Weng, L.-J. 2004. Impact of the number of response categories and anchor labels on coefficient alpha and test-retest reliability. *Educational and Psychological Measurement.* **64,** 956-972.

Westcott, H. & Davies, G. 1996. Sexually abused children's and young people's perspectives on investigative interviews. *British Journal of Social Work.* **26,** 451-474.

Wiebe, S., Guyatt, G., Weaver, B., Matijevic, S. & Sidwell, C. 2003. Comparative responsiveness of generic and specific quality-of-life instruments. *Journal of Clinical Epidemiology.* **56,** 52-60.

Wilkie, D.J., Holzemer, W.L., Tesler, M.D., et al. 1990. Measuring pain quality: validity and reliability of children's and adolescents' pain language. *Pain.* **41,** 151-159.

Wille, N., Badia, X., Bonsel, G., et al. 2010. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Quality of Life Research.* **19,** 875-886.

Williams, B., Brown, T. & Onsman, A. 2010. Exploratory factor analysis: A five-step guide for novices. *Australasian Journal of Paramedicine.* **8,** 1.

Wilson, I.B. & Cleary, P.D. 1995. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *JAMA.* **273,** 59-65.

Wilson-Genderson, M., Broder, H.L. & Phillips, C. 2007. Concordance between caregiver and child reports of children's oral health-related quality of life. *Community Dentistry & Oral Epidemiology.* **35 Suppl 1,** 32-40.

Wogelius, P., Gjorup, H., Haubek, D., Lopez, R. & Poulsen, S. 2009. Development of Danish version of child oral-health-related quality of life questionnaires (CPQ8-10 and CPQ11-14). *BMC Oral Health.* **9,** 11.

Wogelius, P., Rosthoj, S., Dahllof, G. & Poulsen, S. 2011. Oral health-related quality of life among survivors of childhood cancer. *International Journal of Paediatric Dentistry.* **21,** 465-467.

Wong, A.T., McMillan, A.S. & McGrath, C. 2006. Oral health-related quality of life and severe hypodontia. *Journal of Oral Rehabilitation.* **33,** 869-873.

Wong, M.C.M., Lau, A.W.H., Lam, K.F., McGrath, C. & Lu, H.-X. 2011. Assessing consistency in oral health-related quality of life (OHRQoL) across gender and stability of OHRQoL over time for adolescents using Structural Equation Modeling. *Community Dentistry and Oral Epidemiology.* **39,** 325-335.

Woodgate, R.L. 2005. A different way of being: adolescents' experiences with cancer. *Cancer Nursing.* **28,** 8-15.

World Health Organization. 1948. Preamble to the Constitution of the World Health Organization Available: [Accessed

World Health Organization. 1980. International Classification of Impairments, Disabilities, and Handicaps. Geneva: WHO. Available: http://whqlibdoc.who.int/publications/1980/9241541261_eng.pdf [Accessed 20th April 2015].

World Health Organization. 2002. International Classification of Functioning, Disability and Health. ICF. Geneva: WHO. Available: http://www.who.int/classifications/icf/en/ [Accessed 20th April 2015].

World Health Organization. 2007a. Assessing the iron status of populations. Geneva, Switzerland: World Health Organization. Available: http://www.who.int/nutrition/publications/micronutrients/anaemia_iron_deficiency/9789241596107/en/ [Accessed 20th April 2015].

World Health Organization 2007b. Basic documents. World Health Organization.

World Health Organization. 2008. Worldwide prevalence of anaemia 1993–2005. WHO Global Database on Anaemia. Geneva: World Health Organization. Available: whqlibdoc.who.int/publications/2008/9789241596657_eng.pdf [Accessed 20th April 2015].

World Health Organization. 2014. WHO Framework Convention on Tobacco Control. Geneva: WHO. Available: http://www.who.int/fctc/publications/en/ [Accessed 8th October 2015].

World Health Organization. 2015a. Guideline: Sugars intake for adults and children. Geneva: World Health Organization. Available: who.int/nutrition/publications/guidelines/sugars_intake/en [Accessed 20th April 2015].

World Health Organization. 2015b. *Oral Health Database. Global Health of 12-year-olds: 2015* [Online]. University of Malmo. Available: http://www.mah.se/CAPP/Country-Oral-Health-Profiles/According-to-Alphabetical/Global-DMFT-for-12-year-olds-2011/ [Accessed 25th February 2015].

Yang, Y., Brazier, J.E., Tsuchiya, A. & Young, T.A. 2011. Estimating a preference-based index for a 5-dimensional health state classification for asthma derived from the asthma quality of life questionnaire. *Medical Decision Making.* **31,** 281-291.

Yau, D.T., Wong, M.C., Lam, K.F. & McGrath, C. 2015. Evaluation of psychometric properties and differential item functioning of 8-item Child Perceptions Questionnaires using item response theory. *BMC Public Health.* **15,** 792.

Young, N.L., Bradley, C.S., Blanchette, V., et al. 2004. Development of a health-related quality of life measure for boys with haemophilia: the Canadian Haemophilia Outcomes--Kids Life Assessment Tool (CHO-KLAT). *Haemophilia.* **10 Suppl 1,** 34-43.

Young, N.L., Wakefield, C., Burke, T.A., et al. 2013. Updating the Canadian hemophilia outcomes-kids life assessment tool (CHO-KLAT Version2.0). *Value in Health.* **16,** 837-841.

Yusof, Z.Y. & Jaafar, N. 2012. A Malay version of the Child Oral Impacts on Daily Performances (Child-OIDP) index: assessing validity and reliability. *Health and Quality of Life Outcomes.* **10,** 63.

Yusuf, H., Gherunpong, S., Sheiham, A. & Tsakos, G. 2006. Validation of an English version of the Child-OIDP index, an oral health-related quality of life measure for children. *Health & Quality of Life Outcomes.* **4,** 38.

Zernikow, B., Meyerhoff, U., Michel, E., et al. 2005. Pain in pediatric oncology--children's and parents' perspectives. *European Journal of Pain (London, England).* **9,** 395-406.

Zhang, M., McGrath, C. & Hagg, U. 2007a. Patients' expectations and experiences of fixed orthodontic appliance therapy - Impact on quality of life. *Angle Orthodontist.* **77,** 318-322.

Zhang, M., McGrath, C. & Hagg, U. 2007b. Who knows more about the impact of malocclusion on children's quality of life, mothers or fathers? *European Journal of Orthodontics.* **29,** 180-185.

Zhang, M., McGrath, C. & Haegg, U. 2008. Changes in oral health-related quality of life during fixed orthodontic appliance therapy. *Am J Orthod Dentofacial Orthop.* **133,** 25-29.

Zhang, M., McGrath, C. & Hagg, U. 2009. Orthodontic treatment need and oral health-related quality among children. *Community Dent Health.* **26,** 58-61.

Zhou, H., Roberts, P. & Horgan, L. 2008. Association between self-report pain ratings of child and parent, child and nurse and parent and nurse dyads: meta-analysis. *Journal of Advanced Nursing.* **63,** 334-342.

Appendices

The COSMIN checklist with 4-point scale[6]

---

[6] Terwee et al., 2012

COSMIN checklist with 4-point scale

Contact
CB Terwee, PhD
VU University Medical Center
Department of Epidemiology and Biostatistics
EMGO Institute for Health and Care Research
1081 BT Amsterdam
The Netherlands
Website: www.cosmin.nl, www.emgo.nl
E-mail: cb.terwee@vumc.nl

Instructions

This version of the COSMIN checklist is recommended for use in systematic reviews of measurement properties. With this version it is possible to calculate overall methodological quality scores per study on a measurement property. A methodological quality score per box is obtained by taking the lowest rating of any item in a box ('worse score counts'). For example, if for a reliability study one item in the box 'Reliability' is scored poor, the methodological quality of that reliability study is rated as poor. The Interpretability box and the Generalizability box are mainly used as data extraction forms. We recommend to use the Interpretability box to extract all information on the interpretability issues described in this box (e.g. norm scores, floor-ceiling effects, minimal important change) of the instruments under study from the included articles. Similar, we recommend to use the Generalizability box to extract data on the characteristics of the study population and sampling procedure. Therefore no scoring system was developed for these boxes.

This scoring system is described in this paper:

Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, de Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Quality of Life Research 2011, July 6 [epub ahead of print].

Step 1. Evaluated measurement properties in the article

	Internal consistency	Box A
	Reliability	Box B
	Measurement error	Box C
	Content validity	Box D
	Structural validity	Box E
	Hypotheses testing	Box F
	Cross-cultural validity	Box G
	Criterion validity	Box H
	Responsiveness	Box I

Step 2. Determining if the statistical method used in the article are based on CTT or IRT

Box General requirements for studies that applied Item Response Theory (IRT) models					
		excellent	**good**	**fair**	**poor**
1	Was the IRT model used adequately described? e.g. One Parameter Logistic Model (OPLM), Partial Credit Model (PCM), Graded Response Model (GRM)	IRT model adequately described	IRT model not adequately described		
2	Was the computer software package used adequately described? e.g. RUMM2020, WINSTEPS, OPLM, MULTILOG, PARSCALE, BILOG, NLMIXED	Software package adequately described	Software package not adequately described		
3	Was the method of estimation used adequately described? e.g. conditional maximum likelihood (CML), marginal maximum likelihood (MML)	Method of estimation adequately described	Method of estimation not adequately described		
4	Were the assumptions for estimating parameters of the IRT model checked? e.g. unidimensionality, local independence, and item fit (e.g. differential item functioning (DIF))	assumptions of the IRT model checked	assumptions of the IRT model partly checked	assumptions of the IRT model not checked or unknown	

To obtain a total score for the methodological quality of studies that use IRT methods, the ‘worse score counts’ algorithm should be applied to the IRT box in combination with the box of the measurement property that was evaluated in the IRT study. For example, if IRT methods are used to study internal consistency and item 4 in the IRT box is scored fair, while the items in the internal consistency box (box A) are all scored as good or excellent, the methodological quality score for internal consistency will be fair. However, if any of the items in box A is scored poor, the methodological quality score for internal consistency will be poor.

Step 3. Determining if a study meets the standards for good methodological quality

Box A. Internal consistency	**excellent**	**good**	**fair**	**poor**
1 Does the scale consist of effect indicators, i.e. is it based on a reflective model?				
Design requirements				
2 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
3 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
4 Was the sample size included in the internal consistency analysis adequate?	Adequate sample size (≥100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)
5 Was the unidimensionality of the scale checked? i.e. was factor analysis or IRT model applied?	Factor analysis performed in the study population	Authors refer to another study in which factor analysis was performed in a similar study population	Authors refer to another study in which factor analysis was performed, but not in a similar study population	Factor analysis NOT performed and no reference to another study
6 Was the sample size included in the unidimensionality analysis adequate?	7* #items and ≥100	5* #items and ≥100 OR 6-7* #items but <100	5* #items but <100	<5* #items

7	Was an internal consistency statistic calculated for each (unidimensional) (sub)scale separately?	Internal consistency statistic calculated for each subscale separately			Internal consistency statistic NOT calculated for each subscale separately
8	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
Statistical methods					
9	for Classical Test Theory (CTT), continuous scores: Was Cronbach's alpha calculated?	Cronbach's alpha calculated		Only item-total correlations calculated	No Cronbach's alpha and no item-total correlations calculated
10	for CTT, dichotomous scores: Was Cronbach's alpha or KR-20 calculated?	Cronbach's alpha or KR-20 calculated		Only item-total correlations calculated	No Cronbach's alpha or KR-20 and no item-total correlations calculated
11	for IRT: Was a goodness of fit statistic at a global level calculated? E.g. χ^2, reliability coefficient of estimated latent trait value (index of (subject or item) separation)	Goodness of fit statistic at a global level calculated			Goodness of fit statistic at a global level NOT calculated

NB. Item 1 is used to determine whether internal consistency is relevant for the instrument under study. It is not used to rate the quality of the study.

Box B. Reliability: relative measures (including test-retest reliability, inter-rater reliability and intra-rater reliability)

		excellent	**good**	**fair**	**poor**
Design requirements					
1	Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2	Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3	Was the sample size included in the analysis adequate?	Adequate sample size (≥100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)
4	Were at least two measurements available?	At least two measurements			Only one measurement
5	Were the administrations independent?	Independent measurements	Assumable that the measurements were independent	Doubtful whether the measurements were independent	measurements NOT independent
6	Was the time interval stated?	Time interval stated		Time interval NOT stated	
7	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable
8	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate	Time interval NOT appropriate

9	Were the test conditions similar for both measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar
10	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
Statistical methods					
11	for continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated
12	for dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			Only percentage agreement calculated
13	for ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated	Only percentage agreement calculated
14	for ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described		

Box C. Measurement error: absolute measures

		excellent	good	fair	poor
Design requirements					
1	Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2	Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3	Was the sample size included in the analysis adequate?	Adequate sample size (≥100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)
4	Were at least two measurements available?	At least two measurements			Only one measurement
5	Were the administrations independent?	Independent measurements	Assumable that the measurements were independent	Doubtful whether the measurements were independent	measurements NOT independent
6	Was the time interval stated?	Time interval stated		Time interval NOT stated	
7	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable
8	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate	Time interval NOT appropriate

9	Were the test conditions similar for both measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar
10	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
Statistical methods					
11	for CTT: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated based on Cronbach's alpha, or on SD from another population

Box D. Content validity (including face validity)

		excellent	good	fair	poor
General requirements					
1	Was there an assessment of whether all items refer to relevant aspects of the construct to be measured?	Assessed if all items refer to relevant aspects of the construct to be measured		Aspects of the construct to be measured poorly described AND this was not taken into consideration	NOT assessed if all items refer to relevant aspects of the construct to be measured

2	Was there an assessment of whether all items are relevant for the study population? (e.g. age, gender, disease characteristics, country, setting)	Assessed if all items are relevant for the study population in adequate sample size (≥10)	Assessed if all items are relevant for the study population in moderate sample size (5-9)	Assessed if all items are relevant for the study population in small sample size (<5)	NOT assessed if all items are relevant for the study population OR target population not involved
3	Was there an assessment of whether all items are relevant for the purpose of the measurement instrument? (discriminative, evaluative, and/or predictive)	Assessed if all items are relevant for the purpose of the application	Purpose of the instrument was not described but assumed	NOT assessed if all items are relevant for the purpose of the application	
4	Was there an assessment of whether all items together comprehensively reflect the construct to be measured?	Assessed if all items together comprehensively reflect the construct to be measured		No theoretical foundation of the construct and this was not taken into consideration	NOT assessed if all items together comprehensively reflect the construct to be measured
5	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study

Box E. Structural validity

		excellent	good	fair	poor
1	Does the scale consist of effect indicators, i.e. is it based on a reflective model?				
Design requirements					
2	Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
3	Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
4	Was the sample size included in the analysis adequate?	7* #items and ≥100	5* #items and ≥100 OR 5-7* #items but <100	5* #items but <100	<5* #items
5	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study (e.g. rotation method not described)	Other important methodological flaws in the design or execution of the study (e.g. inappropriate rotation method)

Statistical methods					
6	for CTT: Was exploratory or confirmatory factor analysis performed?	Exploratory or confirmatory factor analysis performed and type of factor analysis appropriate in view of existing information	Exploratory factor analysis performed while confirmatory would have been more appropriate		No exploratory or confirmatory factor analysis performed
7	for IRT: Were IRT tests for determining the (uni-) dimensionality of the items performed?	IRT test for determining (uni)dimensionality performed			IRT test for determining (uni)dimensionality NOT performed

Box F. Hypotheses testing

		excellent	good	fair	Poor
Design requirements					
1	Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2	Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3	Was the sample size included in the analysis adequate?	Adequate sample size (≥100 per analysis)	Good sample size (50-99 per analysis)	Moderate sample size (30-49 per analysis)	Small sample size (<30 per analysis)

4	Were hypotheses regarding correlations or mean differences formulated a priori (i.e. before data collection)?	Multiple hypotheses formulated a priori	Minimal number of hypotheses formulate a priori	Hypotheses vague or not formulated but possible to deduce what was expected	Unclear what was expected
5	Was the expected *direction* of correlations or mean differences included in the hypotheses?	Expected direction of the correlations or differences stated	Expected direction of the correlations or differences NOT stated		
6	Was the expected absolute or relative *magnitude* of correlations or mean differences included in the hypotheses?	Expected magnitude of the correlations or differences stated	Expected magnitude of the correlations or differences NOT stated		
7	for convergent validity: Was an adequate description provided of the comparator instrument(s)?	Adequate description of the constructs measured by the comparator instrument(s)	Adequate description of most of the constructs measured by the comparator instrument(s)	Poor description of the constructs measured by the comparator instrument(s)	NO description of the constructs measured by the comparator instrument(s)
8	for convergent validity: Were the measurement properties of the comparator instrument(s) adequately described?	Adequate measurement properties of the comparator instrument(s) in a population similar to the study population	Adequate measurement properties of the comparator instrument(s) but not sure if these apply to the study population	Some information on measurement properties (or a reference to a study on measurement properties) of the comparator instrument(s) in any study population	No information on the measurement properties of the comparator instrument(s)

9	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study (e.g. only data presented on a comparison with an instrument that measures another construct)	Other important methodological flaws in the design or execution of the study
Statistical methods					
10	Were design and statistical methods adequate for the hypotheses to be tested?	Statistical methods applied appropriate	Assumable that statistical methods were appropriate, e.g. Pearson correlations applied, but distribution of scores or mean (SD) not presented	Statistical methods applied NOT optimal	Statistical methods applied NOT appropriate

Box G. Cross-cultural validity					
		excellent	**good**	**fair**	**poor**
Design requirements					
1	Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2	Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	

3	Was the sample size included in the analysis adequate?	CTT: 7* #items and ≥100 IRT: ≥200 per group	CTT: 5* #items and ≥100 OR 5-7* #items but <100 IRT: ≥200 in 1 group and 100-199 in 1 group	CTT: 5* #items but <100 IRT: 100-199 per group	CTT: <5* #items IRT: (<100 in 1 or both groups
4	Were both the original language in which the HR-PRO instrument was developed, and the language in which the HR-PRO instrument was translated described?	Both source language and target language described			Source language NOT known
5	Was the expertise of the people involved in the translation process adequately described? e.g. expertise in the disease(s) involved, expertise in the construct to be measured, expertise in both languages	Expertise of the translators described with respect to disease, construct, and language	Expertise of the translators with respect to disease or construct poor or not described	Expertise of the translators with respect to language not described	
6	Did the translators work independently from each other?	Translators worked independent	Assumable that the translators worked independent	Unclear whether translators worked independent	Translators worked NOT independent
7	Were items translated forward and backward?	Multiple forward and multiple backward translations	Multiple forward translations but one backward translation	One forward and one backward translation	Only a forward translation
8	Was there an adequate description of how differences between the original and translated versions were resolved?	Adequate description of how differences between translators were resolved	Poorly or NOT described how differences between translators were resolved		

9	Was the translation reviewed by a committee (e.g. original developers)?	Translation reviewed by a committee (involving other people than the translators, e.g. the original developers)	Translation NOT reviewed by (such) a committee		
10	Was the HR-PRO instrument pre-tested (e.g. cognitive interviews) to check interpretation, cultural relevance of the translation, and ease of comprehension?	Translated instrument pre-tested in the target population	Translated instrument pre-tested, but unclear if this was done in the target population	Translated instrument pre-tested, but NOT in the target population	Translated instrument NOT pre-tested
11	Was the sample used in the pre-test adequately described?	Sample used in the pre-test adequately described		Sample used in the pre-test NOT (adequately) described	
12	Were the samples similar for all characteristics except language and/or cultural background?	Shown that samples were similar for all characteristics except language /culture	Stated (but not shown) that samples were similar for all characteristics except language /culture	Unclear whether samples were similar for all characteristics except language /culture	Samples were NOT similar for all characteristics except language /culture
13	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study

Statistical methods					
14	for CTT: Was confirmatory factor analysis performed?	Multiple-group confirmatory factor analysis performed			Multiple-group confirmatory factor analysis NOT performed
15	for IRT: Was differential item function (DIF) between language groups assessed?	DIF between language groups assessed			DIF between language groups NOT assessed

Box H. Criterion validity					
		excellent	**good**	**fair**	**poor**
Design requirements					
1	Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2	Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3	Was the sample size included in the analysis adequate?	Adequate sample size (≥100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)
4	Can the criterion used or employed be considered as a reasonable ‘gold standard’?	Criterion used can be considered an adequate ‘gold standard’ (evidence provided)	No evidence provided, but assumable that the criterion used can be considered an adequate ‘gold standard’	Unclear whether the criterion used can be considered an adequate ‘gold standard’	Criterion used can NOT be considered an adequate ‘gold standard’

5	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
Statistical methods					
6	for continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated			Correlations or AUC NOT calculated
7	for dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated

Box I. Responsiveness					
		excellent	**good**	**fair**	**poor**
Design requirements					
1	Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2	Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3	Was the sample size included in the analysis adequate?	Adequate sample size (≥100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)
4	Was a longitudinal design with at least two measurement used?	Longitudinal design used			No longitudinal design used
5	Was the time interval stated?	Time interval adequately described			Time interval NOT described

6	If anything occurred in the interim period (e.g. intervention, other relevant events), was it adequately described?	Anything that occurred during the interim period (e.g. treatment) adequately described	Assumable what occurred during the interim period	Unclear or NOT described what occurred during the interim period	
7	Was a proportion of the patients changed (i.e. improvement or deterioration)?	Part of the patients were changed (evidence provided)	NO evidence provided, but assumable that part of the patients were changed	Unclear if part of the patients were changed	Patients were NOT changed
Design requirements for hypotheses testing					
	For constructs for which a gold standard was not available:				
8	Were hypotheses about changes in scores formulated a priori (i.e. before data collection)?	Hypotheses formulated a priori		Hypotheses vague or not formulated but possible to deduce what was expected	Unclear what was expected
9	Was the expected *direction* of correlations or mean differences of the change scores of HR-PRO instruments included in these hypotheses?	Expected direction of the correlations or differences stated	Expected direction of the correlations or differences NOT stated		
10	Were the expected absolute or relative *magnitude* of correlations or mean differences of the change scores of HR-PRO instruments included in these hypotheses?	Expected magnitude of the correlations or differences stated	Expected magnitude of the correlations or differences NOT stated		

11	Was an adequate description provided of the comparator instrument(s)?	Adequate description of the constructs measured by the comparator instrument(s)		Poor description of the constructs measured by the comparator instrument(s)	NO description of the constructs measured by the comparator instrument(s)
12	Were the measurement properties of the comparator instrument(s) adequately described?	Adequate measurement properties of the comparator instrument(s) in a population similar to the study population	Adequate measurement properties of the comparator instrument(s) but not sure if these apply to the study population	Some information on measurement properties (or a reference to a study on measurement properties) of the comparator instrument(s) in any study population	NO information on the measurement properties of the comparator instrument(s)
13	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study (e.g. only data presented on a comparison with an instrument that measures another construct)	Other important methodological flaws in the design or execution of the study
Statistical methods					
14	Were design and statistical methods adequate for the hypotheses to be tested?	Statistical methods applied appropriate		Statistical methods applied NOT optimal	Statistical methods applied NOT appropriate

Design requirement for comparison to a gold standard					
	For constructs for which a gold standard was available:				
15	Can the criterion for change be considered as a reasonable gold standard?	Criterion used can be considered an adequate ‘gold standard’ (evidence provided)	No evidence provided, but assumable that the criterion used can be considered an adequate ‘gold standard’	Unclear whether the criterion used can be considered an adequate ‘gold standard’	Criterion used can NOT be considered an adequate ‘gold standard’
16	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
Statistical methods					
17	for continuous scores: Were correlations between change scores, or the area under the Receiver Operator Curve (ROC) curve calculated?	Correlations or Area under the ROC Curve (AUC) calculated			Correlations or AUC NOT calculated
18	for dichotomous scales: Were sensitivity and specificity (changed versus not changed) determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated

Interpretability

We recommend to use the Interpretability box to extract all information on the interpretability issues described in this box of the instruments under study from the included articles.

Box Interpretability	
Percentage of missing items	
Description of how missing items were handled	
Distribution of the (total) scores	
Percentage of the respondents who had the lowest possible (total) score	
Percentage of the respondents who had the highest possible (total) score	
Scores and change scores (i.e. means and SD) for relevant (sub) groups, e.g. for normative groups, subgroups of patients, or the general population	
Minimal Important Change (MIC) or Minimal Important Difference (MID)	

Generalizability

We recommend to use the Generalizability box to extract data on the characteristics of the study populations and sampling procedures of the included studies.

Box Generalisability	
Median or mean age (with standard deviation or range)	
Distribution of sex	
Important disease characteristics (e.g. severity, status, duration) and description of treatment	
Setting(s) in which the study was conducted (e.g. general population, primary care or hospital/rehabilitation care)	
Countries in which the study was conducted	
Language in which the HR-PRO instrument was evaluated	
Method used to select patients (e.g. convenience, consecutive, or random)	
Percentage of missing responses (response rate)	

Appendix B

Studies which used a version of the Child Perceptions Questionnaire

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Agou[a] (Agou et al., 2008)	2008	CPQ_{11-14}		English	Canada	Clinical	3-73		Low socioeconomic status = 28.4 High socioeconomic status = 19.4
Agou[a] (Agou et al., 2011)	2011	CPQ_{11-14}		English	Canada	Clinical	0-80		Orthodontic treatment group at baseline = 21.05, At follow up = 16.16 Orthodontic waiting list group baseline = 24.07 At follow-up = 23.14
Aguilar-Diaz[b] (Aguilar-Diaz and Irigoyen-Camacho, 2011)	2011	CPQ_{8-10}	25	Spanish	Mexico	School		16.42	Low malocclusion = 6.39 Moderate malocclusion = 10.93 Severe = 21.03
Baker[a] (Baker et al., 2010)	2010	CPQ_{11-14}	37	Malay	Malaysia	School			
Barbosa[a,b] (Barbosa et al., 2009)	2009	CPQ_{8-10} CPQ_{11-14}	25 37	Portuguese	Brazil	School			
Barbosa[a,b] **(Barbosa et al., 2011**)	2011	CPQ_{8-10} CPQ_{11-14}	25 37	Portuguese	Brazil	School			CPQ_{8-10} TMD group = 20.6 Control = 13.5, CPQ_{11-14} TMD group = 27.6 Control = 16.3
Bekes[a] (Bekes et al., 2011)	2011	CPQ_{11-14}	35	German	Germany	Home survey		12.6	
Bendo[c] (Bendo et al., 2010)	2010	CPQ_{11-14}-ISF:16	16	Portuguese	Brazil	School			
Benson[a] (Benson et al., 2008)	2008	CPQ_{11-14}	31	English	UK	Clinical		20.4	
Brown[a] (Brown and Al-Khayal,	2006	CPQ_{11-14}	36	Arabic	Saudi Arabia	Clinic	0-102	24.22	No malocclusion = 17.38 Moderate/severe

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Cheretakis[a,b] (Cheretakis et al., 2007)	2007	CPQ $_{8-10}$ CPQ$_{11-14}$	25	English	Canada	Clinic	0-58		Neutropenia = 14.2 Control = 2.5
Costa[c] (Costa et al., 2011)	2011	CPQ$_{11-14}$-ISF:16		Portuguese	Brazil	School	0-42		Fixed appliances = 9.5 No active treatment = 11.5
de Oliveira[a] (de Oliveira et al., 2008)	2008	CPQ $_{11-14}$ C-OIDP	 8	English	UK	Clinic		CPQ = 18.41	
Do[a,b] (Do and Spencer, 2007)	2007	CPQ$_{8-10}$ CPQ$_{11-14}$		English	Australia	School			CPQ$_{8-10}$: dmfs/DMFS 0 =10.8 dmfs/DMFS 1-2 =9.2 dmfs/DMFS 3-4 =12.4 dmfs/DMFS 5+=10.2 Fluorosis 0=10.4 Fluorosis 1 =10.7 Fluorosis 2=10.1 Fluorosis 3=15.7 DAI ≤34=9.6 DAI ≥35=12.6 CPQ$_{11-14}$: dmfs/DMFS 0 =11.8 dmfs/DMFS1-2=14.5 dmfs/DMFS 3-4=15.4 dmfs/DMFS 5+=14 Fluorosis 0=15.1 Fluorosis 1=9.2 Fluorosis 2=6.4 Fluorosis 3=12.3 DAI ,34=13 DAI 35+=13.8
Do[a,b] (Do and Spencer, 2008)	2008	CPQ$_{8-10}$ CPQ$_{11-14}$	25 31	English	Australia	Clinic	CPQ $_{8-10}$ = 0-60 CPQ $_{11-14}$ = 0-77	CPQ8-10 = 10.7 11-14 = 13.3	

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Fakhruddin[a] (Fakhruddin et al., 2008)	2008	CPQ$_{11-14}$ (abbreviated version)	10	English	Canada	School			
Fergus[a] (Fergus, 2010)	2010	CPQ (assume CPQ$_{11-14}$)			Montserrat	School		24	Males = 24.28 Females = 23.44
Ferreira[c] (Ferreira et al., 2012)	2012	CPQ (abbreviated version)	14	Portuguese	Brazil	School	0-37	11.32	
Foster Page[a] (Foster Page et al., 2005)	2005	CPQ$_{11-14}$	35	English	New Zealand	School	0-103	17.3	Male = 15.8 Female = 18.9 DMFS 0 = 14.4 DMFS 1 = 16.9 DMFS 2/3 = 16 DMFS ≥4 = 21.8 Malocclusion: Minor/none = 14.8 Definite = 17.2 Severe = 19.6 Handicapping = 21.6
Foster Page[a,c] (Foster Page et al., 2008)	2008	CPQ$_{11-14}$ CPQ$_{11-14}$-ISF/RSF:16 CPQ$_{11-14}$-ISF/RSF:8	37 16 8	English	New Zealand	School	CPQ$_{11-14}$ = 0-110 RSF:16 = 0-52 RSF:8 = 0-23 ISF:16 = 0-47 ISF:8 = 0-24	CPQ$_{11-14}$ = 18.1 RSF:16 = 9.6 RSF:8 = 4.2 ISF:16 = 9.6 ISF:8 = 4.9	
Foster Page[a] (Foster Page et al., 2010)	2010	CPQ$_{11-14}$	37	English	New Zealand	School			
Foster Page[c] (Page et al., 2011)	2011	CPQ$_{11-14}$–ISF:16	16	English, Malay, Portuguese	New Zealand, Brunei, Brazil	School	0-49		Northland = 11.5 Brunei = 16.8 Brazil = 12.4 Otago = 14.6

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Foster Page[a] (Foster Page and Thomson, 2012)	2012	CPQ$_{11-14}$		English	New Zealand	School		At 13 years = 17.1 At 16 years = 17	At 13 years: DMFS 0 = 15.2 DMFS 1,2,3 = 17.1 DMFS ≥4+ = 17.9 At 16 years: DMFS 0 = 15.7 DMFS 1,2,3 = 16.9 DMFS 4+ = 19.1
Foster Page[c] (Foster Page et al., 2013b)	2013	CPQ$_{11-14}$-ISF:16	16	English	New Zealand	School		14.4	Male = 13.2 Female = 15.7 Maori = 11.1 Non-Maori = 14.7 DMFS 0 = 13.3 DMFS 2 = 15.4 DMFS ≥3 = 16 Malocclusion: minor/none = 13.5 definite = 14 severe = 14.3 handicapping = 15.6
Goursand[a] (Goursand et al., 2008)	2008	CPQ$_{11-14}$	37	Portuguese	Brazil	Clinic	0-88	16.23	Without untreated = 12.89 With untreated = 24.50
Gururatana[a,c] (Gururatana et al., 2011)	2011	CPQ $_{11-14}$ CPQ$_{11-14}$-ISF/RSF:16 CPQ$_{11-14}$-ISF/RSF:8	37 16 8	Thai	Thailand	School	CPQ$_{11-14}$ = 1-61 ISF:16 = 1-29 RSF:16 = 1-31 ISF:8 = 1-14 RSF:8 = 0-15	CPQ$_{11-14}$ = 24.3 ISF:16 = 13.4 RSF:16 = 14.1 ISF:8 = 6.8 RSF:8=6.9	
Humphris[b] (Humphris et al., 2005)	2005	CPQ $_{8-10}$	25	English	UK	School			

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Jokovic[a] (Jokovic et al., 2002)	2002	CPQ $_{11-14}$	37	English	Canada	Clinic	1-80	26.3	Paediatric group = 23.3 Orthodontic group = 24.3 Orofacial group = 31.4
Jokovic[a] **(Jokovic et al., 2003a)**	2003	CPQ $_{11-14}$	31	English	Canada	Clinic	2-56	26.2	
Jokovic[b] (Jokovic et al., 2004)	2004	CPQ $_{8-10}$	25	English	Canada	Clinic	1-55	18.6	Paediatric group = 18.4 Orofacial group = 19.1
Jokovic[a] (Jokovic et al., 2005)	2005	CPQ $_{11-14}$	37	English	Canada	Clinic			
Jokovic[a,c] (Jokovic et al., 2006)	2006	CPQ$_{11-14}$ CPQ$_{11-14}$-ISF:16 CPQ$_{11-14}$-ISF:8	37 16 8	English	Canada	Clinic	ISF:16 = 1-40 RSF:16 = 1-37 ISF:8 = 0-24 RSF:8 = 0-22	ISF:16 = 13.8 RSF:16 = 13.6 ISF:8 = 7.4 RSF:8 = 6.6	Paediatric group: ISF:16 = 11.9 RSF:16 = 11.9 ISF:8 = 6.8 RSF:8 = 5.3 Orthodontic group: ISF:16 = 13 RSF:16 = 13 ISF:8 = 7.7 RSF:8 = 6.4 Orofacial group: ISF:16 = 16.5 RSF:16 =15.9 ISF:8 = 8.8 RSF:8 = 8
Kadkhoda[c] (Kadkhoda et al., 2011)	2011	CPQ$_{11-14}$-RSF:8	8	Farsi	Iran	Clinic and school	0-22		Functional appliance = 6.9 Headgear = 6.5 No malocclusion = 2.2
Kohli[a] (Kohli et al., 2011)	2011	CPQ$_{11-14}$	29	English	USA	Clinical	2-80	31.6	Males =27 Females = 39.4
Kok[a] (Kok et al., 2004)	2004	CPQ$_{11-14}$	37	English	UK	School			

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Kolawole[a] (Kolawole et al., 2011)	2011	CPQ$_{11-14}$	37	English	Nigeria	School	0-81	23.44	Male = 22.7, Female = 24.17 DMFT 0 = 23.55 DMFT 1 = 18.57 DMFT 2 = 38 Dental Aesthetic Index: Minor = 25.11 Definite = 20.19 Severe = 20.47 Handicapping = 25.77
Koposova[a] (Koposova et al., 2010)	2010	CPQ (assume CPQ$_{11-14}$)		Norwegian Russian	Norway Russia	School			Russia = 20.2 Norway = 14.2 DMFT 0 = 14.8 DMFT >0 = 20.2
Koposova[a] (Koposova et al., 2012)	2012	CPQ$_{11-14}$	37	Russian Norwegian	Russia Norway	School	Norway = 0-34 Russia = 0-85	Norway = 9.9 Russia = 19.5	DMFT 0 = 15.8, DMFT >0 = 18.6
Kumar[a] (Kumar et al., 2011)	2011	CPQ (assume CPQ$_{11-14}$)	31 (scored 1-5)		India	School and orphanage			With parents = 132.92
Laing[a] (Laing et al., 2010)	2010	CPQ $_{11-14}$	37	English	UK	Clinic	2-80		Hypodontia = 26.82 Control = 28.52
Lau[a,c] (Lau et al., 2009)	2009	CPQ$_{11-14}$ CPQ$_{11-14}$-ISF/RSF:16 CPQ$_{11-14}$-ISF/RSF:8	37 16 8	Chinese	Hong Kong	School			
Locker[a] (Jokovic et al., 2005)	2005	CPQ$_{11-14}$	37	English	Canada	Clinic	1-80	27.8	Orofacial group = 31.4 Dental = 23.2

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Locker[a,c] (Locker et al., 2007)	2007	CPQ $_{11-14}$ CPQ$_{11-14}$-ISF:16 CPQ$_{11-14}$-ISF:8	35 16 8	English	Canada	Clinic	CPQ$_{11-14}$ = 1-75 ISF:16 = 1-41 ISF:8 = 0-19	CPQ$_{11-14}$ = 21.1 ISF:16 = 8.9 ISF:8 = 8.9	Malocclusion (CPQ$_{11-14}$): Minor/none = 16.1 Definite = 19.6 Severe = 17.7 Handicapping = 24.5
Locker[a] (Locker, 2007)	2007	CPQ $_{11-14}$ (abbreviated version)	10	English	Canada	School	10-32	12.9	Number of carious teeth: 0 = 12.8 1 = 12.9 ≥ 2 = 14.6 Aesthetic component IOTN: 1-4 = 12.7 5-7 = 13.1 8-10 = 13.8 Fluorosis: None = 12.7 Very mild = 13.5 Mild = 11.9
Locker[a] (Locker et al., 2010)	2010	CPQ$_{11-14}$	37	English	Canada	Clinic	4-69	22.3	
Luoto[a] (Luoto et al., 2009)	2009	CPQ$_{11-14}$	37	Finnish	Finland	Clinic/school			Attending dentist: Fear = 35.4 No fear = 25 General fear: Fear = 25.3 No fear = 29.9 Fear of treatment: Fear = 33.2 No fear = 24
Marshman[a] (Marshman et al., 2005)	2005	CPQ$_{11-14}$	37	English	UK	Clinic	3-53	18.07	

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Marshman[c] (Marshman et al., 2010)	2010	CPQ$_{11-14}$-ISF:16	16	English	UK	Clinic			
Martins[b] (Martins et al., 2009)	2009	CPQ$_{8-10}$	25	Portuguese	Brazil	Clinic			Caries = 14.77 Malocclusion = 10.7 Both = 21.29
Martins-Junior[b] (Martins-Junior et al., 2012a)	2012	CPQ $_{8-10}$	29	Portuguese	Brazil	School			
Martins-Junior[b] (Martins-Junior et al., 2012b)	2012	CPQ$_{8-10}$	25	Portuguese	Brazil	School	Caries = 0-54 No caries = 2-28		Caries = 23.1 No caries = 12
Massarente[a] (Massarente et al., 2011)	2008	CPQ $_{11-14}$	37	Portuguese	Brazil	Clinic	2-90	28.5	
McGrath[a] (McGrath et al., 2008)	2008	CPQ$_{11-14}$	37	Chinese	Hong Kong	Clinic	7-67	23.1	Paediatric group = 19.3 Orthodontic group = 26.9

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
O'Brien[a] (O'Brien et al., 2006)	2006	CPQ$_{11-14}$	36	English	UK	School			Medians: Male = 50, Female = 54 Townsend deprivation score: Low = 52 High =52 Child's Aesthetic component of IOTN score: 1-3 = 51 4-10=54 Dentist's aesthetic component score of IOTN: 1-5 = 51 6-10 = 53 Dental health component of IOTN: 2-3 = 51 4-5 = 54 No orthodontic treatment = 51 Previous orthodontic treatment = 60 Current orthodontic treatment = 58 On waiting list = 50
O'Brien[a] (O'Brien et al., 2007)	2007	CPQ$_{11-14}$		English	UK	Clinic	0-57	Malocclusion group = 17.6 (median) Control group = 14 (median)	Crowding = 17.1 (median) Overjet = 20 (median) Hypodontia = 18 (median)
Paula[a] (Paula et al., 2012b)	2012	CPQ$_{11-14}$	35	Portuguese	Brazil	School	0-106	23.24	

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Paula[b] (Paula et al., 2012a)	2012	CPQ$_{8-10}$		Portuguese	Brazil	School	Controls = 0-51 Caries = 3-96 Post treatment = 0-100		Controls = 19 Caries = 42 Post treatment = 29
Piovesan[a] (Piovesan et al., 2010)	2010	CPQ$_{11-14}$	37	Portuguese	Brazil	School	0-99	20.9	
Piovesan[a] (Piovesan et al., 2011)	2011	CPQ$_{11-14}$	37	Portuguese	Brazil	School			
Piovesan[a] (Piovesan et al., 2012)	2012	CPQ$_{11-14}$	37	Portuguese	Brazil	School		18.4	Trauma = 21.03, No trauma = 20.85
Porritt[c] (Porritt et al., 2011)	2011	CPQ$_{11-14}$-ISF:16	16	English	UK	Clinic	0-51	15.5 at baseline 12.1 at follow up	
Ramos-Jorge[b,c] (Ramos-Jorge et al., 2012)	2012	CPQ$_{8-10,}$ CPQ$_{11-14}$-ISF:8	25 8	Portuguese	Brazil	School			Interview 1st: CPQ$_{8-10}$= 13.5 ISF:8= 8.7 Interview 2nd: CPQ$_{8-10}$= 21.9 ISF:8= 8.9 Self completion 1st: CPQ$_{8-10}$= 12.4 ISF:8= 8.2 Self completion 2nd CPQ$_{8-10}$ = 15.8 ISF:8= 7.9

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Robinson[a] (Robinson et al., 2005)	2005	CPQ_{11-14}	37	Luganda	Uganda	School			Fluorosis score: >2 = 33.1, 0-2 = 25.4 DMFT 0 = 21.9, DMFT >0 = 32.5
Taylor[a] (Taylor et al., 2009)	2009	CPQ_{11-14}	37	English	USA	Clinic			
Torres[c] (Torres et al., 2009)	2009	CPQ_{11-14}-ISF:16 CPQ_{11-14}-ISF:8	16 8	Portuguese	Brazil	School	ISF:8 = 0-18 ISF:16 = 0-38	ISF:8 = 6.8 ISF:16 = 11.9	
Traebert[a] (Traebert et al., 2012)	2010	CPQ_{11-14}	37	Portuguese and English	Brazil and New Zealand	School			New Zealand = 17.8 Portuguese = 28.1
Traebert (Traebert et al., 2010b)	2010	CPQ_{11-14}	37	English	New Zealand	School	0-110	18.1	Maori = 19.1 Non-Maori = 17.9
Traebert[c] (Traebert et al., 2012)	2012	CPQ_{11-14} short form	16	English	New Zealand	School	0-49	12.4	Trauma =14.6 Without = 9.6
Vargas-Ferreira[a] (Vargas-Ferreira et al., 2010)	2010	CPQ_{11-14}	37	Portuguese	Brazil	School	0-94	18.59	Male = 17.7 Female = 19.25 White = 18.19, Non-white = 20.32 11-12 years = 18.91 13-14 years = 17.94 No erosion = 18.5, Erosion = 19.9

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Wogelius[a,b] (Wogelius et al., 2009)	2009	$CPQ_{8\text{-}10}$ $CPQ_{11\text{-}14}$	25 37	Danish	Denmark	School, clinic			$CPQ_{8\text{-}10}$: Healthy = 8.5 Cleft lip and palate = 7.9 Rare oral diseases =16.3 $CPQ_{11\text{-}14}$: Healthy = 10.5 Cleft lip and palate = 17.8 Rare oral diseases = 17.8 Fixed appliances = 24.4
Wogelius[a,b] (Wogelius et al., 2011)	2011	$CPQ_{8\text{-}10}$ $CPQ_{11\text{-}14}$	25 37	Danish	Denmark	School, clinic	$CPQ_{8\text{-}10}$ = 0-100 $CPQ_{11\text{-}14}$ = 0-148		$CPQ_{8\text{-}10}$: With cancer = 5.6, Without = 8.8, $CPQ_{11\text{-}14}$: With cancer = 12.5 Without = 11.8
Wong[a] (Wong et al., 2006)	2006	$CPQ_{11\text{-}14}$	37	Chinese	Hong Kong	Clinic		29	
Wong[c] (Wong et al., 2011)	2011	$CPQ_{11\text{-}14}$-RSF:8	8	Chinese	Hong Kong	School			
Zhang[a] (Zhang et al., 2007b)	2007	$CPQ_{11\text{-}14}$	31	Chinese	Hong Kong	Clinic		20.1	
Zhang[a] (Zhang et al., 2007a)	2007	$CPQ_{11\text{-}14}$	37	Chinese	Hong Kong	Clinic			Anticipation of changes following treatment = 27.4 Pre-treatment = 20.7 After 1 week of treatment = 25.6 After 1 month of treatment = 22.5 After 3 months of treatment = 21.1 After 6 months of treatment = 20.4

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Zhang[a] (Zhang et al., 2008)	2008	CPQ (assume CPQ$_{11-14}$)	37	Chinese	Hong Kong	Clinic			Pre-treatment = 20.7 After 1 week of treatment = 25.6 After 1 month of treatment = 22.5 After 3 months of treatment = 21.1 After 6 months of treatment = 20.4
Zhang[a] (Zhang et al., 2009)	2009	CPQ$_{11-14}$	37	Chinese	Hong Kong	Clinic			Aesthetic component Need = 24.2 No need = 19.7, Dental Health Component Need = 21.7 No need = 18.3

Appendix B. Studies which used a version of the Child Perceptions Questionnaire with details of version, setting and range and mean scores.

CPQ = Child Perceptions Questionnaire; ISF:8/16 = item impact version of the short form of CPQ$_{11-14}$ (8 or 16 items respectively); RSF:8/16 = regression version of the short form of CPQ$_{11-14}$(8 or 16 items respectively).

TMD = temporomandibular joint dysfunction; dmfs/DMFS = decayed missing and filled surfaces (primary and permanent teeth respectively); dmft/DMFT = decayed missing and filled teeth (primary and permanent teeth respectively); DAI = Dental Aesthetic Index; IOTN = Index of Treatment Need. [a]Included in Section 5.4.1 CPQ$_{11-14}$; [b]included in Section 5.4.1 CPQ$_{8-10}$; [c] included in Section 5.4.1 CPQ short forms.

Appendix C

Studies which used a version of the Child Oral Impacts on Daily Performances index

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Bernabe (Bernabe et al., 2007b)	2007	C-OIDP	8	Spanish	Peru	School		1.3	Proportion with impacts: Malocclusion= 15.5%
Bernabe (Bernabe et al., 2007a)	2007	C-OIDP	8	Spanish	Peru	School	0-62.5	7.8	82% with at least one impact
Bernabe (Bernabe et al., 2008)	2008	C-OIDP	8	Spanish	Peru	School			
Bernabe (Bernabe et al., 2009)	2009	C-OIDP	8	Thai	Thailand	School			
Bianco (Bianco et al., 2010)	2010	C-OIDP	8	Italian	Italy	School	0-30	1.9	66.8% reported at least one impact
Castro (Castro et al., 2008)	2008	C-OIDP	8	Portuguese	Brazil	School		9.2	80.7% with at least one impact
Castro (Castro et al., 2011)	2011	C-OIDP	8	Portuguese	Brazil	School		7.1	88.7% with at least one impact Male = 6.2 Female =7.8 DMFT>1 = 8.2 DMFT 0 = 6.3 Biofilm present =8.3 No biofilm = 6.1 Enamel defects = 9.2 No enamel defects = 7 Trauma = 7.4 No trauma = 7.1
Cortes-Martinicorena (Cortes-Martinicorena et al., 2010)	2010	C-OIDP	8	Spanish	Spain	School			11-12 years = 2.69 13-14 years = 3.08

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Dumitrache (Dumitrache et al., 2009)	2009	C-OIDP	8		Romania	School	Total negative effects 57.4%		
Gherunpong (Gherunpong et al., 2004a)	2004	C-OIDP	8	Thai	Thailand	School			Perceived treatment need = 18.1 (median) No perceived treatment need = 5.6 (median) Perceived oral health problems: None/little = 5.6 Moderate = 13.9 Severe = 33.3
Gherunpong (Gherunpong et al., 2004b)	2004	C-OIDP	8	Thai	Thailand	School	0-59.7	8.8	89.8% with at least one impact
Gherunpong (Gherunpong et al., 2006a)	2006	C-OIDP	8	Thai	Thailand	School			89.8% with at least one impact 20.3% of impacts related to malocclusion
Gherunpong (Gherunpong et al., 2006b)	2006	C-OIDP	8	Thai	Thailand	School			
Krisdapong (Krisdapong et al., 2009)	2009	C-OIDP	8	Thai	Thailand	School	0-68.1	7.8	85.2% with at least one impact
Krisdapong (Krisdapong et al., 2012a)	2012	C-OIDP	8	Thai	Thailand	School			26% had impacts related to calculus/gingivitis
Krisdapong (Krisdapong et al., 2012b)	2012	C-OIDP	8	Thai	Thailand	School			24.7% of impacts related to recurrent apthous ulceration
Krisdapong (Krisdapong et al., 2012c)	2012	C-OIDP	8	Thai	Thailand	School			81.7% had a least one impact
Krisdapong (Krisdapong et al., 2012d)	2012	C-OIDP	8	Thai	Thailand	School			85.1% with at least one impact

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Mashoto (Mashoto et al., 2010)	2010	C-OIDP	8	Kiswahili	Tanzania	School		T0 = 1.8 Follow up = 1	T0: ART = 1.3 ART and extractions = 3.9 OHE = 1.6 Follow-up: ART = 1.5 ART and extractions = 1.7 OHE = 0.9
Mbawalla (Mbawalla et al., 2010)	2010	C-OIDP	8	Kiswahili	Tanzania	School			48.2% with at least one impact
Mbawalla (Mbawalla et al., 2011)	2011	C-OIDP	8	Kiswahili	Tanzania	School			Oral impacts: Arusha = 0.7% Dar es Salaam = 28.6% Decayed teeth: 0 = 0.5 >0 = 0.8 Missing teeth: 0 = 0.6 >0 = 0.8 Oral hygiene: Good = 0.5 Bad = 0.7 Calculus: Good = 0.6 Poor= 0.7
Mtaya (Mtaya et al., 2007)	2007	C-OIDP	8	Kiswahili	Tanzania	School		1.2	DMFT 0 = 1.1 DMFT >0 = 1.5
Mtaya (Mtaya et al., 2008)	2008	C-OIDP	8	Kiswahili	Tanzania	School			28.6% had at least one oral impact
Nurelhuda (Nurelhuda et al.)	2010	C-OIDP	8	Arabic	Sudan	School	0-9	1.5	54.6% had a least one impact 53.4% of those in public school 64% in private school
Pau (Pau et al., 2008)	2008	C-OIDP	8	English	Pakistan	School		3.4	

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Raymundo de Andrade (Raymundo de Andrade et al., 2011)	2011	C-OIDP	8	Portuguese	Brazil	Clinic	0-33	6.1	Pain = 8.97 No pain = 4.4 Sensitivity = 7.65 No sensitivity = 4.48 Caries = 7.03 No caries = 5.34
Rosel (Rosel et al., 2010)	2010	C-OIDP	8	Spanish	Spain	School			Face to face interview 1st =4.38 Face to face interview 2nd = 4.21 Self-administered 1st = 4.2 Self-administered 2nd = 4.46
Tsakos (Tsakos et al., 2006)	2006	C-OIDP	8	Thai	Thailand	School			Presence of impact: Caries = 50.6% Trauma = 4.6% Enamel defect/anomaly = 8.7% Malocclusion = 20.3% Prosthodontic = 0.7%
Tsakos (Tsakos et al., 2008)	2008	C-OIDP	8	English	UK	Clinic			Mean number of impacts = 3.16
Tubert Jeannin (Tubert-Jeannin et al., 2005)	2005	C-OIDP	8	French	France	School		6.3	Not satisfied with oral health = 7.5 Satisfied with oral health = 2 Oral problems: None = 0 Some = 4.2 Many = 19.2 No treatment need = 1 Treatment need = 5.2 Global oral health: Very bad = 14.6 Bad = 19.2 Fairly bad = 8 Fairly good = 4.2 Good = 1.7 Excellent = 0.2
Yusuf (Yusuf et al., 2006)	2006	C-OIDP	8	English	UK	School			40.4% had at least one impact

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Yusof (Yusof and Jaafar, 2012)	2012	C-OIDP	8	Malay	Malaysia	School		13.2	66.7% had at least one impact

Appendix C. Studies which used a version of the Child Oral Impacts on Daily Performances index with details of version, setting and range and mean scores.

C-OIDP = Child Oral Impacts on Daily Performances index; dmft/DMFT = decayed missing and filled teeth (primary and permanent teeth respectively); ART = atraumatic restorative technique; OHE = oral health education; T0 = baseline measurement.

Appendix D

Studies which used a version of the Child Oral Health Impact Profile

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Ahn (Ahn et al., 2012)	2011	COHIP	34	Korean	Korea	School	46-132	103.3	Caries = 106.8 No caries = 104.2 Ortho need = 100.4 No need = 104.9
Bos (Bos et al., 2010)	2010	COHIP	30	Dutch	Netherlands	Clinic		112.12	Female = 110.92 Male = 113.28
Bos (Bos and Prahl, 2010)	2010	COHIP	30	Dutch	Netherlands	Clinic		119.29	Female = 118.2 Male = 119.81 8-12 years = 121.77 >12 years = 116.19 Cleft palate = 116.29 Cleft lip = 121.65 Unilateral cleft lip and palate = 119.2 Bilateral cleft lip and palate = 119.25
Broder (Broder et al., 2007)	2007	COHIP		English, French, Spanish	USA/Canada	Clinic			
Broder (Broder	2007	COHIP	34	English,	USA/Canada	Clinical/sch	28-135	99	Paediatric group = 97.7

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Broder (Broder et al., 2012)	2012	COHIP short form	19	English Spanish	USA	Agricultural population/ clinical			Caries: COHIP = 103.5 Short form = 56.2 Malocclusion: COHIP = 102.2 Short form = 57.8, Craniofacial group: COHIP = 98.3 Short form = 53.7
Calis (Calis et al., 2009)	2009	COHIP	38	Dutch	Netherlands	School			
Dunlow (Dunlow et al., 2007)	2007	COHIP	34	English	USA	Clinic			
Geels (Geels et al., 2008a)	2008	COHIP	38	Dutch	Netherlands	School		1.78	
Geels (Geels et al., 2008b)	2008	COHIP	38	Dutch	Netherlands	Clinic	45-122	74.1	

Author	Year	Measure used and version	Number of items analysed	Language of version	Study country	Study population	Item/score distribution	Total mean score	Subgroups mean score/ proportion with impacts
Ravaghi (Ravaghi et al., 2011)	2011	COHIP	34	Farsi	Iran	School	0-82	28	Perceived dental need = 32.6 No perceived need = 22.8 Pain = 35.1 No pain = 25.6 Dissatisfied with oral health = 38.5 Satisfied with oral health = 25.1 DMFT 0 = 20 DMFT >0 = 29.6 DMFT 0-2 = 25.8 DMFT >3 = 30.4
Wilson-Genderson (Wilson-Genderson et al., 2007)	2007	COHIP	34	English	Canada/USA	Clinic			

Appendix D. Studies which used a version of the Child Oral Health Impact Profile with details of version, setting and range and mean scores.

COHIP = Child Oral Health Impact Profile; dmft/DMFT = decayed missing and filled teeth (primary and permanent teeth respectively).

Appendix E

Patient Information Leaflet

Participant information leaflet: in-depth interviews

The
University
Of
Sheffield.

Teeth and dentists: what do you think?

Who am I?

My name is Fiona, I work at the University. This is me

What am I doing?

I want to find out what children think about their teeth.

Can you help me?

I would like to talk to you about your teeth. It should be fun!

What happens after?

Everything you tell me will be a secret.

You can choose a different name, so that no one will know what you've said.

I will write a story to tell other people what I've found out.

What if you don't want to join in anymore?

If you don't want to do it anymore, you can stop at any time.

No one will be cross.

What do you do now?

There is a sheet to fill in. If you would like to join in, please tick the box on the sheet and get someone to post it to me.

If you have any questions or are unhappy, you can phone me on 0114 271 7885 on or email f.gilchrist@sheffield.ac.uk or Mrs Tracey Plant, Clinical Hospital Manager, Charles Clifford Dental Hospital, Sheffield S10 2SZ or the Patient Services Team on 0114 271 2400 or email PST@sth.nhs.uk

Thank you for reading this. I hope to meet you soon.

Appendix F

Assent forms

Assent form for both qualitative and quantitative studies (5-10 years)

Name of young person to be involved in the research: ____________________________

Participant identification number for this project: ____________________________

	Please put a circle round the one
1. Has someone read to you information about this study?	Yes / No
2. Has somebody else told you what this study is about?	Yes / No
3. Do you understand what this study is about?	Yes / No
4. Have you asked all the questions you want?	Yes / No
5. Have your questions been answered OK?	Yes / No
6. Do you understand it's OK to stop taking part at any time?	Yes / No
7. Are you happy to take part?	Yes / No

If any answers are 'no' or you don't want to take part, don't write your name!

If you <u>do</u> want to take part, you can write your name below.

Name of child **(please PRINT name)**	Date	Child to write name here
Name of person taking consent	Date	Signature

Teeth and dentists: what do you think?

Participant assent form (qualitative, 11-16 years)

Young person, please complete this sheet.

Name of young person to be involved in the research:

____

Participant identification number for this project:

____

Please initial box:

1. I confirm that I have read and understood the information sheet for the above project and that I have had the opportunity to ask questions ☐
2. I understand that tapes will be made and that the reason they will be used has been explained in terms which I have understood ☐
3. I understand that my participation is entirely voluntary and that we are free to withdraw at any time, without giving a reason and without our medical care or legal rights being affected ☐
4. I understand that any information will be used for research purposes only; including research publications and reports. Anonymity and confidentiality will be preserved at all times. ☐
5. I give my permission to be contacted again within a three year period ☐
6. I agree to take part ☐

Name of young person: ______________________________

Signature: ______________________________

Date: ______________________________

Name of researcher: ______________________________

Signature: ______________________________

Date: ______________________________

1 copy for parent/guardian, 1 for researcher

Appendix G

Item reduction data collection sheet

Item reduction

Participant number:

Item	Experienced?	Most frequent/most severe
Felt annoyed		
Not been able to do things		
Schoolwork		
Felt grumpy		
Felt tired		
Cried		
Food stuck		
Eat slowly		
Couldn't eat some foods		
Eat on other side		
Eat carefully		
Kept awake		
Pain		
Medicine		
Talking		
Brown marks		
Toothbrushing		

Appendix H

Questionnaire design stage:

Version One

These questions ask how you feel about your teeth. Read all the answers and see which one is most like you.

Put a tick in the box next to it, like this ☑. Only tick one box for each question.

For example: I feel very happy about my teeth, so I will tick this box:

Happy

I don't feel happy about my teeth ☐

I feel a bit happy about my teeth ☐

I feel very happy about my teeth ☑

Now please think about your teeth and answer the questions on the next pages.

Please tick one answer for each question.

1. Hurting

My teeth do not hurt ☐

My teeth hurt a bit ☐

My teeth hurt a lot ☐

2. Eating food

My teeth don’t stop me eating any foods ☐

My teeth stop me eating some foods a bit ☐

My teeth stop me eating some foods a lot ☐

3. Getting food stuck

I don’t get food stuck between my teeth ☐

I get food stuck between my teeth a bit ☐

I get food stuck between my teeth a lot ☐

4. Eating differently

I don't have to eat on the other side of my mouth because of my teeth .... ☐

I sometimes have to eat on the other side of my mouth because of my teeth .... ☐

I have to eat on the other side of my mouth because of my teeth a lot .... ☐

5. Sleeping

I don't get kept awake by my teeth .... ☐

I get kept awake by my teeth a bit .... ☐

I get kept awake by my teeth a lot .... ☐

6. Annoyed

I don't feel annoyed by my teeth .... ☐

I feel annoyed by my teeth a bit .... ☐

I feel very annoyed by my teeth .... ☐

Appendix I

Questionnaire design stage:

Version Two

These questions ask how you feel about your teeth. Read all the answers and see which one is most like you.

Put a circle round the answer like this ◯ . Only make one circle for each question.

For example: I feel very happy about my teeth, so I will circle this:

How happy do you feel about your teeth?

Not at all A bit [✓ A lot]

Now please think about your teeth and answer the questions on the next pages.

1. How much do your teeth hurt you?

Not at all A bit A lot

2. How much do your teeth stop you eating some foods?

Not at all A bit A lot

3. Do you have to eat on the other side of your mouth because of your teeth?

Not at all A bit A lot

4. How much do you get food stuck between your teeth?

Not at all A bit A lot

5. How much do you get kept awake by your teeth?

Not at all A bit A lot

6. How much do your teeth annoy you?

Not at all A bit A lot

Appendix J

Topic guide for interviews

Topic guide in depth interviews

- Attending the dental hospital
    - Going to the dental hospital, what was that for?
    - Who told them?
    - What was it like?
- Experience of having caries
    - Did they have any symptoms?
    - Did it stop them doing anything?
    - How did it make them feel?
    - When symptoms stopped, why did it feel better?
    - What do they think it looks like?
- Treatment
    - Have they had any treatment?
    - What was it?
    - What was it like?
    - How did it make them feel?
    - Did it make it better?
- Describing it to a friend
    - How would they describe caries to a friend
    - How would they describe the treatment they had to a friend
- Things other children have said
    - Stops them eating on those teeth
    - Stops them talking
    - Stops them eating some foods
    - Makes them eat slower
- Anything else?

Appendix K

Item generation stage participant characteristics

Pseudonym	Gender	Age	Deprivation quintile	Ethnicity	dmft/DMFT	Presence of anterior caries	Pain experienced
Marie	F	12years, 11months	2	White British	8/3	No	Yes
Annie	F	8 years, 5 months	5	White British	6/1	No	Yes
Maria	F	8 years, 6 months	4	White British	4/0	No	Yes
Wayne	M	9 years, 5 months	4	White British	4/2	No	Yes
Steve	M	12 years, 11 months	3	White British	8/1	No	No
Justin	M	7 years, 3 months	5	White British	6/0	No	Yes
Precious	F	11years, 3 months	5	White British	0/3	No	Yes
Didier	M	8 years 7 months	5	White British	2/0	No	No
Bob	M	11 years, 3 months	3	White British	3/0	No	Yes
Jack	M	11 years, 11 months	1	White British	1/0	No	Yes
Lucy	F	10 years, 7 months	5	White British	3/0	No	Yes
Isabelle	F	7 years, 1 months	3	White British	7/0	No	No
Elisabeth	F	7 years, 6 months	5	Bangladeshi	8/0	No	No
Ahmed	M	5 years, 6 months	5	Black African	12/0	Yes	No
Lily	F	12 years, 3 months	5	White British	0/5	No	Yes
Leah	F	6 years, 11 months	5	White British	8/0	Yes	Yes
Mark	M	8 years, 7 months	2	White British	1/1	No	Yes
Liam	M	13 years, 11 months	5	White British	0/7	Yes	No
Brodie	M	9 years, 6 months	3	White British	6/3	No	Yes
Danesh	M	6 years, 9 months	5	Pakistani	1/0	No	No

Appendix K. Interview and focus group participant characteristics.

dmft = number of decayed, missing and filled primary teeth, DMFT = number of decayed, missing and filled permanent teeth. Deprivation quintiles = 1 is least deprived and 5 most deprived.