THE USE OF THE EQ-5D-Y HEALTH RELATED QUALITY OF LIFE INSTRUMENT IN CHILDREN IN THE WESTERN CAPE, SOUTH AFRICA: PSYCHOMETRIC PROPERTIES, FEASIBILITY AND USEFULNESS

By
DESIREE SCOTT
Student number: THMDES001

SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE
MASTER OF SCIENCE IN PHYSIOTHERAPY

Date of submission: 18 August 2015

Supervisor: Professor J Jelsma
University of Cape Town
Faculty of Health Sciences
Department of Health and Rehabilitation Sciences
Division of Physiotherapy

Co-supervisor: Dr G Ferguson
University of Cape Town
Faculty of Health Sciences
Department of Health and Rehabilitation Sciences
Division of Physiotherapy

Declaration

I, Desiree Scott, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise). I declare that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: Signed by candidate

Date: 18/08/2015

Acknowledgements

I would like to acknowledge that this research was conducted out of a need which was identified by the EuroQoL Research Foundation, following the original reliability and validity study, of the paediatric outcome measure, the EQ-5D-Y (2010) (1). This study used in the main, healthy typically developing children to determine reliability and validity of the measure and it was suggested that these psychometric properties also be assessed when used in a clinical sample. This research study was therefore undertaken to determine the performance of the EQ-5D-Y in clinical sample and typically developing children. The funding received from the EuroQoL Research Foundation (a non-profit organisation) is gratefully acknowledged.

I would like to express my enormous gratitude to my main supervisor, Professor Jennifer Jelsma, for her enthusiastic support and invaluable guidance throughout this research study. Without her encouragement and belief in my capabilities this dissertation would not have been possible. I have been honoured to have had the opportunity to learn from such an expert researcher, who has instilled a keen interest in research in me.

My sincere thank you to my co-supervisor, Dr Gillian Ferguson, for her insightful input and suggestions, particularly during the difficult write up phase.

To the clinicians at the various research facilities and to the teachers at the Main Stream school, thank you all for your assistance in the data collection and for allowing me to interact with the children.

A special mention needs to made of the wonderful children who participated in this study. Without their willingness to engage, there would not have been a research study.

My husband also deserves an enormous thank you for his support and encouragement, during the many long weekends of work.

Potential conflict of interest:

The research was funded by the EuroQoL Research Foundation and hence the EQ-5D-Y was the primary Health Related Quality of Life outcome measure used. The Foundation, however, had no involvement in the analyses of data or interpretation of results.

Abstract

Aim:

The overall aim of the study was to investigate the performance of the EQ-5D-Y, a self-reported Health Related Quality of Life (HRQoL) outcome measure, in children between eight and twelve years of age. The study objectives were to examine the measure's psychometric properties of criterion validity, discriminant and concurrent validity, when used on children with different health statuses, and to determine its ability to detect change within the different groups over a period of time. The study also set out to determine whether a life event had an impact on HRQoL, and whether children and their therapists or parents shared the same perceptions of HRQoL. The feasibility of using the EQ-5D-Y as a routine, additional, physiotherapy assessment tool was monitored. The study also assessed the usefulness of the collected data to the therapists administering the measure to children under their management.

Method:

A longitudinal, analytical descriptive study design was used. Typically developing children attending a Main Stream (MS) school (105), children with lifelong physical disabilities at a Special School (SS) (35), chronically ill children at an institution (CI) (32) and acutely ill children in hospital (AI) (52) were recruited. The EQ-5D-Y was the primary outcome measure, and was administered at baseline and again at three monthly intervals, or, in the case of AI children, at admission and discharge. The PedsQL as a parallel HRQoL measure, the WeeFim as a functional measure and the Faces Pain Scale (FPS) to monitor pain were used. A self-designed questionnaire was completed by the therapists treating the children to assess feasibility and usefulness of the EQ-5D-Y.

Data analysis:

Descriptive statistics were used to describe the sample and the health conditions of the participants. Reliability of the measures was determined at different time intervals by Cohen's kappa coefficient for dimension scores. Spearman's rho and Intraclass Correlation (ICC) were used to determine reliability of Visual Analogue Scale (VAS) scores and also total scores of the measures over time. The same analysis was used to compare self-reports and proxy reports. Kruskal-Wallis ANOVA by ranks, median scores and mean rankings were used to examine discriminant validity between known groups, using the same outcome measure and convergent validity between similar dimensions on different outcome measures. Responsiveness was described by examining the effect size of the Wilcoxon Signed-rank test. The VAS score was compared against the ranking of different levels of the dimensions, across groups, using Kruskal Wallis H statistic. A discrepancy between changes in VAS and changes in Worried, Sad or Unhappy (WSU) dimension were examined after three months to determine whether these were related to life events and/or changes in management of health condition. The clinical feasibility of using the EQ-5D-Y and its usefulness as an additional evaluation tool in providing a holistic assessment of the child's condition was established by analysing the frequency of positive responses on the questionnaire.

Results:

A total of 224 children were recruited. The level of problems on the dimensions was associated with institution and in all cases, apart from Mobility, the AI children reported more problems. The EQ-5D-Y only demonstrated discriminant validity between the MS children and AI children. The MS group scored significantly lower ranked scores on all dimensions and a significantly higher VAS (better overall HRQoL) compared to the AI group with more problems on each dimension and lower VAS. When comparing VAS across the mean ranking on each dimension, it was found to be significantly correlated at the AI only.

Convergent validity between EQ-5D-Y and PedsQL was evident only at the AI for all similar dimensions. The other groups demonstrated convergent validity with some, but not all of the dimensions. Convergent validity was evident between the EQ-5D-Y VAS and total scores of PedsQL and WeeFIM ($p<.05$ in all cases) across institutions.

The treatment effect over time was largest in the AI.

For all groups, there was limited agreement between proxy and self-report at a dimension level, except for Mobility with moderate to good agreement. Even though the proxy and self-report VAS scores demonstrated good (.58) ICC overall, at an institutional level, this was only significant in the MS children.

The EQ-5D-Y only took five minutes to complete. Six of the nine therapists who took part in the study, found the measure easy to apply, used the information in the management of the child and would continue to use it in future.

Conclusion:

The performance of the EQ-5D-Y, as determined by the psychometric properties, was variable. It could discriminate between children with an acute illness and children in the MS school. In addition, good convergent validity was demonstrated in the AI children and the largest treatment effect was observed in these children. However, it does not perform as well in children with no health condition or chronic conditions and should be used with caution in these groups. HRQoL did not appear to be linked to a life event. It is recommended that both proxy and self-report measures be taken into account when assessing a child's HRQoL but these should not be used interchangeably. It appears to be feasible and useful to include the EQ-5D-Y in routine assessments. It was concluded that the EQ-5D-Y self-report can be used with confidence as an outcome measure for acutely-ill children.

Table of Contents

Declaration ii
Acknowledgements iii
Abstract iv
Table of Contents vi
List of Tables xi
List of Figures xiv
Glossary of terms xv
1 Chapter 1: Introduction and Scope of Thesis 1
1.1 Introduction 1
1.2 Aims and objectives 2
1.2.1 Aim 2
1.3 Rationale and significance of the study 3
1.4 Outline of the study 3
2 CHAPTER 2: LITERATURE REVEIW 3
2.1 Introduction 3
2.2 Quality of Life 4
2.3 Health Related Quality of Life 5
2.3.1 HRQoL in children 6
2.4 Reasons for assessing HRQoL 7
2.4.1 Clinical practice and patient empowerment 7
2.4.2 Research and clinical trails 8
2.4.3 Evaluating cost effectiveness and use of Quality Adjusted Life Years (QALYs) 8
2.4.4 Health utility measurements 8
2.5 Generic or specific HRQoL measures 9
2.6 Paediatric HRQoL (p-HRQoL) measures 10
2.6.1 Essential elements of p-HRQoL measures 11
2.6.2 Value of p-HRQoL measures 11
2.6.3 The EQ-5D-Y 11
2.6.4 Challenges in p-HRQoL 12
2.7 Psychometric properties of HRQoL outcome measures 14
2.7.1 Reliability 15
2.7.2 Validity (content, criterion and construct) 17

2.7.3 Responsiveness (and ceiling/floor effect) 20
2.7.4 Interpretability 20
2.8 Review of some generic HRQoL Paediatric Outcome measures 21
2.8.1 Introduction 21
2.9 Limitations in previous assessment of psychometric properties of EQ5D-Y 31
2.10 Summary of literature review 31
3 CHAPTER 3: METHODOLOGY 32
3.1 Introduction 32
3.2 Research Settings 32
3.2.1 Geography and Demographics 32
3.2.2 Description of facilities from which the children for the study were drawn 33
3.3 Study participants 34
3.3.1 Sample size calculation 35
3.3.2 Inclusion and exclusion criteria 36
3.3.3 Sample for pilot testing 36
3.4 Instrumentation 36
3.4.1 Self-designed questionnaires 37
3.4.2 Standardised outcome measures used 38
3.4.3 Translations 38
3.4.4 The EuroQoL Five Dimension – Youth Version (EQ-5D-Y) 39
3.4.5 The Paediatric Quality of Life Inventory (PedsQL4.0) 41
3.4.6 The Paediatric Functional Independence Measure (WeeFIM) 42
3.4.7 The Faces Pain Scale-Revised (FPS-R) 43
3.5 Summary of Pilot study 44
3.6 Procedure followed for main study 45
3.6.1 Ethical approval 45
3.6.2 Research personnel and their role in the study 45
3.6.1 Recruitment and requirements from participants 45
3.6.2 Baseline measures 47
3.6.3 Repeated measures 48
3.6.4 Proxy measures 48
3.6.5 Usefulness and feasibility questionnaire 48
3.7 Data management, safety and monitoring 49
3.7.1 Data analysis 49

3.8 Ethical considerations 53
3.8.1 Autonomy 53
3.8.2 Confidentiality 54
3.8.3 Beneficence/non-maleficence 54
3.8.4 Justice 54
4 CHAPTER 4: RESULTS 55
4.1 Sample 55
4.1.1 Demographic details of the respondents 57
4.1.2 Health status of participants 58
4.1.3 Length of stay 60
4.2 Test-retest reliability of EQ-5D-Y dimension scores 60
4.3 Test-retest reliability of VAS scores 61
4.4 Baseline assessment of outcome measures across all institutions 61
4.4.1 Baseline EQ-5D-Y Dimension scores: 62
4.4.2 Index scores across institutions 65
4.4.3 Baseline EQ-5D-Y VAS scores 65
4.4.4 Proxy scores per EQ-5D-Y dimension 67
4.4.5 Discriminant validity of EQ-5D-Y at baseline 70
4.4.6 Comparison of VAS against the three levels of the dimensions, across institutions. 74
4.5 Baseline PedsQL measure across all institutions 74
4.5.1 PedsQL dimension scores across all institutions 76
4.5.2 PedsQL dimensions sub scores across institutions 78
4.5.3 PedsQL total scores across institutions 79
4.5.4 Summary of performance of PedsQL across the different health conditions, indicating discriminant validity 80
4.6 Baseline WeeFIM measure across all institutions 80
4.6.1 WeeFIM dimension scores across institutions 81
4.6.2 WeeFIM total score per institution 84
4.6.3 Summary of performance of WeeFIM across the different health conditions at the different institutions, indicating discriminant validity. 85
4.7 Baseline Faces Pain Scale measure across all institutions 86
4.8 Convergent validity 87
4.8.1 Convergent validity of the EQ-5D-Y Mobility and PedsQL Health and Activities dimensions across institutions 87
4.8.2 Convergent validity of the EQ-5D-Y UA and PedsQL total 88

4.8.3 Convergent validity of the EQ-5D-Y P/D dimension with PedsQL "I hurt" item across institutions 88
4.8.4 Convergent validity of the EQ-5D-Y WSU dimension with PedsQL Feelings dimension, at each institution 89
4.8.5 Convergent validity of the EQ-5D-Y Mobility and WeeFIM Mobility dimensions, at each institution. 89
4.8.6 Convergent validity of the EQ-5D-Y LAM and WeeFIM Self-care dimension, at each institution 90
4.8.7 Convergent validity of the EQ-5D-Y P/D dimension with Faces Pain Scale, at each of the institutions 90
4.8.8 Summary of Convergent validity between similar dimensions for all outcome measures, across institutions 90
4.8.9 Convergent validity between the sets of total scores 91
4.9 Repeated measures over the study period 92
4.10 The EQ-5D-Y scores over time, across institutions 94
4.10.1 Stability of dimension scores over time, per institution 94
4.10.2 Correlations of EQ-5D-Y VAS scores at different time intervals 96
4.10.3 Stability of VAS over baseline, 2nd assessment and 3rd assessment 96
4.11 The PedsQL scores over time, across institutions 97
4.11.1 Stability of PedsQL scores at different time intervals 97
4.12 Responsiveness of the different outcome measures 98
4.13 Correlations between dimension scores and VAS 100
4.14 Life events and changes in HRQoL 102
4.15 Feasibility and Usefulness of EQ-5D-Y outcome measure 103
4.15.1 Time taken to complete EQ-5D-Y measure 104
4.15.2 Therapists responses to usefulness questionnaire 104
4.16 Summary of the psychometric properties of the EQ-5D-Y, PedsQL and WeeFIM as assessed on the different health profiles of the children. 106
4.17 Summary of performance of EQ-5D-Y when used on children with different health states 112
4.17.1 Performance of the EQ-5D-Y dimensions 113
4.17.2 Summary of EQ-5D-Y Index Score 113
4.17.3 Performance of the EQ-5D-Y VAS 113
4.17.4 Relationship between EQ-5D-Y proxy and self-report 114
4.17.5 Feasibility and usefulness of the EQ-5D-Y 114
5 CHAPTER 5: Discussion 115
5.1 Introduction 115

5.2 Sample 115
5.3 Responses to the EQ-5D-Y 116
5.3.1 Typically developing children at a MS school 116
5.3.2 Children with health conditions 118
5.4 Evaluation of the psychometric properties of the EQ-5D-Y 119
5.5 Reliability of the EQ-5D-Y 121
5.5.1 Measuring instrument error 121
5.5.2 Variability of the characteristic (HRQoL) 122
5.5.3 Agreement between individuals taking the measure 123
5.6 Validity of the EQ-5D-Y 124
5.6.1 Discriminant validity based on dimensions 124
5.6.1 Discriminant validity of the VAS 125
5.6.2 Convergent validity 125
5.7 Responsiveness to change in HRQoL over time 126
5.8 Response shift 127
5.9 Feasibility and Usefulness of EQ-5D-Y outcome measure 129
6 CHAPTER 6: CONCLUSION 131
6.1.1 Recommendations for practice: 132
6.1.2 Recommendations for research: 132
6.1.3 Recommendations for use of the EQ-5D-Y outcome measure in economic valuations 132
REFERENCES 133
APPENDICES 151

List of Tables

Table 1: Commonly used generic, paediatric HRQoL measures 22
Table 2: Development and psychometric properties of some paediatric HRQoL measures. 23
Table 3: Socio-economic information of the two suburbs in which the facilities are located 32
Table 4: 1-Way ANOVA: Power Calculation 35
Table 5: Power calculation based on root-mean-square error (RMSSE), as calculated above. 36
Table 6: Contextual information sheet 37
Table 7: Usefulness and feasibility questionnaire (X refers to the number of responses) 38
Table 8: Outline of procedure 46
Table 9: Data analysis 49
Table 10: Gender distribution in each institution 57
Table 11: Distribution (%) of educational grade at each institution 58
Table 12: Post-hoc Tukey test determining where the difference in educational level lay 58
Table 13: Chronic or acute health status of participants at each institution 59
Table 14: Number and type of general health conditions of participants at each facility 59
Table 15: Length of stay of participants at each institution at baseline 60
Table 16: Agreement between first and second EQ-5D-Y dimension scores for pilot study. 60
Table 17: Comparing Index Scores across institutions 65
Table 18: Comparison of ranking of VAS across institutions 67
Table 19: Proxy EQ-5D-Y dimension scores per institution 67
Table 20: Kappa values comparing Proxy EQ-5D-Y dimensions scores and child self- report dimensions scores 68
Table 21: Percentage of children and proxies reporting some or severe problems in each domain 68
Table 22: Mean rank of proxy Index Scores across institutions 69
Table 23: Comparison of ranking of proxy VAS across institutions 69
Table 24: P value and Spearman's Rho for Proxy: Child self-report for VAS 70
Table 25: Kruskal-Wallis ANOVA by Ranks for Mobility, per institution 71
Table 26: Multiple comparisons of mean ranking of Mobility problems, per institution 71
Table 27: Kruskal-Wallis ANOVA by Ranks for LAM problems, per institution 71
Table 28: Multiple comparisons of mean ranking of LAM problems, per institution 71
Table 29: Kruskal-Wallis ANOVA by Ranks for UA, per institution 72
Table 30: Multiple comparisons of mean ranking of UA, per institution 72
Table 31: Kruskal-Wallis ANOVA by Ranks for P/D, per institution 72
Table 32: Multiple comparisons of mean ranking of P/D, per institution 72
Table 33: Kruskal-Wallis ANOVA by Ranks for WSU dimension, per institution 73

Table 34: Multiple comparisons of mean ranking of WSU, per institution 73
Table 35: Comparison of VAS against the three levels of the dimensions, across institutions. 74
Table 36: Baseline PedsQL of children who “never had a problem” in the various items in each dimension 75
Table 37: Kruskal-Wallis H and p values for each PedsQL dimensions subtotal, across all groups 76
Table 38: PedsQL Activity mean ranking across institutions 77
Table 39: PedsQL Feelings mean ranking 77
Table 40: PedsQL mean ranking for schooling 78
Table 41: Comparison of ranking of the PedsQL total across institutions 80
Table 42: Baseline WeeFIM for “complete independence” of all items in each dimension 81
Table 43: Differences in WeeFIM Self-care ranking across Institutions 82
Table 44: WeeFIM Mobility ranking across institutions 83
Table 45: Differences in mean ranking of WeeFIM cognition score between the institutions 83
Table 46: Weighted percentage scores for WeeFIM self-care, mobility and cognition 84
Table 47: Differences in WeeFIM total score ranking between the institutions, excluding MS 85
Table 48: Baseline Faces Pain Scale 86
Table 49: Kruskal-Wallis ANOVA by Ranks; Faces Pain Scale across institutions 86
Table 50: Comparing similar dimensions of EQ-5D-Y, PedsQL, WeeFIM and Faces Pain Scale 87
Table 51: Comparing EQ-5D-Y Mobility and PedsQL Activities dimensions 87
Table 52: Comparing EQ-5D-Y UA dimension with PedsQL total across institutions 88
Table 53: Comparing EQ-5D-Y P/D dimension with PedsQL “I hurt” item 88
Table 54: Comparing EQ-5D-Y WSU dimension and PedsQL Feelings dimension across institutions 89
Table 55: Comparing EQ-5D-Y Mobility and WeeFIM Mobility dimension, across institutions 89
Table 56: Comparing EQ-5D-Y LAM and WeeFIM Self-care dimensions, across institutions 90
Table 57: Comparing Faces Pain Scale with the independent variable, EQ-5D-Y P/D 90
Table 58: Convergent validity between similar dimensions 91
Table 59: Correlations between the different instruments total scores, all groups 91
Table 60: Summary of significant correlations between the different instruments per institution 92
Table 61: Summary of discriminant ability of the outcome measures, between the different groups 92
Table 62: Length of time between assessment periods 93
Table 63: Number of participants with a deteriorating health condition 93
Table 64: Kappa values for agreement in EQ-5D-Y dimensions at all time intervals 94
Table 65: Correlations between baseline and second assessment Index Scores across institutions 96
Table 66: Spearman’s Rho and p value for correlations in VAS at different time intervals 96
Table 67: Friedman ANOVA showing differences between VAS over time, at all institutions 96
Table 68: Friedman ANOVA showing differences between VAS scores at MS over time 97

Table 69: Friedman ANOVA showing differences between VAS at SS over time 97

Table 70: Spearman's Rho and p value for correlations between PedsQL totals at different time intervals..... 98

Table 71: The Wilcoxon Signed rank test and effect size of each outcome measure per institution, between baseline and 2nd assessment dimension scores 99

Table 72: The Wilcoxon Signed rank test and effect size based on each outcome measure's total score 100

Table 73: Spearman Rank Order Correlations between VAS and Index Scores across institutions 101

Table 74: Spearman Rank Order Correlations between Change in VAS and Change in Index Scores across institutions 102

Table 75: Number of participants who experienced an additional life event between baseline and 2nd assessment 103

Table 76: Effect of life event on EQ-5D-Y WSU dimension 103

Table 77: Effect of life event on VAS 103

Table 78: Time taken in minutes to complete EQ-5D-Y..... 104

Table 79: Responses on usefulness of EQ-5D-Y..... 105

Table 80: Summary of the psychometric properties of the outcome measures 106

List of Figures

Figure 1: Relationships between psychometric properties (25) 15

Figure 2: Adapted from "COSMIN taxonomy of relationships of measurement properties" - Mokkink et al 2010 (25). 15

Figure 3: Assembled from "A psychometric toolbox for testing validity and reliability" – DeVon et al 2007(103) and "COSMIN taxonomy of relationships of measurement properties" - Mokkink et al 2010 (25). 17

Figure 4: Flow diagram indicating the number of participants at each stage of the study 56

Figure 5: Mean age per institution 57

Figure 6: Mean educational level across institutions 58

Figure 7: Test-retest EQ-5D-Y VAS scores for pilot study 61

Figure 8: Percentage of Mobility dimension at each institution 62

Figure 9: Percentage of Self-Care dimension (LAM) at each institution 63

Figure 10: Percentage of UA dimension at each institution 63

Figure 11: Percentage of P/D dimension at each institution 64

Figure 12: Percentage of WSU dimension at each institution 64

Figure 13: Median Index Score per institution 65

Figure 14: VAS scores categorised by Institution 66

Figure 15: Median VAS per Institution 66

Figure 16: Comparing proxy Index Scores across institutions. 69

Figure 17: Self-report VAS versus proxy VAS, per institution 70

Figure 18: PedsQL median for Activity subtotal per institution 76

Figure 19: PedsQL median for Feelings across institution 77

Figure 20: PedsQL median for socialising across the institutions 78

Figure 21: PedsQL Schooling median at each institution 78

Figure 22: PedsQL subtotal for activity, feelings, socialising and schooling at each institution 79

Figure 23: Histogram indicating distribution of PedsQL total scores 79

Figure 24: WeeFIM median Self-care score per institution, excluding MS 82

Figure 25: WeeFIM median Mobility scores per institution, excluding MS 82

Figure 26: WeeFIM Cognition median score per institution 83

Figure 27: Distribution of WeeFIM total scores across institutions, excluding MS 84

Figure 28: WeeFIM median score per institution, excluding MS, at baseline 85

Figure 29: Scatter plot showing correlations between 2nd assessment PedsQL total and baseline PedsQL total; categorised by institution 98

Figure 30: Scatterplot of VAS versus Index Score 101

Figure 31: Scatterplot of Change in VAS versus Change in Index Score 102

Glossary of terms

AI	Acute Institution
CI	Chronic Institution for children with chronic health conditions
Chronic health conditions	Includes diabetes mellitus, malnutrition and failure to thrive, HIV infections, long standing respiratory conditions, Guillian Barre Syndrome and Fanconi Syndrome
COSMIN	Consensus Based Standards for the Selection of Health Measurement Instruments
ES	Effect Size
ERES	Empirical Rule Effect Size
EQ-5D	European Quality of Life (EuroQol) Five Dimension - Adult Version
EQ-5D-Y	European Quality of Life (EuroQol) Five Dimension – Youth Version
EQ-5D-Y dimensions	Mobility (*walking about*) LAM (Looking After Myself – *washing and dressing myself*) UA (Doing usual activities - *for example, going to school, hobbies, sports, playing, doing things with family or friends)* P/D (having pain or discomfort) WSU (feeling worried, sad or unhappy)
EuroQoL Group	The EuroQol Group Association comprises a network of international, multilingual, multidisciplinary researchers
EuroQol Research Foundation	A non-profit organisation researching and developing instruments that describe and value health.
FPS	Faces Pain Scale
HUI2 and HUI3	Health Utilities Index Mark 2 and 3
HRQoL	Health Related Quality of Life
HIV/AIDS	Human immunodeficiency virus
HR-PRO	Health Related- Patient reported Outcome
ICC	Intraclass correlation coefficient
LAM	Looking After Myself

MCID	Minimally Clinical Important Difference
MS school	Main Stream school for typically developing children
NS	Not significant
P/D	Pain or Discomfort
PedsQL4.0	Paediatric Quality of Life Inventory Version 4.0 Generic Core Scales
PRO	Patient-Reported Outcome
Physical disabilities / functional disabilities / chronic disabilities	Congenital disorders resulting in limited mobility such as spina bifida, congenital muscular dystrophy, Duchenne's muscular dystrophy , osteogenesis imperfecta and cerebral palsy
QALY	Quality Adjusted Life Years
QoL	Quality of Life
QWB-SA	Quality of Well Being Self Administrated measure
RS	Response Shift is the contradictory findings which occur because of changes in people's internal standards, values, or their conceptualisation of HRQOL
SS	Special School for children with physical disabilities
Std. Dev.	Standard Deviation
SG	Standard Gamble
SF-6D	Short Form-6 Dimensions
SF-36	Short Form Health Survey – 36 items
TTO	Time Trade-Off
TD	Typically Developing
UA	Usual Activities
VAS	Visual Analogue Scale
WeeFIM	Paediatric Functional Independence Measure
WHO	World Health Organisation
WHOICF	World Health Organisation International Classification of Functioning Disability and Health
WHOQOL BREF	WHO Quality of Life-BREF, an abbreviated instrument to measure

	quality of life , with 12 items
WHOQOL-100	A longer QoL instrument with 100 items
WSU	Worried, Sad or Unhappy

1 Chapter 1: Introduction and Scope of Thesis

1.1 Introduction

Quality of life (QoL) is a concept which reflects many aspects related to the lived human experience and subjective sense of well-being. Most QoL definitions refer to how well an individual's needs within a particular context are met and how satisfied or dissatisfied the individual is in various life areas. These life areas are variously known as domains or dimensions (2)(3)(4).

Health Related Quality of Life (HRQoL) is a multi-dimensional concept, which concentrates on the impact of disease or a health condition and its treatment on a person's daily life. HRQoL outcome measures are typically questionnaires on which an individual self-reports on their experience in the various dimensions. Common dimensions in HRQoL outcome measures are physical, psychological and social (5)(6)(7). The essence of HRQoL measures is that they are Patient-Reported Outcomes (PRO) (8)(9), focussing on how a health condition impacts an individual's daily life, from their own perspective (10).

There are fewer outcome measures available for assessing HRQoL in children compared to the wide range of measures available for adults (12)(11). However, since about 2005, a growing interest in HRQoL of children has led to the development of various new outcome measures for use in the paediatric population (12). One of these outcome measures developed specifically for children between the ages of eight and twelve years is the European Quality of Life (EuroQoL) Five Dimension – Youth Version (EQ-5D-Y).

The EQ-5D-Y is a short and easily-administered questionnaire in which children self-report on the extent of their problems in five dimensions, namely "Mobility" (walking about), "Looking After Myself" (LAM) (self-care), "Doing Usual Activities" (UA) (for example going to school, hobbies, sports, playing, doing things with family and friends), "Having Pain or Discomfort" (P/D) and "Feeling Worried ,Sad or Unhappy" (WSU). Scores are given for each dimension, and another score on a graduated Visual Analogue Scale (VAS) is provided for overall health status (13). The EQ-5D-Y, is one of the few paediatric HRQoL measures which have taken the views of the children for whom it is intended into account, during the development process (13). This was achieved through cognitive interviews with the children to assess their understanding of the wording and whether they found it acceptable to use.

The scores generated by HRQoL outcome measures may be used in epidemiological and clinical research studies, for the evaluation and planning of the medical management of an individual, for assessing the performance of health care systems, and for determining resource allocation (13)(1)(11). In order to utilise the EQ-5D-Y for these purposes, this measure must be reliable and valid when used in children with different health conditions. Following the initial development of EQ-5D-Y, Ravens-Sieberer et al (2010) (1) went on to assess the feasibility, reliability and validity of the measure when used with typically developing (TD) children from the general population. It was then recommended that the measure should be applied in a longitudinal, clinical study. The current research study sets out to examine the performance of the measure by comparing the responses of healthy children and those with chronic and acute health conditions in response to this recommendation. The primary research question was: How reliable, valid and responsive to change is the EQ-5D-Y when used with children who have a health condition? Secondary questions were: Do the EQ-5D-Y dimension scores

correlate to the VAS in children with chronic health conditions? Is the EQ-5D-Y a useful and feasible instrument to use routinely within a clinical setting?

1.2 Aims and objectives

1.2.1 Aim

The aim of the study was therefore to investigate the psychometric properties of the EQ-5D-Y and, thereby, to establish whether the measure is reliable, valid, and responsive, as well as feasible to use in children with different health statuses (as grouped in different institutional settings).

1.2.1.1 Specific objectives

This study recruited a sample of children between eight and twelve years of age from four different institutional settings; [1] a Main Stream (MS) school, all of whom were typically developing (TD); [2] a Chronic Care Institution (CI) for children with a chronic health condition; [3] a Special School (SS) for children with a chronic physical disability; and [4] acutely-ill children from an Acute Institution (AI). The specific objectives of this study were:

- To determine the test-retest reliability of the EQ-5D-Y (performed in pilot study, Appendix 14).
- To compare the intra-rater reliability of EQ-5D-Y and Paediatric Quality of Life Inventory (PedsQL) when repeated over a period of time, thus establishing the stability of the measures in the groups hypothesised to have a stable HRQoL (i.e. children attending MS school with no serious health conditions and children attending a SS with stable, chronic, physical disabilities)
- To compare inter-rater reliability of the EQ-5D-Y between self-report and proxy report by determining if their description and perception of HRQoL differ (i.e. between children with a health condition and their physiotherapists and between MS children and their parents).
- To investigate construct validity by examining both discriminant and convergent validity.
    - The discriminant validity of the EQ-5D-Y was examined by comparing the HRQoL profiles of the different groups of children.
    - Convergent validity was examined by comparing dimensions of EQ-5D-Y to similar dimensions on:
        - ❖ The PedsQL, another paediatric measure of HRQoL,
        - ❖ The WeeFIM , measure of functional independence
        - ❖ The Faces Pain Scale (FPS), a paediatric measure of pain.
- To compare the responsiveness to change of the EQ-5D-Y, PedsQL, WeeFIM and FPS over time. It was hypothesised that there would be a *large* improvement in HRQoL over time in the AI group of children and *slight* improvement in CI group. No change was expected in the MS or SS groups.
- To establish whether there is concordance between reporting of problems in the five EQ-5D-Y dimensions and the Visual Analogue Scale (VAS). It was anticipated that the VAS, which is an indication of overall global HRQoL, should improve in tandem with improvement in the five EQ-5D-Y dimensions, over time, as an indication of improved health condition or function.
- To establish whether changes in HRQoL were related to life events and/or changes in management of health condition. The specific research questions related to this objective included whether family related incidents, introduction or cessation of specific treatment regimes, change in pain management or surgical intervention influenced HRQoL. It was anticipated that HRQoL would mirror changes in these variables to a certain extent.
- To assess the clinical feasibility and usefulness of using the EQ-5D-Y by determining the time taken to complete the measure, whether there were any difficulties in collecting the data and whether the data gathered has been used by the clinical physiotherapists in decision making regarding prioritisation of problems and management of the child's condition.

1.3 Rationale and significance of the study

This research study was conducted to further the development of the EQ-5D-Y outcome measure, by examining whether the measure performed equally well in TD children; children with a chronic disability; with a chronic health condition without a physical disability and acutely-ill children.

A few studies subsequent to the original Feasibility, Reliability and Validity study of the EQ-5D-Y (1), have used the measure to compared HRQoL in children with a chronic health condition and healthy TD children. These studies found that the children with a chronic health condition reported a higher overall HRQoL on the VAS than those of MS children, whereas their dimension scores indicated greater problems (14)(15)(16). This finding raised the question as to whether this might have been an example of the occurrence of RS. If so, RS may have an effect on the psychometric properties of the HRQoL measure and could influence clinical management and decision making.

The ability of the EQ-5D-Y to detect a response to change in HRQoL over a period of time, has not yet been fully investigated. This study will examine which group of children demonstrated responsiveness on the EQ-5D-Y.

Even though the EuroQol website states that the EQ-5D-Y VAS, overall global HRQoL, can be used as a quantitative measure of health outcome as judged by individual respondents, there is little data concerning the psychometric properties of the EQ-5D-Y when applied to children with health conditions or disabilities.

If the EQ-5D-Y is found to be useful by the physiotherapists administering the EQ-5D-Y, in targeting their management to better address the problems identified by the child, routine inclusion of the EQ-5D-Y in assessments at each research facility could be implemented. In this case, under-graduate physiotherapy students should also be taught to use the measure as an outcome measure in paediatrics.

1.4 Outline of the study

In preparation for this dissertation, a comprehensive review of the literature was conducted. The primary themes explored included QoL and HRQoL, especially within the paediatric population, as well as the use and psychometric properties of some commonly-used paediatric HRQoL outcome measures. This is presented in Chapter 2. The methodology of the study, which followed a longitudinal, descriptive, analytical design, is presented in Chapter 3. The data were subjected to various statistical analyses depending on the particular research question being addressed and the results are reported in Chapter 4 followed by a discussion of the results for each research question in Chapter 5. Finally, a conclusion section, including recommendations for practice and future research, completes the dissertation (Chapter 6).

2 CHAPTER 2: LITERATURE REVEIW

2.1 Introduction

This chapter describes a detailed review of related literature for the various objective of the study. Data were sourced from the following online databases: PubMed, CINAHL, EBSCO and Google Scholar. Key words included: quality of life, health related quality of life, paediatric HRQoL measures, EQ-5D-Y, PedsQL, WeeFIM, psychometric properties, reliability, validity and responsiveness. Only full-text

journal articles and web articles published in the English language were considered. A time limit for articles was not set as research in the area of HRQoL increased in interest in the 1990's. However emphasis was placed on paediatric HRQoL literature published between 2005 and 2015, as this was the period of prolific research on the subject.

The review will define QoL and, particularly, HRQoL. A summary of the reasons for assessing HRQoL will be presented. This will be followed by a description of the different types of paediatric HRQoL measures, as well as a discussion concerning their use. The general psychometric properties pertaining to HRQoL measures will also be described.

2.2 Quality of Life

To understand the concept of HRQoL, it is necessary to first understand what is meant by QoL, how it is measured and factors influencing the reporting on the concept. The concept of QoL is widely used in the literature, but it is well documented that agreement on its definition is lacking (3)(17)(18)(19)(20). The most commonly-used definition is that of the World Health Organization (WHO) which defines QoL 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 goals, expectations, standards and concerns*" (21) (page e299). Within this definition the term "*individual's perception*" implies that it should be a self-report, determined by the individual (whether adult or child), concerning their own experience of their "*position in life*" or particular set of circumstances in which they live. The use of "*in the context of the culture*" suggests that the person's ethnic customs and traditions may have an influence on their perceptions and "*value systems*". This further implies that their moral standards and beliefs may influence their perception of QoL. All of these personal traits will shape the individual's "*goals*" or ambitions, and "*expectations*" or outlook on life. A person's perception of QoL is also influenced by their "*standards*" or values and "*concerns*" or anxieties regarding the life situation in which they find themselves. These many factors may influence an individual's satisfaction, well-being or happiness with their situation in life and all need to be taken into account when assessing QoL.

There are many other varied definitions describing QoL as it would not be possible for this complex, abstract concept which measures objective needs and subjective satisfaction in the many contexts in which it is used, to have a single definition (17)(22)(23). The multiple definitions of QoL depend on which vocation/profession is defining it, whether community, group or individual well-being is being examined and which indicators of QoL are being are assessed(24)(2)(8). For example, the psychology profession would define QoL with emphasis on psychological and emotional well-being, while the economic or environmental vocations might place more emphasis on material well-being or work related satisfaction. The medical profession would place more emphasis on health conditions and how disease and its management affect well-being.

There is no single theoretical model of QoL as it is multi-faceted, covering numerous life areas and whether an individual's requirements and aspirations in these areas are met. The different dimensions examined in QoL outcome measures vary in significance depending on what indicators or items of a life situation are being assessed. Measurable, objective dimensions may include indicators of socio-economic status, educational and literacy level, physical and/or mental health and life expectancy and the extent to which an individual's needs in these areas are met (2)(4). Subjective dimensions include indicators of personal values and beliefs, philosophies, traditions, politics and family and social relationships and require self-report on how satisfied, happy or fulfilled the individual is in these areas. The objective and subjective domains are interlinked; for example, an individual's socio-economic status may enhance or limit their satisfaction in social relationships (2)(3). QoL may

also be affected by the environment; for example, a disabling or enabling environment plays a role in the QoI of a person with a disability. However, their dissatisfaction or satisfaction could change as their expectations change or if they are given the resources to adapt to the environment to meet their needs (2)(4).

There is some concern among researchers as to whether or not QoL can in fact be defined and assessed, because little is known about how individuals reach a qualitative judgement regarding satisfaction or well-being. Research has shown that cognitive and linguistic processes are used by individuals in making this qualitative judgement (17). Individuals are generally required to respond to abstract concepts such as "to what extent is your problem?" or "how satisfied are you?" and, therefore, need to have a good understanding of the expressive language being used. Figurative language is also sometimes used in QoL items, such as "I feel sad or blue" or "it is hard to walk more than one block". Not all individuals may be familiar with these expressions or language use and they may respond differently, which could affect the quantitative analysis of the qualitative data (17).

As the use of QoI outcome measures increases, continued development of the theoretical models is needed, ensuring that they are cross-culturally valid. Different cultures may have different interpretations of the same word or even have different opinions of what constitutes satisfaction and well-being, which could affect the measures psychometric properties (25)(26).

2.3 Health Related Quality of Life

In the area of health, the impact of disease or treatment of disease on QoL is relevant and a specific aspect of QoL, termed Health Related Quality of Life (HRQoL) has been identified. HRQoL is a multi-dimensional concept, which focusses on the physical, psychological and social impact of a health condition on a person's daily life socially and it is ever-changing (5) (6) (7). The physical dimension may include items such as mobility, activities of daily living, energy levels and pain. The psychological dimension may include positive and negative feelings, self-esteem, memory, concentration and anxiety. Relationships with family, friends and within a community may constitute the social dimension (25). Other indicators of QoL, such as socio-economic status, the environment in which the individual functions and access to education may also influence perceived satisfaction of the individual's HRQoL (3). HRQoL refers to an individual's well-being in the above dimensions, despite certain limitations imposed on them by their health condition or disability. At this point, it is necessary to define "health" in order to understand what constitutes "health" and how it relates to QoL.

The current WHO definition of health was first described in 1948 and has never been changed. It describes health as, "*a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity*" (27) (page 1). With regard to HRQoL, this definition has certain limitations which were discussed by Huber et al (2011) in a paper following a conference of international health experts in the Netherlands (28). The main limitation was in the absoluteness of the words "complete well-being", which may not be applicable as no-one is completely healthy all of the time. With increasingly advanced screening for diseases, some sub-clinical health conditions may now be identified before symptoms negatively affect a person's QoL. However, in these cases, the term "complete wellness" would also not apply. Disease patterns have changed since 1948 with improved healthcare. Populations are living longer, often with chronic health conditions which are well managed with improved pharmaceutical drugs and have minimal negative impact on QoL. However, according to the WHO definition, people living with a chronic disease or disability would be defined as ill. This definition of health does not take into account the ability of an individual to adjust to life's changing physical, emotional and social challenges and still live a fulfilling life (28)(4).

Huber et al (2011) proposed that the definition of "health" be changed to include, "*the ability of an individual to adapt and self-manage in the face of social, physical and emotional challenges*"(28) (page 1). In terms of their perceived HRQoL, the ability of an individual with a health condition to change their own methods of managing daily life and/or their environment in order to become more independent is seen as more important than their complete recovery from the condition (28).

Another limitation of the WHO definition of health is the phrase "*A state of complete physical, mental and social well-being and not merely the absence of disease or infirmity*" (27) (page 1). This implies that children born with a health condition or disability or develop the condition very early in life and therefore do not know any other way of life, would also be defined as ill, which is not how they perceive themselves. As most HRQoL assessments are based on the WHO definition of health, children born with a disability would have a different frame of reference to the concept of health (29), compared to TD school children, who have never experienced living with a disability or chronic health condition. This could affect the validity of the measures and could lead to under-estimation of their QoL (30).

Health care professionals, researchers and policy makers are becoming more interested in the impact of disease or treatment of disease on perceived QoL of patients and their families. HRQoL measures provide health care professionals with a self-report, completed by the patient, concerning the patient's perspectives on how disease or the treatment of disease positively or negatively impacts their life. As such, HRQoL assessments may be used in conjunction with objective outcome measures which only monitor clinical changes, to holistically assess a patient. Self-reporting on HRQ0L may also reveal subjectively experienced problems not apparent to an outside observer (31)(11). Assessing the patient's subjective experience of how the health condition affects their perceived QoL can assist clinical decision making by health care professionals. This may also be useful when examining the effectiveness of clinical trials in research and can assist health economists in making resource allocation decisions. However, before using HRQoL outcome measures for these purposes, it is important to verify that they have robust psychometric properties of validity, reliability and responsiveness (33).

2.3.1 HRQoL in children

The measurement of HRQoL in children between eight and twelve years, is the focus of this thesis and will depend on each child's experiences and perceptions and on how important the life dimension being assessed is to the individual child. Two children with the same health condition may report a different HRQoL but the emphasis should always be on self-reporting from the individual's perspective (32). Following on from the key components in the WHO definition of HRQoL, the dimensions assessed in the HRQoL self-reported outcome measures are typically physical, mental, social and environmental (2). However, the items in each dimension vary depending on which HRQoL questionnaire is used.

The physical dimension includes functional items which are assessed as the ability to perform activities required for daily living (33). Function refers to the execution of a task, but does not imply that everyone performs the task in the same manner. Three separate studies investigating HRQoL and function in youths and young adults with spina bifida found that limitations in physical functioning did not negatively affect self-perceived HRQoL or limit the range of life experiences lived by the participants (34) (35) (36). A study determining associations between function and well-being in ambulatory children with cerebral palsy found that functional measures were good at predicting

functional well-being in HRQoL, but poor at predicting overall HRQoL, as there were no direct associations between the two (37). Another study investigating perceived HRQoL in children found that children with a disability and limited functional ability, did not perceive their HRQoL to be any worse than healthy, TD children (16). Therefore, there is a need to further investigate the reliability and validity of HRQoL measures in children with a chronic disability or health condition and whether these children who would be expected to report significant problems on individual dimension levels, report a high overall HRQoL or not. If they do report a high overall HRQoL despite problems on individual dimensions, it could affect the measurement properties, as it is generally expected that the level of problems in the dimensions would be reflected in the evaluation of overall HRQoL (38)(39).

In contrast, three studies examining the impact of pain on function and HRQoL, found significant relationships between results on a pain measure, a functional measure and HRQoL (40)(41)(42). It would seem that while functional limitations in children with a disability might not negatively affect their overall HRQoL, pain is more likely to affect function and overall HRQoL.

2.4 Reasons for assessing HRQoL

Self-reported HRQoL has become increasingly recognised as an important supplementary measurement in assisting health professionals in clinical practice (11)(43)(44), in empowering patients (45)(46)(47), in research and clinical trials (48)(49)(31) and in analysing cost effectiveness of health programs (50)(51)(52).

2.4.1 Clinical practice and patient empowerment

HRQoL assessments can be useful in clinical practice to enable health professionals to improve both the 'quality of care' and 'health care provider-patient' communication, especially in the paediatric population (53)(11).

Using appropriate age related HRQoL measures to obtain information on how the health condition affects the child's life, from the child's perspective and using this information when planning an intervention, improves the effectiveness and sustainability of the intervention. It ensures that the patient's priorities are met which may be surprisingly different from that of the health care professional. This shared decision making has been shown to improve compliance in the treatment process (47). Studies by Wang et al 2011 (45) and Izard et al 2014 (54) demonstrated how health care providers could use HRQoL measures to empower their patients. The authors emphasised that discussing how different health care options could influence their HRQoL and by allowing the patient to be involved in choosing the option best suited for their perceived needs could empower them to achieve satisfaction in their life.

Routine use of subjective self-report measures on health could ensure that health care providers do not overlook any functional, psychological or social problems patients might be experiencing, but may be hesitant to disclose (43). Some patients might be unwilling to initiate discussion of these issues, but by completing a HRQoL questionnaire the issues become apparent to the health care provider and communication with the patient improves (11). Using HRQoL measures to guide and encourage discussions could also result in improving patients' understanding their health condition as well as their sense of control in the management of their condition (54).

HRQoL may be measured in both the chronically- and acutely-ill patients. Even though the primary outcome for acutely-ill individuals is survival, followed later by restoration to their previous health state, self-reported measures may be administered to assess sudden and relatively short lived changes in HRQoL while in hospital. Acutely-ill individuals often experience initial deterioration in some or even

all the dimensions of HRQoL. This reduced HRQoL may be temporary and generally improve with treatment of the acute condition and pain management, or may persist after discharge from the acute care setting (55) (56). Frequent assessment of pain in the acutely-ill patient is necessary in order to effectively manage it. One method of assessing pain may be by self-reporting on a pain scale (57). The effect of an acute illness on the long term HRQoL could also be assessed (55)(58).

Patient self-reporting may be used, in conjunction with objective clinical findings, to monitor the efficacy of interventions with regard to an improved or worsened health condition (7)(11)(44). Wilson and Cleary (1995) (59) developed a conceptual model of HRQoL linking HRQoL measures with clinical measures, such as biological and physiological function. They emphasised the importance in clinicians linking clinical findings to reported QoL, in order to plan effective treatment strategies.

2.4.2 Research and clinical trails

Researchers are interested not only in the clinical outcomes of new interventions, but also in how the intervention affects the patient's HRQoL. Direct reporting by the patient on their subjective experiences of the impact a health condition has on their everyday life adds to the holistic information researchers need when evaluating new technologies (31)(7)(46). A new drug may prolong the life of a patient, but is also important to investigate whether it has improved the patient's HRQoL because prolonged and improved HRQoL may be a more desirable outcome (48). When evaluating new technologies, objective measures may indicate improvement in body structures, but may not necessarily indicate improvements in daily functioning or perceived HRQoL. These concepts could be evaluated directly with HRQoL measures (7).

2.4.3 Evaluating cost effectiveness and use of Quality Adjusted Life Years (QALYs)

Both the cost and benefit of health care interventions need to be evaluated when economists decide on resource allocation (60)(52) and HRQoL assessments could be used to guide this analysis. It has been recognized that basing resource allocation on objective clinical outcomes alone may be misleading (50) and that quality of life and length of life are both important when determining cost value of any program. The QALY, 'Quality Adjusted Life Years' was developed as a unit indicating the effectiveness of a health care program in terms of length and quality of life (50)(61). The QALY is a single index obtained by combining the length of time spent in a particular health state and the HRQoL weighting or utility score given to that health state (3). Some HRQoL measures include standardised quantitative scales or quality of life index scores which were developed using different methods of attaching preferences or utility scores to different health states. Measures which produce a single index score (utility preference score) on a 0 to 1 or +1 to -1 scale (which is needed to calculate 'Quality Adjusted Life Years' – QALY) are the most useful, where 1=full health and 0=death. Negative scores mean a health state worse than death. The use of QALYs allows for the comparison across many health states (60) (51)(62) of cost effectiveness of interventions, in terms of length of life and HRQoL.

2.4.4 Health utility measurements

The different methods used to produce a health utility preference score (index score) or HRQoL weighting are described by Whitehead and Ali (2010) (51). There are direct and indirect methods, the latter being more appropriate as it is based on pre-scored, generic, preference-based measures. Of the direct methods the Visual Analogue Scale (VAS) is the simplest. A single vertical scale, the top and bottom of which indicates the 'best health" and 'worst health', respectively, is used by the individual to mark where on the scale a particular health state would be.

The Time Trade-Off (TTO) method asks individuals to choose between living for 10 years in an impaired health state and living for fewer years in full health. They are then asked to determine how

much time they would be prepared to lose to avoid living in a diminished health state. At this point the HRQoL can be weighted e.g. if the cut-off point is 6 years, then the weighting is 0.6 (6 divided by 10).

The Standard Gamble (SG) method asks the individual to choose between the certainty of remaining in a particular health state or taking a gamble of either being in full health or a chance of dying. The probability of death is altered until the individual is indifferent between certainty and the gamble (51).

The data generated from some HRQoL measures, such as the EQ-5D-Y are ordinal and for research purposes it is useful to use value indexes to develop a summary measure against which to compare with other measures which do sum dimension scores (63).

The adult EuroQol 5 Dimensions (EQ-5D), the Short Form-6 Dimensions (SF-6D), and the Health Utilities Index Mark 2 and 3(HUI2 and HUI3) are examples of generic HRQoL measures providing a single index score (64)(61). The EQ-5D used the TTO valuation method and the SF-6D and the HUI used the SG method (51).

Until recently, the HUI 2 was one of the few HRQoL measures that produced preference weightings for the different dimensions, with which to summarise a child's HRQoL, on a QALY scale (61). One USA health valuation study used SG to produce a summary Paediatric Asthma Health Outcome Measure score on a QALY scale (65), and another more recent study by Craig et al (2015) (66) used adult preferences for child health states on the EQ-5D-Y, to produce a summary score on a QALY scale (66). The QALY values produced by Craig et al were based on the values of 4155 adult respondents who participated in an exercise choosing between losses in HRQoL and/or life span, in children with a health condition. Even though the values developed by Craig et al were used in this study to allow for the calculation of a composite score, the Index Score, it should be noted that they have not yet been formally adopted by the EuroQoL Group. In addition, as the scores are said to be QALY values and not weights to be used in QALY calculations, it would be difficult to use these in economic evaluation in which a weight should be multiplied by the time in which the condition is spent.

A question arises as to whose valuations should be utilised, child or parent proxy. Adults are ultimately responsible for the child's health and have to make the resource allocation decisions in paediatric health care so their valuations would be valid. The cognitive ability of the child to complete the valuation task might be a challenge and therefore the task would probably then fall back on the parent proxy, which in essence would be adult valuation (67).

There are many reliable and validated HRQoL instruments varying in health and disease aspects, age ranges and population groups, and in length and time needed to complete. In research, an important consideration when choosing an appropriate HRQoL measure is the study design and population being assessed, as different measures may significantly affect the statistical analysis of HRQoL (68).

2.5 Generic or specific HRQoL measures

HRQoL assessments may be generic or disease specific. Generic measures tend to assess basic function required for physical, emotional and social well-being (69)(20). They usually cover a range of dimensions, including function, functional incapacity and emotional distress that apply to most health conditions (70). Generic HRQoL measures enable comparison across different health conditions and with healthy populations, by providing a health profile of different aspects of QoL, in one outcome measure. This is achieved by a scoring system, which may generate a score for each dimension or even a total score, commonly known as an index. They are normally used to assess general health status and are able to depict changes in overall health status after treatment, meaning they are responsive to change over time (20)(71)(72)(20).

Another benefit of generic, preference- based measures, of HRQoL is that they can be used as a utility measure, which is used to analyse the cost effectiveness of an intervention related to the number of Quality Adjusted Life Years (QALY) gained. A QALY is generated by combining a description of the health state with survival. Patients' preferences for an intervention and its outcome are determined by measuring QoL as a single number somewhere between 0.0 (being death) and 1.0 (being full health). This can be achieved by one of two methods. The first being when a patient is asked a number of questions about their health and function and then classifying them into a particular category based on their responses. Each category has a QoL number between 0.0 and 1.0 assigned to it, which was determined by previous group's ratings for that health state. The other approach is to ask patients to give a single rating to all aspects of their QoL. This can be achieved by using the Standard Gamble method, which requires patients to choose between remaining in their present health state or to take a gamble between dying immediately or being in full health for the remainder of their lives. The probabilities between immediate death and full health are varied until the patient chooses a particular QoL state. A simpler method of this is used in the Time Trade Off method, which requires patients decide on how many years in their present health state they would be willing to trade for a shorter life span in full health (70) (72).

A limitation of generic measures is that they do not provide information needed to detect changes in specific disease related symptoms (24).

Therefore in contrast to generic measures, the dimensions in disease specific measures of HRQoL are specific to one particular health condition and more sensitive or responsive to changes in symptoms for that particular condition, but are unable to demonstrate comparisons between different diseases or comparison with the healthy population (11) (24)(73) (71). Disease specific measures are also not suitable if the patient has more than one health condition (32).

Generic measures were preferable for this study as they enabled comparisons across different health conditions and have been shown to capture responsiveness to treatment over time.

2.6 Paediatric HRQoL (p-HRQoL) measures

The age of the respondent also informs the choice of outcome measure and there are several measures specific to children of different ages. Child health interventions need to be assessed, the same as adult health interventions and it is useful to know more about the child's HRQoL and the impact of disease from their own perspective. This has been found to differ from the adult's perspective of the child's HRQoL in some cases (74)(75)(76)(76). As the essence of HRQoL measures is to assess the impact of a health state on the individual's QoL from their own perspective, measures appropriate for the child's developmental age are essential. Research has shown that paediatric HRQoL measures have rendered results, especially in the psychological dimension, that health service providers might not otherwise have been aware of (77)(11)(76). There has been a general under identification of psychological problems experienced by children. The need to include how this affects children's daily life has been emphasized in a study in children with diabetes mellitus type 1(78).

Different frameworks than those used for adults need to be considered when assessing children's HRQoL. It is important to consider the context in which the child functions, as well as developmental changes that take place as the child matures (77)(79)(24). Children function within many different settings, such as within the family, at school, with friends and within the community and each of these environments will have an impact on the child's HRQoL. The child is best able to report on the impact

of the health state within a school context and with friends, as this is often not directly observed by a parent (77).

2.6.1 Essential elements of p-HRQoL measures

Paediatric HRQoL measures need to take the cognitive development of a child into account (74)(79)(29). Children's understanding of health and their perceptions of how the health condition impacts their lives will change as they mature, as will their priorities on what is important to their HRQoL (80).

Language comprehension and reading level of the targeted age group is another consideration (77). It has been found that children as young as five years can reliably and validly self-report on their HRQoL with age appropriate measures (81)(77)(82)(83). Studies conducted on eight year old children have found that they are reliably able to respond to a Likert Scale (77)(82)(84); have a reliable recall period of up to four weeks (77); and, as they have a longer attention span than younger children, are able to complete longer, more complex measures (77). Eiser & Morse (2001) (32) and Rajmil et al (2004) (20) reviewed the content of some generic HRQol measures for children. They found that items in the physical dimension of the various instruments were age related, including [a] physical activities requiring maximal effort, such as sport; [b] participation in everyday activities, such as walking; and, [c] physical symptoms, such pain and tiredness. Mental or psychological dimensions included [a] negative feelings such as feeling worried, sad, difficulty with sleeping or being teased; [b] positive feelings such as happiness, ease of making friends and being able to do the same things as other children; and, [c] cognitive abilities where assessed mostly by asking about difficulty with paying attention in the classroom and coping with schoolwork. The items in the social dimension of the various instruments were very varied, but most included [a] the ability to make friends; [b] school functioning; and, [c] family involvement (32)(20).

2.6.2 Value of p-HRQoL measures

The child's experiences within these contexts will have a bearing on their psychological and social development as they mature into adulthood. Likewise, as the child with a chronic health condition develops into adolescence and then adulthood, generic longitudinal monitoring of HRQoL may provide researchers with a better understanding of the changes, the influences, and adaptations that occur over time to form the child's concept of a good or deteriorating HRQoL. This might be more useful than objective changes in specific indicators of the health condition (23)(11)(79). Health care providers and clinicians would also benefit from tracking the changes in HRQoL in children with chronic conditions in order to focus and adapt their management appropriately as the maturing child's priorities change (11) (85). Physical, cognitive and emotional growth will also be taking place in children with a chronic health condition, which could affect their perceptions of the impact of their health condition on their HRQoL and measures need to take this into consideration (86).

2.6.3 The EQ-5D-Y

The EQ-5D-Y measure has specifically been developed by the EuroQol Group, keeping in mind all the considerations mentioned above (82) (20). Wille et al (2010) (13) described the development of the EQ-5D-Y by an international task team from seven different countries. The adult version, EQ-5D, was modified in language and comprehensibility to be appropriately in line with an eight to twelve year old child's cognitive level (81)(20). This modified version was translated into Swedish, Spanish, Italian and German. The EQ-5D and the EQ-5D-Y were then randomly assigned to just under 3000 children and adolescents from German, Spain and South African children. The children completed either the adult or youth version and it was found that the EQ-5D-Y was generally easier to complete, with fewer

missing values. This was followed by cognitive interviews with the children to determine their understanding of the measure. Following further minor revisions, the English language version was adopted as the source version.

2.6.4 Challenges in p-HRQoL

Self-reporting by children has however raised a number of issues regarding the reliability and validity of the child's reporting (12) and, there is a need to investigate this further, particularly in children with a health condition, as most research has been conducted on typically developing children. Some of these challenges are:

2.6.4.1 Proxy reporting measures

In some cases, when the child is unable to report a reliable and dependable account of their own perception of HRQoL, a proxy report may be needed. This could arise if the child is too young to understand the concept of HRQoL, is cognitively impaired or too ill to reliably self-report (87)(88)(89)(90). Any individual who knows the child fairly well may provide a proxy report. This could be a parent, caregiver, school teacher or health care provider(91). However it has been found that there is often poor agreement between the proxy report and the child's self-report, in cases where it has been possible to compare the two (16)(92)(15)(77). It would seem that parents consider disease-related symptoms and physical ability or lack thereof as factors negatively affecting their child's HRQoL, whereas the child generally places less importance on this. Proxies may not always be aware of the extent of the emotional impact a health condition has on a child and they tend to under-report on this dimension (81)(26)(93). Parent proxy reporting may also be influenced by their own hopes for their child, their own perceptions of QoL and their own psychological health. This may cause a discrepancy between child and parent's perceived impact of the disease on the child's HRQoL and lends weight to the importance of HRQoL being reported by the child directly, whenever possible (16)(15)(94)(95).

As mentioned earlier, there is debate as to who should report on HRQoL for evaluating health benefits for economic evaluations and resource allocation; that is, the parent or the child, particularly in light of the self-report/proxy discrepancies (52)(15). However, the ability of a child with a disability to adapt to their condition and reconceptualise their concept of HRQoL, the phenomenon known as RS, may lead to under reporting of the effect of that health condition on HRQoL and, as a result, may lead to a smaller resource allocation(15). Therefore, there is argument against using the child's self-report for economic analysis.

2.6.4.1 Response shift (RS) / Disability paradox

There is a concern, when measuring HRQoL in individuals with a chronic disability, that ambiguous or paradoxical results may be found. This phenomenon is known as RS or the disability paradox and needs to be taken into account when measuring HRQoL in these individuals as it could affect the psychometric properties of HRQoL outcome measures thus rendering confusing results. RS refers to the change in perceived HRQoL that takes place when people with a disability adapt to a reduced level of functioning. Albrecht and Devlieger first described this in 1999 in an attempt to explain why many people with a long standing disability reported that they experienced good QoL, whereas people without a disability perceived this state as unsatisfactory (96). A study on successful aging, by Van Faber et al 2001, also discussed this phenomenon of disability paradox. The study found that elderly people viewed successful aging as an ongoing adaptation to all the changes and obstacles they encountered. Limitations in their physical and mental domains were viewed as less important to their sense of well-being and QoL, than being able to participate socially (97).

Huber et al (2011) (28) maintained that people with a health condition who were able to "adapt and self-manage" in order to maintain a balanced physiological state and still participate in work and social activities generally reported a good HRQoL. The same was true for people with a mental disorder who were able to "adapt and self-manage" their condition.

This ability of people with a chronic health condition to adapt to their situation, known as RS (98) or hedonic adaptation (10), is achieved by either changing their internal standards and values or their conception of HRQol. When analysing response shift it is useful to refer to the Spranger and Schwartz response shift model of HRQoL (38). This model refers to a "catalyst" (alpha change), usually an observable physiological change in the health condition, that initiates the need for a "recalibration" (beta change) in the individual's internal standards of measurement, resulting in "reconceptualization" (gamma change) or change in the meaning or importance of the construct being measured (38)(98)(99).

RS could be related to a sudden or gradual change in a health condition, an improving or deteriorating health condition or even the severity of the condition. Therefore, response shift is described in this model as the interaction between the "catalyst" (change in health condition), the "antecedent" (which could include demographics such as age, gender, socioeconomic group, education, personality type or spiritual belief) and the "mechanisms a person relies on to adapt" to change (such as individual coping, social and community support structures, and ability to change expectations)(38). RS could affect the reliability and validity of an outcome measure because, despite indicating a lot of problems in one or more dimensions of HRQoL, the person might still rate their overall HRQoL as good (39).

This has been noted in some HRQoL studies on children with chronic functional disabilities due to spina bifida, who, as mentioned previously, reported high levels of satisfaction with overall HRQoL despite having functional challenges in individual domains (29)(35)(34). RS may also occur in a child with a health condition who does not report a change in their sense of well-being or overall life satisfaction, despite obvious observable changes in the condition (15).

Schwartz (2010) (99) highlighted the impact of RS on the psychometric properties of HRQoL outcome measures which need to be able to accommodate changes in perceived Qol. Reliability, validity and responsiveness of HRQoL measures can be affected by the participants' differing conceptions of QoL. For example, for high internal consistency (reliability) and cross measure correlation (convergent validity), it is assumed that the participants all have the same conception of what is being measured and similar experiences to refer to. This is, however, unlikely between individuals with a long standing health condition, individuals with a sudden acute health condition and individuals with no health condition (99).

Likewise, for high inter-rater agreement (reliability), participants would need to experience similar participation levels. However, some individuals with limited physical abilities may have adjusted their internal references and rate their participation in social activities as high, whereas an individual without physical limitations would not rate this level of participation as high (99). As discussed earlier, it would seem that more emphasis is placed on functional ability in HRQoL assessments, resulting in a person with a disability scoring low in this particular area, but their self-perceived overall HRQoL may be scored higher leading to paradoxical findings if RS is not considered (39).

Therefore, some ambiguity may arise when interpreting the responsiveness of a measure. A measure might not reflect a change in HRQoL, despite evident objective changes in the health condition, as a result of well-developed coping mechanisms and this could influence the interpretation of clinical

findings (99). Likewise, a person with a disability may have poor coping mechanisms and this may affect many dimensions, without objective changes being evident (99)(39).

RS may also contribute to reported low correlations between self and proxy reports of HRQoL, as was evident in a study by Jelsma and Ramma (2010) (15).

G. Norman (2003) (100) proposed that 'RS' was not the only theoretical model to explain the ambiguity that occurred when some people assessed their HRQoL over time, but that 'The Implicit Theory of Change' was another method. According to Norman, when patients were asked to rate their health at six months and again at two years and then to rate their health retrospectively (thentest) for the earlier time period, they consistently re-rated the earlier time lower than they did initially. RS theory views the retrospective measure of health as being the more valid measure because new insight or new standards were available at the second testing. This is in contrast to 'The Implicit Theory of Change', which asks people to rate their HRQoL today and then work backwards and rate how they felt six months ago. Implicit Theory implies that accurate recall is not possible and that the retrospective rating of the initial health state was reconstructed and contextual changes may since have occurred. Therefore according to Implicit Theory the prospective rating is more valid. As a result it is unclear which theory is more accurate and RS remains a conundrum.

Further research into developing guidelines for detecting response shift and its effect on the psychometric properties of HRQoL outcome measures is being conducted by Schwartz et al (2013) (98).

2.7 Psychometric properties of HRQoL outcome measures

One approach used in developing self-reported HRQoL measures is the psychometric approach, which ensures that the scores generated are reliable (with regard to internal consistency, test-retest and intra-rater), valid (construct and criteria) and responsive to change (101)(102)(32). The diagram below (Figure 1) demonstrates the relationships between the different psychometric properties of a HRQol measure.

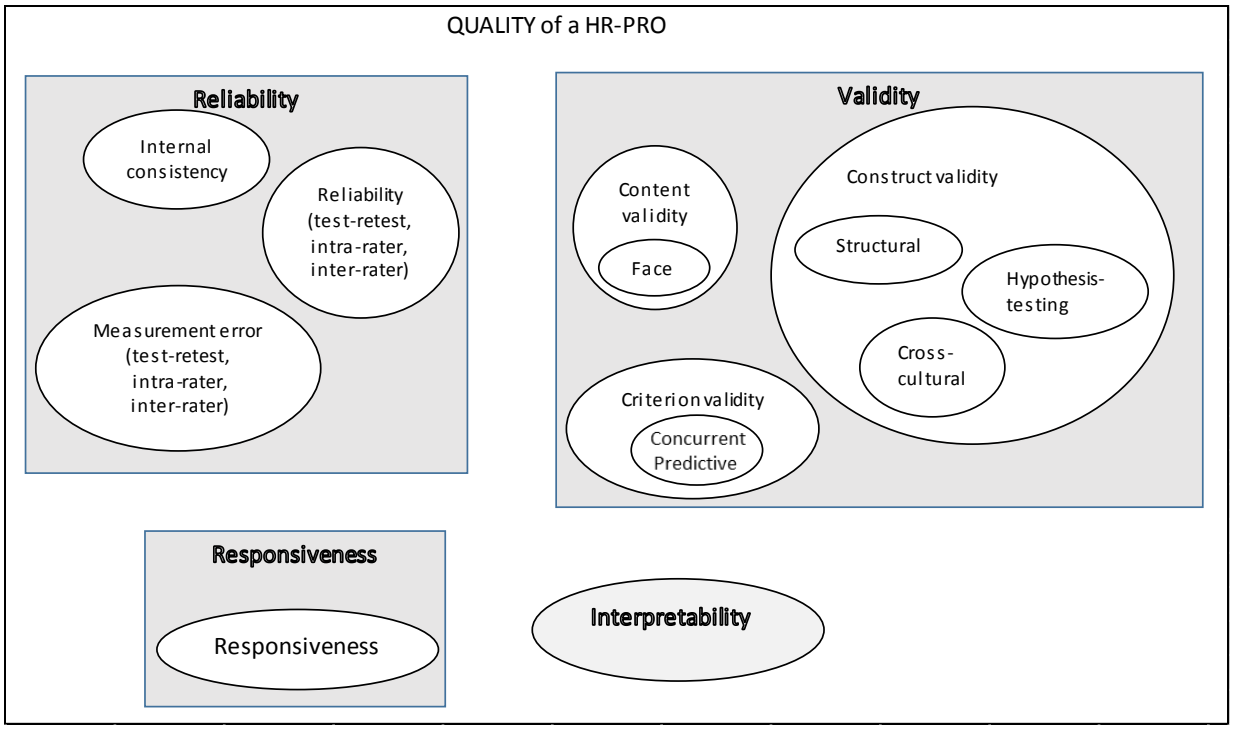


Figure 1: Relationships between psychometric properties (25)

As decisions on the management of a health condition or the effect of a treatment may be based on the scores obtained from the HRQoL measure, it is essential that the scores are precise and not biased (102). Therefore, before a HRQoL measure can be used clinically or for research, its reliability, validity and responsiveness has to be fully investigated. These investigations need to be of a high methodological quality and, as HRQoL measures essentially assess abstract constructs which are not directly measureable, the establishment of the reliability and validity of the constructs are vital.

A checklist based on the numerical criteria such as the **Co**nsensus-based **S**tandards for the selection of health **M**easurement **I**nstruments (COSMIN) is useful when designing a research study on the psychometric properties of a HRQoL measure, such as the EQ-5D-Y. The COSMIN checklist was developed to evaluate the methodological quality of studies assessing the psychometric properties of HRQol outcome measures. Fifty seven international psychology, epidemiology, statistical and clinical experts participated in a four-round Delphi study, resulting in a consensus regarding measurement criteria. This checklist that ensued provides information on design requirements, which measurement properties should be included and how these properties should be evaluated, using appropriate statistical analysis (102)(25). Therefore, COSMIN was used in the discussion of this research study. The full checklist is in Appendix 1.

A COSMIN taxonomy of relationships of HRQoL measurement properties was developed, and this encompassed three main areas, namely reliability, validity and responsiveness. Even though interpretability is not a measurement property, it was included as a characteristic (25). Each of these will be discussed as they pertain to HRQoL measures.

2.7.1 Reliability

According to COSMIN taxonomy, reliability, being the consistency with which a HRQoL measure produces the same results and is free from random error, should be examined by assessing internal consistency, test-retest reliability, intra-rater reliability and inter-rater reliability (Figure 2).

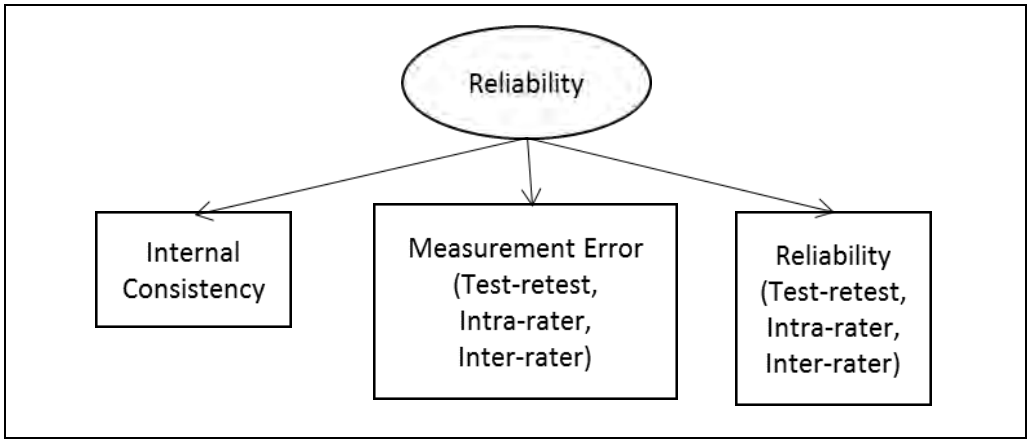


Figure 2: Adapted from "COSMIN taxonomy of relationships of measurement properties" - Mokkink et al 2010 (25).

2.7.1.1 Internal consistency

The design requirement for internal consistency refers to whether the HRQoL measure contains indicators or items which accurately describe the dimension being assessed. The statistical methods used should assess the extent to which the measure is consistent within itself. An internal consistency statistic should be assessed for each separate dimension. Relationships between the individual dimensions and overall HRQoL may also be examined for internal consistency. Cronbach's alpha

coefficient, a reliability index, can be used to calculate internal consistency reliability. Cronbach's Alpha gives a score of between 0 and 1, with 0.7 or above generally accepted as a sign of acceptable reliability (102)(103)(104)(105)(25).

In the development of the KIDSCREEN HRQoL questionnaire, the internal consistency was high. Cronbach's alpha for the dimensions of the 52-item measure ranged from 0.77 to 0.89 (106). Cronbach's alpha for The Child Health Questionnaire with 50 items (CHQ-PF50) ranged from 0.39 to 0.96 for the subscales, while the 28 item measure (CHQ-PF28) demonstrated satisfactory internal consistency ranging from 0.56 to 0.92 for the summary scales but internal consistency for individual subscales was low (107).

As most HRQoL measures are multi-trait measures, the internal consistency of the various dimensions/subscales is expected to be low because different indicators are being assessed (for example, anxiety should not necessarily be related to mobility) (21)(22)(20). In this case, low consistency would indicate a reliable measure. In particular, the EQ-5D-Y dimensions do not share the same underlying latent construct and, therefore, would have a low internal consistency.

2.7.1.2 Measurement error

Measurement error may affect the reliability of an outcome measure. Measurement error is the difference between the actual value and the value obtained through measurement. Some degree of measurement error is inevitable and is caused by either the accuracy limit of the measuring instrument (the EQ-5D-Y outcome measure), the variable nature of the characteristic being measured (HRQoL) or by errors reported by the respondents (the sample of children recruited and proxies). Whether or not the error is random (factors randomly affecting the measure or sample) or systemic error (introducing bias), it may not affect the overall result or the effect may be limited. Random error, such as an error resulting from an individual's mood, is not consistent and will vary the data but not affect the average overall performance of any group. Systemic error, a consistent error that affects all participants, will either raise or lower the results, causing bias. It is possible to apply statistical procedures to adjust for measurement error or to use multiple measures of the same construct (108)(103)(109).

All these aspects of measurement error will be assessed in the research study.

2.7.1.3 Test-retest reliability

External consistency is assessed by test-retest methods which examine the stability of the measure over time. Test-retest reliability refers to the ability of a measure to produce the same results when conducted at different times. The design requirements are that testing is performed under the same conditions each time and on the same group (102). The statistical analysis for ordinal data, as in the EQ-5D-Y dimensions, to assess percentage of agreement between test and retest responses for each dimension, should make use of Cohen's kappa coefficient. Kappa values, interpreted according to Landis and Koch guidelines (110) may be either poor, slight, fair, moderate or good. The intraclass correlation coefficient (ICC) can be calculated to determine agreement between test and retest VAS scores. An ICC of >0.7 is considered reliable and a p value of <0.05 is statistically significant.

The time between the tests should be no longer than two weeks, as children's recall becomes unreliable after this period (102)(103). In this research study, test-retest reliability was assessed on consecutive days and the reason for doing so will be discussed later.

2.7.1.4 *Intra-and inter- rater reliability*

The test-retest scores give an indication of intra-rater reliability (being the consistency a child demonstrates in scores at different time intervals) as long as the child's health is stable and the test conditions are the same (25)(111). This will be evaluated in the study.

Inter-rater reliability may be assessed for the extent of concordance between child self-report and parent proxy report, as described in some research articles (91)(92) (112). The difference in perspective between the self-report and proxy report is known as the inter-rater gap (112). Other research articles refer to the comparison of reports from different proxies, on behalf of the patient as inter-rater reliability (72)(113). This may be comparing proxy reports from both parents or parent and clinician proxy (113). The design requirement for this test is that the administrations are independent. Issues which may affect this reliability have been discussed as a discrepancy between the two groups' different perceptions of HRQoL. Parents seem to relate overall HRQoL more to problems reported in individual items, whereas children do not link the two aspects. In addition, parents place more emphasis on functional disability, while children focus on what they can achieve (16)(83)(114)(15)(92). There also seems to be better agreement between observable characteristics of health, than subjective perceptions of overall health (83). This can be calculated using ICC or Kappa coefficients (92)(115)(24). Proxy and self-report measures will be compared in the study.

The reliability of a measure is affected by a ceiling effect, which occurs when the scores of an independent variable are all grouped together at the highest level. This might be observed in HRQoL measures with limited response categories, which would prevent all variations being captured (90)(116)(1). In addition, if aspects relevant to a particular indicator of HRQoL, for a particular population, were excluded from a measure, it may not be able to discriminate between groups as a ceiling effect would occur, with all groups responding similarly (117). In the event of a ceiling effect in the three possible levels of response in measures such as the EQ-5D-Y (that is, all the responses fall into one level of "no problems"), it may not be possible to confirm reliability using kappa coefficients as the intra-rater reliability will automatically be reduced to only one level of response (111).

2.7.2 Validity (content, criterion and construct)

Validity is the extent to which a measure tests the characteristics it sets out to measure. This can be assessed by examining content, criterion-related properties of the measure and construct validity.

Error! Reference source not found. Figure 3: Assembled from "A psychometric toolbox for testing validity and reliability" – DeVon et al 2007(103) and "COSMIN taxonomy of relationships of measurement properties" - Mokkink et al 2010 (25).

2.7.2.1 *Content validity*

The general requirements needed to examine content validity are whether the various items in the measure cover all aspects or all symptoms of a particular health issue (103)(76). In order to ensure this, a full search of the literature would need to be conducted, followed by population sampling and the acquisition of expert opinions. The Patient-Reported Outcomes Measurement Information System (PROMIS) is developing item banks to ensure that all relevant items (for example, all subjective indicators of physical, mental and social well-being) for a particular construct are available when developing core sets for HRQoL measures (118)(1). To assess content validity of HRQoL measures, each item in the measure is rated by an expert in the field. These ratings also validate the appropriateness and completeness of the items as indicators of HRQoL (102)(103)(45). Focus groups of a larger sample of the population with or without a particular health issue may also be used to rate the items.

Face validity is a subjective evaluation of the measure in which the measure looks as though it measures the constructs it is supposed to. It is the weakest method of assessing validity.

2.7.2.1 Criterion validity

Criterion validity refers to the relationship between the characteristics of an outcome measure and the criteria it is measuring. It is usually assessed by comparing the outcomes of the HRQoL measure with the outcomes of a separate measure. This may make use of a number of methods:

Concurrent validity:

Concurrent validity refers to the relationship between the response to an item on one outcome measure and the corresponding criteria of an item on another measure when tested at the same time (103)(86)(67)(67)(119). Criterion validity is often used in the development of a new measure to assess the relationship between items of a well-established measure, often called the 'gold standard' or 'most accurate validated measure', against items of the new measure. The measures are tested at the same time and a consistent relationship between the two item scores deduces concurrent validity (24).

Predictive validity:

This refers to degree to which initial test scores on an outcome measure predict performance on test scores when tests are taken again in the future (103).

Convergent validity:

Convergent validity refers to the degree to which two measures that are theoretically similar are in fact, related, but they need not be tested concurrently and one measure does not need to be the 'gold standard'. Therefore convergent validity tests whether constructs of HRQoL that are expected to be similar are, in fact, related and can be examined by correlating items from one HRQoL measure with similar items on another validated HRQoL measure (1). The inter-item correlation coefficients, Spearman's rank for dimensions and Pearson's correlation coefficients for overall HRQoL would be high between measures that have convergent validity (103)(25). The convergent validity between similar items on various outcome measures will be assessed in the study.

Discriminant validity:

Discriminant validity tests that constructs which should be different are, in fact, different and do not have any relationship. For example, literature indicates that children with a disability tend to demonstrate passive behaviour, so scores on a measure assessing functional disability should not relate to scores on a measure assessing personality traits (40). The inter-item correlation coefficients would be low between measures that have discriminant validity (102)(103)(40). Discriminant validity may also be assessed by hypothesising how differently a group with acute illnesses and a group with chronic health conditions might respond to the same HRQoL measure. Testing a measure's ability to discriminate between two known groups with different health conditions may, therefore, rely on hypotheses based on expert knowledge or previous literature pertaining to the expected differences between groups (86)(40). This method will be used to determine discriminant validity of the measures used in the study.

Convergent and discriminant validity are similar but complementary concepts.

A potential problem may exist when comparing different HRQoL measures. HRQoL measures are multi-dimensional and multi-item, which are scored on an ordinal scale. For comparison purposes, total scores of dimensions, typically generated by summing item scores (which are ordinal), are used. Mathematically, ordinal data should not be summed but, rather, weightings should be attached to the

individual item scores in order to rank them. Investigating the relationship between the rank ordering of the dimensions is preferential in order to limit measurement errors and misrepresentation of results. Bias may also occur due to categorisation error if the HRQoL measure has only two to four dimensions with ordinal scores. This bias would be negligible if the sample size is large enough to detect a change in HRQoL indicators (109)(120). As the EQ-5D-Y data is ordinal, each dimension was described individually and the summed total was not used.

2.7.2.1 Construct validity

Construct validity, first defined by Cronbach and Meehl in 1995, determines whether the items in a HRQoL measure correlate with all theoretical concepts of the aspect of health being investigated.

Structural validity:

Construct validity may be assessed by looking at structural validity, which refers to whether the structure of the items in the dimensions are in fact indicators of the construct being measured (86)(25). Confirmatory factor analysis, also known as the multi-trait multi-method approach, analyses relationships among a large number of items in the HRQoL measure to derive a different smaller number of items and discover the most important factors. Items that do not relate to the important aspect of HRQoL being assessed should be removed. All items in a particular dimension should assess the same construct (102)(103)(104)(102)(121).

Hypothesis testing:

In 'hypothesis testing', a hypothesis based on a theoretical framework would indicate the direction in which the scores are expected; that is, the scores indicate consistency with the hypothesis and confirm or reject it (122).

Construct validity can also be measured using the 'known group method' in which two groups known to be different should produce significantly different scores for the construct being measured (88)(52)(81)(123).

A measure with both convergent and discriminant validity, described under criterion validity, would have excellent construct validity.

Cross cultural validity refers to whether the measure has been assessed across different cultures in different countries (124)(1).

2.7.3 Responsiveness (and ceiling/floor effect)

Responsiveness refers to the ability of a HRQoL measure to accurately detect a change in health status, however small. The design requirement is that the change in health status, which may be as a result of clinical changes, should be assessed over a described period of time. The deterioration or improvement in the health condition may also be due to the effect of treatment. Usually a hypothesis regarding the expected change in HRQoL scores are made prior to data collection (123)(125)(125)(126)(25).

There are two types of responsiveness, internal and external. Internal responsiveness is when a HRQoL measure is administered before and after an intervention which is known to produce change. Internal responsiveness may be assessed using the distributional approach, which evaluates the statistical significance of the change in the pre- and post-intervention scores (126). However, a statistically significant change may not be of clinical significance and vice versa, which limits the ability of an instrument to measure internal responsiveness (125).

It is important to determine the clinical relevance of a change in self-reported HRQoL, and what causes this the change. One method of determining this is to use the 'Minimally Clinical Important Difference Approach (MCID approach). This uses the patient's own evaluation of the smallest meaningful change in their health status which would result in a change in HRQoL score (48)(18). This is also known as the 'Anchor-Based Approach', as it is anchored to the patient's own perception of change in HRQoL (107)(127). Another method used to assess the clinical significance of a change in HRQoL score is the Effect Size (ES) approach, which relies on the use of statistical analysis and standard deviation (Std. Dev.) (48). ES is the ratio between change in mean scores and the Std. Dev. of baseline scores and is described as small (0.20), medium (0.50) or large (0.80) in the HRQoL literature (128) (81)(88)(117) (129)(49).

There is no 'gold standard method' for calculating MCID, but ES seems to be the most common in HRQoL research (12).

External responsiveness is when the changes depicted in one HRQoL measure are related to the changes depicted by another valid outcome measure; for example, when scores for mobility on a HRQoL measure are related to scores on a measure of functional ability. This is also known as the anchor-based approach (125)(130)(126)(131).

As is the case with reliability, the responsiveness of the measure could be affected by a ceiling or floor effect, which occurs when a HRQoL variable is not able to show a change as the scores are at the highest (ceiling) or lowest (floor) possible level. If the participants initially score the highest or lowest possible level, the measure would not be able to detect change in their HRQoL as there is no room for improvement or decline (1)(132)(7)(119). For the measure to be of use to a health care provider, it needs to be able to detect change due to the effect of treatment.

In this study, effect size was used to examine responsiveness of the various outcome measures, and ceiling effects were discussed.

2.7.4 Interpretability

The Cosmin requirement for interpretability was whether scores and changes in score were presented for each individual group being assessed, which was done in this study (25).

2.8 Review of some generic HRQoL Paediatric Outcome measures

2.8.1 Introduction

This section presents an alphabetical summary of some of the more commonly-used, generic measures for assessing HRQoL in children between 8 and 12 years (Table 1). In addition, this section provides details concerning the development of these measures, as well as their psychometric properties (Table 2).

Table 1: Commonly used generic, paediatric HRQoL measures

Measure		Country of origin	Age range (years)	Dimensions	Number of items	Proxy version	Time to complete (minutes)
1. Child Health and Illness Profile – Child Edition (133)	CHIP-CE	USA	6-11	1. Satisfaction (with self and health) 2. Comfort (emotional, physical and limitations) 3. Resilience 4. Risk avoidance 5. Achievement (social roles)	45-self 76-parent	yes	20
2. Child Health Questionnaire (134)	CHQ	USA	10-18 (self) 5-18 (parent)	1. Physical function 2. Pain 3. role/social-physical 4. General health 5. Perception 5. Role/social emotional behaviour 6. Mental health 7. General behaviour 8. Self-esteem 9. Parental emotional impact 10. Parental time impact 11. Family impact	87 (self) 98,50 or 28 (parent)	yes	30
3. DISABKIDS Chronic Generic Measure (86)	DCGM	Europe	4-16	1. Mental 2. Social 3. Physical	6, 12 or 37	yes	
4. EuroQol Five Dimension – Youth Version (13)	EQ-5D-Y	Multi national	8-12	1. Mobility (Walking about) 2. LAM (looking After Myself) 3. Doing my UA (Usual Activities) 4. P/D (Pain or Discomfort) 5. WSU (Worried, Sad or Unhappy) 6. VAS (Visual Analogue Scale)	5	yes	5
5. Generic Children's Quality of Life Measure (135)	GCQ	UK	6-16	Single scale	25	no	
6. Health Utilities Index (61)	HUI Mark III	Canada	2-18	1. Vision 2. Hearing 3. Speech 4. Ambulation 5. Dexterity 6. Emotion 7. Cognition 8. Pain	45	yes	5-10
7. How Are You (136)	HAY	Holland	7-13	1. Physical function 2. Cognitive function 3. Social function 4. Physical complaints 5. Happiness	80	yes	30
8. KIDSCREEN (106)		Europe	8-18	1. Physical well-being 2. Psychological well-being 3. Moods and emotions 4. Self-perception 5. Autonomy 6. Parent relation and home life 7. Social support and peers 8. School environment 9. Social acceptance (bullying) 10. Financial resources	52	yes	

9. KINDL ^{R} (137)		Germany	8-16	1. Physical well-being 2. Emotional well-being 3. Self-esteem 4. Family 5. Friends 6. School	24	yes	10
10. Paediatric Quality of Life Inventory 4.0 Generic Core Scales (123)	PedsQL 4.0	USA	5-18	1. Physical functioning 2. Emotional functioning 3. Social functioning 4. School functioning	23	yes	5-10
11. Patient Reported Outcome Measurement Information System Paediatric Global Health Measure (138)	PROMIS PGH-7	USA	5-17	General health Quality of life Physical health Mental health Feel sad Fun with friends Parents listen to ideas	7	yes	1-2
12. Quality of My Life Questionnaire (139)	QoML	Canada	8-12	Two Visual Analogue Scales and a categorical item assessing change in QoL	3	yes	<5
13. TNO-AZL Child-Quality-of –Life (140)	TACQOL	Holland	8-11	1. Physical complaints 2. Motor functioning 3. Autonomous function 4. Social function 5. Cognitive function 6. Positive emotions 7. Negative emotions	56	yes	5-10

Of the 13 p-HRQoL measures summarised, only five are specifically targeted at an age range of between eight and twelve years, while the others cover child and adolescent problems with the same measure. The EQ-5D-Y is one of the measures that is specifically developed for children within the eight to twelve year developmental range. The number of items in the different measures range from seven to 98. A questionnaire containing 98 measures would take a child far more than five minutes to complete. However, a child typically takes five minutes to complete the five dimension EQ-5D-Y.

Table 2: Development and psychometric properties of some paediatric HRQoL measures.

1. Child Health and Illness Profile – Child Edition (CHIP-CE)(133)(76)(133)

Purpose: To assess the health of 6 to 11-year-old children in epidemiological studies, and to provide a basis for measuring the impact of changes made to health services.
Method of development: Based on the literature, this measure was developed and reviewed by a panel of experts, followed by testing on focus groups of teenagers.
Scoring: Initially, items are summed. Then, a sub-score for each of the five dimensions is obtained, followed by a total score.
Reliability: The internal consistency reliability of all items was found to have an acceptable α coefficient of 0.7 or higher. The ICC for test-retest reliability of total scores was above 0.6 levels.

Validity: Construct validity in the four groups of children known to have different health statuses was assessed by determining whether the mean scores for sub-dimensions were different according to illness, age or gender. These scores were, in fact, different.
Ability to detect change: No responsiveness studies found.

2. Child Health Questionnaire (141)(107)(75)(142)(134)

Purpose: To measure health-related quality of life (HRQOL) in healthy children and adolescents (aged 5–18 years) and those with chronic or acute health conditions.
Method of development: The CHQ development was based on the structure and methodology used by the Short Form 36 Health Survey (SF-36). Traditional item scaling analysis was used (6).
Scoring: A CD on the CHQ Scoring and Interpretation Manual is required, but the individual items on each of the 14 dimensions can be summed to provide an overall health profile, or divided into two summary scores for physical and psychosocial functioning. Sub-scores and overall score can be transformed to a 0–100 scale, where 0 indicates worst possible health and 100, best health
Reliability: Cronbach's α indicated adequate internal consistency for the two summary scales, ranging from 0.69 - 0.92. Test–retest mean scores were not significantly different, demonstrating test-retest reliability.
Validity: Good construct validity was demonstrated by lower scores for children with no health condition and higher scores for those with chronic illnesses. Convergent validity with the HUI was acceptable with correlations between 0.21–0.49, for similar dimensions. Good convergent validity was also found with the PedsQL Discriminant validity was moderate to strong between children with and without a chronic health condition.
Ability to detect change: The CHQ demonstrated good sensitivity to clinical change in children whose health condition improved, moderate for those who worsened and small those whose health did not change.

3. DISABKIDS Chronic Generic Measure(107)(86)

Purpose: To assess HRQol in chronically-ill children between 4 and 16 years
Method of development: In order to identify themes, focus groups were held with children, adolescents, parents and health professionals. This was followed by item selection and pilot testing on different age groups.
Scoring:

Scoring of the main dimensions is on an ordinal scale ranging from 1 to 5. Raw scores are summed and transformed into a 0-100 scale, with 0 indicating worst health and 100, best health, respectively.
Reliability: Cronbach's α of 0.70–0.87 indicated good internal consistency on the six subscales across various health conditions. An intraclass correlation coefficient of 0.71–0.83 indicated satisfactory test–retest reliability.
Validity: Satisfactory internal consistency was found for each of the 6 subscales. Moderate convergent validity was demonstrated with validated Children's General Health Perceptions-Child Report and Paediatric Quality of Life Inventory. Discriminant validity: Significantly lower HRQOL was found in children with severe chronic diseases and those from low socio-economic backgrounds, demonstrating differences where expected.
Ability to detect change: No information

4. EuroQol Five Dimension – Youth Version (EQ-5D-Y)(13)(1)

Purpose: To enable children to self-report on their health.
Method of development: A multi-national task team of experienced HRQoL researchers, all with a paediatric background, reviewed the suitability of the adult EQ-5D dimensions for use in children. This was followed by a revision of the wording and design of the adult version to produce a version appropriate for children. The EQ-5D-Y was then piloted. The English piloted version was translated into German, Spanish, Italian and Swedish following the forward, backward, forward process. The measure was then tested on mostly healthy children, followed by cognitive interviews with respondents to determine their understanding of the measure. After discussions and adaptations, the English EQ-5D-Y was accepted as the source version. Thereafter the EQ-5D and EQ-5D-Y versions were tested for comparison of results.
Scoring: Each item has three levels of responses, 1 being 'no problem', 2 'some problems' and 3 ''lots of problems'. The VAS is a vertical graduated scale from 0 (worst possible health) to 100 (best health), giving an overall rating of both physical and psychological health.
Reliability: Fair to moderate test-retest reliability was found, with most children reporting similar levels of problems on the dimensions and satisfactory ICC (0.82) with respect to VAS. High ceiling effects in the Mobility, Self-care and Doing Usual Activity dimensions limited the ability to fully assess reliability.
Validity: Convergent validity was evident between EQ-5D-Y VAS and overall global health on KIDSCREEN, the General Health Item and Life Satisfaction Ladder, with correlation coefficients from 0.33-.0.65. Associations between EQ-5D-Y Mobility and KIDSCREEN physical well-being and PedsQL Physical Functioning scale were, however, low. Children with chronic conditions reported more problems on the dimensions than healthy children, but only children with severe problems were identified.

Ability to detect change: No information available.

5. Generic Children's Quality of Life Measure (GCQ)(135)

Purpose: To examine differences in HRQoL between chronically-ill children and healthy children
Method of development: Focus groups were held with children to determine what affected their HRQoL. The measure was drawn up and pilot-tested on healthy children.
Scoring: On a Likert type scale, scores are first given for the child character in the story that the real child feels they are most like (self-perceived score) and then scores are given for the child character they would most like to be. The discrepancy between the two scores is calculated to determine child's QoL.
Reliability: Self-perceived and QoL scores showed acceptable reliability with Cronbach's α of 0.74 and 0.78, respectively.
Validity: Discriminant validity was not indicted by different scores across the different age ranges.
Ability to detect change: No information.

6. Health Utilities Index (HUI lll) (61)(143)(91)

Purpose: HUI was designed to provide detailed descriptions of widespread health states and to provide a HRQL summary score for children with childhood cancer.
Method of development: This measure, guided by theoretical and experimental evidence, evolved from the HUI and is used as a 'gold standard' for other HRQoL measures.
Scoring: Scoring was based on ratings for health states on eight dimensions of health, using a 0-100 VAS and then assessed for an overall score of health states using Standard Gamble method.
Reliability: Assessed over the last 30 years in many clinical studies world-wide, found at http:// www.healthutilities.com
Validity: Construct validity was assessed on patients with type 2 diabetes and the measure was found to differentiate between patients with different pain levels and different emotional qualities. Concurrent validity was tested with CHQ scores for similar dimensions and measured by Spearman's rank correlation coefficients as the data was not normally distribution.

Discriminant validity was demonstrated by the measure's ability to discriminate between the presence or absence of chronic conditions.
Ability to detect change: Responsiveness was assessed in patients undergoing total hip arthroplasty and the measure was able to detect Minimally Clinical Important Difference (MICD).

7. How Are You (HAY)(20)(144)

Purpose: To examine the difference between actual achievement and expectant achievement in children's functioning.
Method of development: Development was based on self-discrepancy theory i.e. measuring the difference between how life really is, compared to how one expects it to be.
Scoring: Children first rate frequency of an activity, then ability to perform it, then the importance of the task
Reliability: For Cronbach's α, internal consistency of dimensions was reported as 0.77 – 0.93.
Validity Construct and convergent were assessed.
Ability to detect change: No information

8. KIDSCREEN(145)(124)

Purpose: A cross cultural measure of children's HRQoL across Europe.
Method of development: Developed as a shortened version of the original 52 item, to ensure its use clinically or as a screening tool for developmental problems.
Scoring: Items are scored for frequency of feelings or intensity, for each of the five dimensions, on a 5-point scale.
Reliability: Internal consistency of dimensions using Cronbach's α, ranged from 0.77 to 0.99 for different dimensions. Test-retest reliability demonstrated ICC values of 0.56 to 0.77.
Validity: Small to medium correlations (0.4) were shown between similar dimensions on PedsQL and Kindl. Discriminant validity was demonstrated by low correlations between dimensions expected to be different on KIDSCREEN and PedsQL and Kindl.

Ability to detect change: No information.

9. KINDL^{R} (137)(146)

Purpose: To measure HRQoL in healthy and ill children, aged 4 to 16 years.
Method of development: Developed from a conceptual model of the four main concepts of HRQoL, namely physical function, everyday activities, psychological well-being and social relationships. KINDL was tested on healthy children in two pilot studies.
Scoring: Scoring of each item is on a 5 point ordinal scale. Item scores are reversed before being summed to produce subscales, which can be combined to produce a total score, or transformed to values between 0 and 100. Computer scoring software needed to interpret scores.
Reliability: Satisfactory internal consistency for subscales and good reliability for total scores were reported, with Cronbach's α >0.7. Test-retest reliability correlation of scores were r=0.8.
Validity: Convergent validity with similar dimensions on the CHQ, Short Form 36 Health Survey and Life Satisfaction Questionnaire (adapted for children) was reported, with ICC 0.7. Convergent validity between the child and parent KINDL-R was also found. Discriminant validity: Low correlations between the KINDL-R and opposing dimensions of the SDQ were found.
Ability to detect change: A change in environmental living conditions demonstrated a change in scores.

10. Paediatric Quality of Life Inventory 4.0 Generic Core Scales (PedsQL) (123)(88)

Purpose: To measure HRQoL in children and adolescents from 2 to 18 years
Method of development: Data from paediatric cancer patients was collected with the aim of designing a generic HRQOL instrument to be used for all paediatric populations. This was followed by field testing on healthy children and their parents and then testing on children with juvenile rheumatic diseases.
Scoring: Items on each of the four dimensions were rated on a 5 point ordinal scale. Items can be reversed scored and linearly transformed onto a 0 -100 scale.
Reliability: The measure was tested and found to exceed the internal consistency reliability α coefficient of 0.7 for child self-report and parent proxy report scales. The individual child's total score approached or exceeded the reliability criterion of 0.9 across all ages.
Validity:

Construct validity was determined using the known group method, comparing scores between children with juvenile rheumatic diseases and healthy children. The healthy children scored significantly higher.
Ability to detect change: Medium to large effect sizes between interventions in children with juvenile rheumatic diseases were recorded.

11. Patient Reported Outcome Measurement Information System – Paediatric Global Health (PGH-7) measure (147)

Purpose: To develop a paediatric global health measure for clinical, research and quality improvement purposes.
Method of development: An item pool was developed from a computerised item bank of Patient –Reported Outcome Measures (PRO) to form the paediatric version of PROMIS, PGH-7, followed by interviews with children and parents to determine their understanding of global health and experiences of ill health. Thereafter field testing of the measure was conducted.
Scoring: Individual items are scored between 1 and 5 and an overall score can be obtained.
Reliability: Internal consistency reliability α was 0.88 for the child/youth sample and test-retest reliability coefficient was 0.73, indicting good stability.
Validity: Not assessed yet
Ability to detect change: Not assessed yet

12. Quality of My Life Questionnaire (QOML)(139)(107)

Purpose: To assess Quality of Life and Health-Related Quality of Life as two separate concepts in children.
Method of development: The measure was developed from a paediatric rheumatology sample, basing the use of the VAS on previously validated HRQoL assessments in cancer patients. Pilot testing was performed on healthy children and children with various forms of rheumatology.
Scoring: Children score their QoL on a 0-100 VAS by responding to 'Overall my life is…' and HRQoL by responding to 'Considering my health, my life is…' and answering 'Since the last time I was here, my life is…' on a five point ordinal scale ranging from much worse to much better.
Reliability: No information available.
Validity:

Convergent validity was performed by comparing scores on the QoL VAS and HRQoL VAS with scores on Childhood Health Assessment Questionnaire. Convergent validity was determined as good, as the expected relationships between scores were noted.
Ability to detect change: Responsiveness was determined by asking the child to rate two hypothetical questions on the QoL and HRQoL VAS. The questions pertained to 'How would a small improvement in health affect the scoring on the VAS?' and 'How would a small deterioration in health affect the VAS'? Numerical values were given for MCID and the measure assessed was able to depict changes in QoL and HRQoL

13. TNO-AZL Child-Quality-of –Life (TACQOL)(140)(148)

Purpose: To meet the needs for reliably and validly assessing HRQoL in children.
Method of development: Development followed focus group discussions with paediatricians, parents and developmental psychologists. The measure was tested on healthy children and children with a chronic condition, as well as their parents.
Scoring: The frequency of occurrence of a problem in each item is assessed by rating on a 4 -0 scale, with 4 being 'never' and 0 'feeling bad'. Item scores for each dimension are added.
Reliability: When comparing groups for internal consistency reliability, Cronbach's α of 0.65-0.84, indicated acceptable reliability, but this was not the case in individual scores and caution was advised as the measure did not appear stable for all the children. Test-retest ICC 0.3-0.91
Validity: The construct validity was considered good as Pearson's correlation coefficient between dimension scores were low. Concurrent validity was assessed by comparing scores on similar dimension on the KINDL and correlation coefficients were low, indicating limited concurrent validity, which was felt to be due to the difference in time frame being assessed in the questionnaire. Chronic illness had a negative effect on the measure's scores, indicating the measure's ability to discriminate between healthy and chronically ill children.
Ability to detect change: Clinical studies are being conducted.

The EQ-5D-Y is unique in that it the population it was aimed at was included in cognitive interviews during the development of the outcome measure. The youth version has been adapted from the adult version, which results in consistency in the dimension assessed across the life-span.

Most of the above measures have been investigated for reliability and validity, but responsiveness to change when a change occurred was only examined in the Child Health Questionnaire, Health Utilities Index (HUI lll), KINDL, Paediatric Quality of Life Inventory 4.0 Generic Core Scales (PedsQL) and Quality of My Life Questionnaire.

2.9 Limitations in previous assessment of psychometric properties of EQ5D-Y

To date, the psychometric properties (reliability, validity and responsiveness) of the EQ-5D-Y remain relatively unknown among acutely-ill children and children with a chronic health condition. Therefore, there was a need to test the performance of the EQ-5D-Y on participants with either an acute or chronic health condition and to compare this to healthy TD school children. In this latter population, a marked ceiling effect was observed in the original feasibility, reliability and validity study (1).

A longitudinal study was needed to assess the measure's ability to detect change in the participants over time, and to determine which factors are associated with the change. Comparing the information gained from the EQ-5D-Y with another validated outcome measure of the child's physical, psychological and social health, the PedsQL, an objective measure of function, the WeeFIM, and a paediatric P/D scale, the Faces Pain Scale, could further validate the use of this measure in routine health assessments of children. In addition, determining whether or not children with differing levels of illness and disability demonstrate RS could direct the use of the EQ-5D-Y.

2.10 Summary of literature review

HRQoL assessments focus on how disease impacts a person's daily life physically, psychologically and socially. Self-reported HRQoL has been recognised as an important measurement in research and clinical trials in order to improve clinical practice and patient management, as well as for economic evaluations aimed at improving health care services.

Self-reporting from children on their HRQoL, using measures that take the child's developmental age into account, such as the EQ-5D-Y and PedsQL, may identify unexpected health problems, especially psychological problems, which might otherwise be overlooked. Sometimes, the child's perspective of their own HRQoL has been found to differ from their care-givers' perspective. Children also function in different contexts from adults (for example, in play, at school, and with friends) and, therefore, there is a need to develop child-appropriate HRQoL measures with good psychometric properties.

Few generic HRQoL measures for children have been validated in a clinical population and even fewer have examined the ability of the measure to detect change over time i.e. responsiveness. Therefore, this study aimed to validate the EQ-5D-Y in clinical samples and to determine the responsiveness of the measure over time.

3 CHAPTER 3: METHODOLOGY

3.1 Introduction

An observational, analytical cohort study with repeated measures was used to investigate the psychometric properties of the EQ-5D-Y. The research questions called for repeated assessments of HRQoL over six months using the main outcome measure, the EQ-5D-Y and comparing it's performance to another HRQoL measure, the PedsQL; a measure of functional independence, the WeeFIM and a paediatric pain measure, FPS. Measures were taken at baseline and repeated three months later or shortly before discharge for the acutely ill children and again three month later i.e. seven months post baseline. The study design was an observational cohort study in that the same participants were followed over the research period. The analytical component of the study design refers to the testing of the psychometric properties of the EQ-5D-Y in the children with different health statuses, including whether the measure would depict different changes in the HRQoL in the groups expected to show a change, over time (138)(139).

It was hypothesised that the children from the Chronic Institution (CI) group would show a slow, but steady improvement in HRQoL with improved management of their chronic health condition, as evidenced from changes in their EQ-5D-Y dimension scores and VAS scores. This change was expected to be minimal for the Special School (SS) group, whose chronic disabilities were more stable. Those in the Acute Institution (AI) were expected to demonstrate a larger, swifter improvement with treatment and once their medication took effect. The Main Stream (MS) school children were expected to maintain a more or less static HRQoL in terms of the VAS and the various dimensions on the EQ-5D-Y, as they were typically developing (TD) with no serious health problems.

3.2 Research Settings

3.2.1 Geography and Demographics

The current study was conducted at four different facilities for children, within the same middle to low socio-economic area in the Western Cape (socio-economic information in Table 3 below). Cape Town, the largest city in the Western Cape, has a population of 3,740,0225, according to the latest population consensus carried out in 2011 (149). The four facilities were located in two Cape Town suburbs, which are closely linked geographically. In the suburb of Athlone, the population size is 45 048. People whose home language is English or Afrikaans account for the majority of this population, while 8.5% are first language isiXhosa speakers. The population of the suburb of Bonteheuwel is 52 956, of which 4.1% are home language isi-Xhosa speaking.

Table 3: Socio-economic information of the two suburbs in which the facilities are located

	Athlone	Bonteheuwel
Total population	45 048	52 956
Households	11 739	11 037
Average household size	3.84	4.8
Percentage of population 20 years and older having completed Grade 12 educational level or higher.	59	23
Percentage of population who are employed.	88	73
Percentage of population with a monthly income of R3 200 or less.	31	52
Percentage of population living in a formal dwelling.	95	87

Monthly household income:	Number	%	Number	%
No income	1 254	10.7	1 275	11.6
R 1 - R 1 600	1 239	10.6	2 217	20.1
R 1 601 - R 3 200	1 185	10.1	2 220	20.1
R 3 201 - R 6 400	1 392	11.9	2 427	22.0
R 6 401 - R 12 800	1 890	16.1	1 899	17.2
R 12 801 - R 25 600	2 337	19.9	759	6.9
R 25 601 - R 51 200	1 638	14.0	189	1.7
R 51 201 - R 102 400	606	5.2	21	0.2
R 102 401 or more	195	1.7	24	0.2
Unspecified	0	0.0	0	0.0
TOTAL	11 736	100.0	11 031	100.0

2011 Census – Cape Town Suburb Profile (December 2012). Compiled by Strategic Development Information and GIS Department, City of Cape Town, using 2011 Census data supplied by Statistics South Africa.

3.2.2 Description of facilities from which the children for the study were drawn

In South Africa, the Department of Education is responsible for centres for early childhood learning, public schools, private schools and special needs schools. The state provides the schools with a subsidy depending on the level of poverty of the neighbourhood in which the school is situated. Therefore the eight public primary schools in Athlone and the 14 in Bonteheuwel receive more money per child, than schools in more affluent suburbs. One of the primary schools for TD children in the Bonteheuwel area was chosen as a research facility.

Special Need Schools cater for children with various health conditions, providing them with classes of smaller numbers, feeding schemes and health care to ensure optimal learning occurs in these facilities. Some Special Schools follow the normal school curriculum and are also government-funded. They do their own fundraising to supplement their running costs because the children's parents only pay what they can afford.

The Athlone area includes Special Schools for the blind; for children with cerebral palsy and other physical, as well as learning problems; a Special School for children with mostly congenital disabilities, and a school attached to a home for children with chronic illnesses. Children from the latter two schools were included in this research study.

An Acute Care Hospital in close proximity to Athlone, which admits children from a wide geographical area and from mid to low income bracket families, was also chosen as a research facility.

The four research facilities were all clinical sites where physiotherapy students from the University of Cape Town carry out clinical training. From a practical point of view these facilities were chosen as the researcher was familiar with the sites from supervising students and the facilities were known to be amenable to interaction from the university.

3.2.2.1 A Main Stream (MS) primary school for typically developing (TD) children

A Main Stream (MS) public primary school in Bonteheuwel, catering for about 350 TD children, ranging from an age of five to approximately 13 years, was one of the research settings. All the children live in the surrounding area and none presented with disabilities or serious health conditions. A few relatively minor health issues such as asthma, eczema and mild learning difficulties were reported. These children's health status was not expected to change over the study period. For a complete list of the minor health conditions acquired by these learners, see Table 14 and Appendix 2.

3.2.2.2 A Special School (SS) for children with physical disabilities

A Special School (SS) for about 200 learners with special needs is located in Athlone. Children from the age of five to 16 years attend this school which caters for children with primarily physical disabilities (such as spina bifida, various muscle diseases, spinal cord lesions and cerebral palsy). Many of these children are wheelchair bound or walk with assistive devices, and require management of their condition throughout their lives (See Table 14 and Appendix 2 for full range of conditions). A nursing service is provided at the school, as well as rehabilitation therapies. The children's disabilities were expected to remain stable over the study period. At the time of this study, approximately 50/200 (25%) of the learners were between the ages of eight and 12 years. Most of them lived on the premises during school terms and followed a main stream school curriculum.

3.2.2.3 A facility for children with chronic health conditions (CI)

A long-stay, chronic care institution (CI), also in Athlone, which cares for about 145 children aged from infancy to 18 years, was also recruited for this study. The children present with a wide range of conditions including diabetes mellitus, neoplasm and neurological disorders and heart, renal and respiratory impairments. They are admitted to the facility if their families are no longer able to manage their chronic health condition at home. Fewer of these children rely on mobility aids compared to the children at the SS, but they require nursing management of their chronic health condition. Most of the health conditions are acquired. They face more emotional challenges in adjusting to their new health status, than the children born with a disability. (See Table 14 and Appendix 2 for full range of conditions.) Children are typically admitted for six months or longer, 45% for up to a year. Being a convalescent home, the majority of the children live on the premises while either recovering from a severe illness or suffer a chronic health condition that needs supervised medical care and rehabilitation. At the time of this study, approximately 50/145 (~34%) of the children were between eight and 12 years of age and cognitively able to attend school on the premises.

3.2.2.4 An acute care children's hospital (AI)

A referral hospital, situated near Athlone, offers a comprehensive range of specialist paediatric services to acutely ill children, from low to middle income areas, but covering a large geographical catchment area. At the time of this study, there were about 350 in-patients who suffered a large range of conditions in either medical or surgical wards, in intensive care, trauma or oncology wards. Most of the children at this facility were typically developing (TD) before being admitted to hospital for management of an acute condition and most were in pain. See Table 14 and Appendix 2 for a complete list of health conditions. Although the majority of the patients were younger than eight years of age, children up to the age of 16 were admitted to this acute institution (AI) for short term care of an acute health condition. The average length of stay was about one week.

At the start of this research study, no standard instrument was being used by the physiotherapists at any of the institutions to assess or monitor changes in children's HRQoL, and this aspect of child health was not being taken into account.

3.3 Study participants

The sampling frame included all children between eight and twelve years of age with a health condition at the three institutions catering for the needs of these children. Children from eight to twelve years in grades three to seven[1] at a MS school for TD children, were also included as a

[1] Originally the sample was to be drawn from children in Grades three to six, but this was amended as a result of the pilot study. See Section 3.2.2

comparator. One child of thirteen years was included in the sample as he had been assessed by a school psychologist as being of a cognitively developmental level of eleven years. The sample was one of convenience in that there was no random selection of either the facility or the individual participants.

The MS school records indicated that there were 201 children between the ages of eight and 12 years. All of these children were considered eligible for recruitment and were given an informed consent form to take home for parents to sign and were asked to sign an informed assent form themselves. Only the 105 children who returned legible, signed consent and assent forms were included in the study.

The records at the SS and the CI indicated 45 and 47 children, respectively, between the ages of eight and 12 years with a health condition attended each facility and they were all given an informed consent form to be signed by a parent, when they returned home for the weekend or the parent visited. Only 35 children at the SS and 32 children from the CI returned signed consent and assent forms and were included in the study.

At the AI, all children admitted during the six week data collection period allocated to the institution, fulfilling the inclusion criteria, were recruited two days post admission. They were included in the study if both consent and assent forms were signed.

All the physiotherapists treating children with a health condition at the various facilities, and all parents of MS children were invited to complete the EQ-5D-Y proxy measure.

3.3.1 Sample size calculation

The sample size was calculated based on the number of children with an acute or chronic health condition who were anticipated to indicate a change in overall HRQoL with repeated testing, compared to children with no health condition. There was no available literature on which to base how many points would make a meaningful difference in HRQoL when using the EQ-5D-Y on these children. Therefore the calculation was based whether a minimum of 30 children per group would be sufficient to demonstrate a change of 5 points on the EQ-5D-Y VAS, with a Standard Deviation (Std. Dev.) of 18, from baseline to repeated measure. The VAS scores and Std. Dev. were taken from a study using the measure in other South African children attending another MS school and the same SS for children with disabilities (16), as used in this study. While a larger difference in VAS, may be of more clinical relevance, it was decided to calculate the sample size needed to measure a small difference, so that even slight changes would be included.

Table 4: 1-Way ANOVA: Power Calculation

Type I Error Rate (Alpha):	0.05
Number of Groups:	4
Group Sample Size (N):	30
RMSSE:	0.929

Table 5: Power calculation based on root-mean-square error (RMSSE), as calculated above.

	Value
Number of Groups	4.000
Group Sample Size (N)	30.000
RMSSE	0.929
Non centrality Parameter (Delta)	77.678
Type I Error Rate (Alpha)	0.050
Effect Df	3.000
Error Df	116.000
Critical Value of F	2.683
Power	1.000

Therefore a minimum of 30 children per group are required to ensure the sample size is sufficiently powered to detect changes in HRQoL VAS scores as specified.

3.3.2 Inclusion and exclusion criteria

The children had to be able to respond independently to the EQ-5D-Y questionnaire, the other HRQoL self-report measure and the pain measure. There is evidence in the literature that children as young as five years can reliably and validly report on their health and this increases from seven years, so long as the measure has been specifically developed for the age of the children (82)(80). Most studies have found good reliability and validity in self-reporting on health by children of eight years and older (1)(127)(117)(125)(140). It has been recommended that children need to have attained a developmental age of eight or older in order to respond appropriately to the numerical Visual Analogue Scale (VAS) (1) (141). Children in South Africa may start formal schooling at age six and it is compulsory to start in the year in which they turn seven. Therefore, by grade three, most children will be eight years or older. It is assumed that these children, having successfully passed two years of schooling, should be of a developmental age of at least eight years because they came from a low to middle socio-economic background and were not exposed to extreme poverty or malnutrition, which may affect cognitive development (142). The upper age limit of the measure was 12 years and as most 12 year old children were in grade seven, (as was evident following the pilot study which only included children up to grade six), grade seven children were also included in the main study.

Children who were in the terminal stages of illness or were critically ill, with unstable and/or abnormally high or low vital statistics were excluded as it was likely that they and their parents might find the task too distressing.

Any child who was unable or unwilling to respond was excluded.

3.3.3 Sample for pilot testing

A sub-sample of convenience of 38 children who met the above criteria participated in the pilot study and test-retest reliability testing of the EQ-5D-Y (Appendix 14).

3.4 Instrumentation

Information was gathered using two standardized measures of HRQoL developed for children between eight and twelve years, an observational measure of functional independence, a paediatric pain measure and two self-designed questionnaires.

3.4.1 Self-designed questionnaires

3.4.1.1 Contextual information sheet

Information related to the demographics, general health status and management of the health condition was entered into a contextual information sheet for each child. This sheet was designed by the researcher in order to capture the demographic information (institution, date of birth, age, gender, educational level and diagnosis) and specific information which could affect HRQol in the participants (Table 6). The full contextual information sheet can be seen in Appendix 3.

Table 6: Contextual information sheet

ITEM	REASON
Reason for admission	To identify the reason for needing special health care or schooling and its impact on HRQoL.
Medication	In order to assess its action e.g. relieves pain or makes participant nauseous or drowsy and therefore impact on HRQoL.
Health status: Acute, chronic or typically developing	To enable differentiation between the groups and to identify a participant with a chronic health condition, who might have a superimposed acute illness
Assistive device and type	In order to identify whether a child with mobility problems is on bed rest (immobile and high impact on HRQoL) or in a wheelchair or walking with crutches (more mobile and less impact on HRQoL)
Urinary catheter or not, self-catheterised or not	Self-managed catheterisation would have less of an impact on HRQol, than an indwelling catheter
Seen a doctor (other than for routine check-up) and/or cut down on activity level in last two weeks	If so, this would indicated a deterioration in health condition, adversely affecting perceived HRQoL
Medical management	To assess the level of intervention and whether it may be expected to influence any aspect of HRQoL, presently and/or over time
Surgical management	Any recent surgical intervention would be expected to negatively influence all present aspects of HRQol, but might improve HRQoL over a longer period
Physiotherapy management	Might contribute to improved function overtime.
Occurrence of a life event in the last three months e.g. death in the family, divorce, moving home, new baby.	To monitor whether the occurrence of a life event affects HRQoL in any way
Included in life event, is whether child is a border at school during the term and only goes home during vacation time. If child is acutely ill and hospitalised, do family visit regularly	To assess whether being separated from family during the term affects HRQoL. To assess whether the lack of family visits affects HRQoL
Time child took to complete the self-report measures	To determine feasibility of using the measures routinely

The information to populate this questionnaire was obtained from medical files, nursing staff, therapists, teachers and direct questioning of children.

3.4.1.2 *Usefulness and feasibility questionnaire*

The clinical physiotherapists who administered the EQ-5D-Y to the participants or who completed the proxy version filled in a questionnaire designed by the researcher, asking them to comment on the feasibility and utility of the EQ-5D-Y as a standard outcome measure in the future (Table 7). Even though most of them only completed the proxy version, it was felt that they could comment on its potential usefulness. The full questionnaire can be seen in Appendix 4. The number of positive responses to the questions was recorded.

Table 7: Usefulness and feasibility questionnaire (X refers to the number of responses)

Number of self-reports administered?	x
Number of proxy report completed?	x
Ease of administration?	Very easy; moderately easy or difficult
Reason if not very easy?	Time constraints; child did not understand; age group not understanding
Which dimensions did children find difficult to understand? VAS?	Mobility; LAM; UA; P/D/ discomfort; WSU
Was a relationship noticed between child's response and clinical signs in any dimension?	Mobility; LAM; UA; P/D; WSU
Did outcome measure assist with planning management of child?	Yes or No
Which dimension was most useful when planning management of child?	Mobility; LAM; UA; P/D; WSU
Did the measure provide additional information on child's health status?	Yes or No
Which was the most useful dimension in providing additional information	Mobility; LAM; UA; P/D; WSU
Will you continue to use measure routinely?	Yes or No

In order to determine content validity, the self-designed questionnaires were examined by a panel consisting of three paediatric physical therapists, with knowledge in research and HRQoL. The feasibility of gathering the content was examined during the pilot study (Appendix 14)

3.4.2 Standardised outcome measures used

The primary outcome measure of HRQoL was the EQ-5D-Y. The reason for using this measure and its psychometric properties will be discussed. Justification for the selection of the other outcome measures, PedsQL, WeeFIM and Faces Pain Scale are also included, below.

3.4.3 Translations

Prior to using the two self-reported HRQoL outcome measures, EQ-5D-Y and PedsQL, translated versions were required.

As the most common languages, apart from English, spoken in the Western Cape are Afrikaans and isiXhosa, validated Afrikaans and isiXhosa translations of the EQ-5D-Y were sourced from the EuroQol Foundation, all of which had been through the rigorous translation process required by the EuroQol Foundation.

There were, however, no Afrikaans and isiXhosa versions of the PedsQL4.0 Generic Core Scale. The translation process was undertaken following the forward, backward, forward method, prior to the start of the study. For this process, two translators bilingual in Afrikaans and English and two translators bilingual in isiXhosa and English were sourced.

The English PedsQL version was translated forward into the two languages. Following the backward translation (back into English), discussions between the researcher and the translators were conducted in order to produce the best reconciled version. The reconciled Afrikaans version made use of the most appropriate wording from each translator's version, to preserve the original English meaning as closely as possible.

The reconciled isiXhosa version, agreed upon after similar discussions, combined the more contextually accurate phrases and phrases that were more commonly used by isiXhosa speaking people, from each translator's version.

A report was then sent to Dr Varni at the Mapi Research Group and, following their recommendations, further slight changes were made to the wording of both the Afrikaans and isiXhosa versions. These changes mostly concerned finding a more appropriate word for "depressed" which was used in both the Afrikaans and isiXhosa languages and was felt to be too strong for "I feel sad". This was changed to a more appropriate translation of the word "sad "in each language. "No energy" or "lack of energy" in the isiXhosa version was also felt to be too strong for "I have low energy" and this was changed to "I feel tired".

Cognitive interviews, using the newly translated measures were then conducted on five Afrikaans speaking children, between eight to twelve years of age, who all easily understood the form and could complete it independently.

However, when tested on the isiXhosa speaking children, it was discovered that even though they spoke their home language fluently, most of them could not read isiXhosa as they attended English or Afrikaans speaking schools and had learnt to read in one of these languages. Five children who attended an isiXhosa school and who could read isiXhosa were sought and the form was tested on them. They also easily understood the form and could complete it without difficulty.

The forms were proof-read and finalised, and a final report was submitted to Mapi Research Group. The translated PedsQL forms were considered appropriate for the research participants.

The full report on the PedsQL translation process can be found in Appendix 5.

3.4.4 The EuroQoL Five Dimension – Youth Version (EQ-5D-Y)

The EQ-5D-Y was the primary outcome measure of HRQoL (Appendix 6), as the research study was conducted out of a need, identified by the EuroQoL Group, following the reliability and validity study of the measure by Ravens-Sieberer et al (2010) (1). The authors suggested that the performance of the EQ-5D-Y be examined further on a clinical sample of children, as the original study used in the main, normal TD children. One of the principle aims of this research study was to assess the psychometric properties of the EQ-5D-Y (reliability, validity and responsiveness) in children with different health states.

The EQ-5D-Y assessed HRQoL by requiring children to subjectively, self-reporting on five dimensions, namely:

- Mobility (walking about)

- Looking After Myself (LAM) (self-care, washing and dressing))
- Doing Usual Activities (UA) (going to school, hobbies, sports, playing, doing things with family or friends)
- Having Pain or Discomfort (P/D)
- Feeling Worried, Sad or Unhappy (WSU) (anxiety or depression)

Each dimension, which consists of only one item, has three possible levels of response:

1 “no problems”
2 “some problems”
3 “a lot of problems”

It should be noted that the five dimensions do not share the same underlying construct (they are multi-dimensional) and the intervals between the three problems levels are not necessarily equal, so the dimensions were assessed individually. Until recently, there was no single index score for the EQ-5D-Y which summarised the level of problems reported on each dimension. Recently Craig et al (2015) (66) developed a summary EQ-5D-Y on a QALY scale which was used to determine a composite score, the Index Score.
Craig et al (66) used adults to weight losses and gains in children’s HRQoL to produce a QALY weighting to allow for comparison across different measures using a composite score, which takes into account the importance of different health states. For example, 1 year with ‘a lot’ of problems walking about, washing or dressing, and doing UA (i.e., physical problems) is worth the same as 1 year with ‘some to a lot’ of P/D, from the adults’ perspective.

The EQ-5D-Y has five dimensions (each with three levels) and measures 10 possible losses in HRQoL (e.g., ‘no’ to “some” problems walking about).

The weighting produced by Craig et al (66) describe the values of losses in child HRQoL, assuming that each lasted 1 year. For each dimension, the value of ‘no to some’ problems is less than ‘some to many’ problems. As the two largest losses are ‘some to a lot’ of P/D (4.0 QALYs, 95% CI, 3.8–4.4) and ‘a bit to very WSU (2.0 QALYs, 95% CI, 1.9–2.3), index scores of up to 10 are possible, using Craig’s QALY weightings, which are actually a decrement in QALY’s. This assumes that some states are worse than death.

In addition to the dimensions, there is a Visual Analogue Scale (VAS) on which subjective, self-rated overall health status is rated on a graduated VAS scale, with 0 indicating worst health state imaginable and 100 indicating best health imaginable (1). The measure takes five minutes to complete and refers to the child’s health on the day of completing the questionnaire.

The previously Afrikaans translated versions for both self- and proxy reports were used for all Afrikaans speaking participants. No isiXhosa speaking participants used the translated version, as the majority of these participants were unable to read isiXhosa, but were all able to read English.

The EQ-5D-Y is a copyrighted instrument. Therefore, prior written consent to use the measure was obtained from the EuroQol Executive Office. The study was registered at the EuroQol website and permission was granted to use the paper version of the EQ-5D-Y (Appendix 7).

3.4.4.1 Validity and reliability studies

As mentioned in the literature review, the original reliability and validity of the EQ-5D-Y was conducted on a multi-national group of mostly healthy, (TD) children, including South African children (1). Some subsequent studies have compared the performance of the EQ-5D-Y between TD children

and children with a health condition. In 2009, a study on German children with cystic fibrosis found the measure to be valid in depicting differences in HRQoL as this chronic disease state progressed, even though it was not compared with any other group (123). A South African study in 2010 compared the EQ-5D-Y self-report and parent proxy report in children with a disability at a SS and children at an Open School for TD children. Even though there were differences between the parent respondents and the children's self-report, the EQ-5D-Y was deemed to demonstrate inter-rater reliability in both groups. Validity between VAS and dimensions was evident with strong relationships between the two at the Open School sample, but not in the SS sample (16). Discriminative validity was examined in a Swedish study in 2011 to determining whether the EQ-5D-Y could discriminate between children with a prior known health condition, such as asthma, rhinitis, diabetes and coeliac disease and severe illness and/or handicap (which was not specified), and children with no reported health condition. The measure was found to have discriminate validity, however only 16% of the sample group had a health condition (144). An Italian study in 2011 found the measure to be reliable and valid when tested on healthy children and children with Acute Lymphoblastic Leukaemia (107). Again, the percentage of children with the acute health condition was very low (only 6% of the sample).

None of these studies has assessed the performance of the measure across chronically-ill children with or without a physical disability, acutely-ill children and TD children. There have also been no longitudinal studies to examine the measure's responsiveness to changes in HRQoL in these groups of children overtime.

3.4.4.2 EQ-5D-Y Proxy version

The EuroQol Group, the body responsible for the development of the HRQoL measures, has also developed a proxy version, which requires the adult caregiver, or anyone else who knows the child well such as the teacher or therapist, to complete the form in the same way they think the child would complete the form (16) (17). Even though proxy measures are most commonly used when the person is unable to complete the form themselves, as in cases when the child is too young to understand the concepts of HRQoL being investigated or the respondent is either physically or mentally incapacitated in some way (87), they may be useful to determine whether the proxy respondent is in agreement with the child as to what extent the health condition affects their QoL (16) (73).

3.4.5 The Paediatric Quality of Life Inventory (PedsQL4.0)

The 23 item PedsQL Generic Core Scales, another generic paediatric HRQoL measure (Appendix 8), specifically developed and validated for children between the ages of eight and twelve years (123), was used to compare results with the EQ-5D-Y.

The PedsQL Generic Core Scales consists of self-reporting on 23 items, in four dimensions:

- About My Health and Activities (Physical Functioning) – eight items
- About My Feelings (Emotional Functioning) – five items
- How I Get Along With Others (Social Functioning) – five items
- About School (School Functioning) – five items

The scoring ranges from 0-4 with:
0 being "never a problem",
1 "almost never a problem"
2 "sometimes a problem"
3 "often a problem" and

4 “almost always a problem”

The scores may be added together to produce a dimension sub-score and also a total score of all dimensions. Therefore the higher the score, the greater the problem in that dimension and the lower the HRQoL (150).

3.4.5.1 Reliability and validity study

The PedsQL was found to be reliable and valid for use as a paediatric health outcome, following a study in 2003 on children from the general population between different age ranges, including those between eight and twelve years. It was tested on a large sample of children, 86.3% of whom were healthy and 5.4% with a reported chronic health condition, such as asthma, ADHD (Attention Deficit Hyperactive Disorder) and diabetes (123).

The PedsQL is longer than the EQ-5D-Y and it takes about 15 minutes to complete. All items refer to how much of a problem was experienced in each item in the ‘past month’, as opposed to “today” in the EQ-5D-Y.

Similar to the EQ-5D-Y translated versions all Afrikaans speaking participants used the translated version and no isiXhosa speaking participants used the translated version

A User Agreement form was completed and returned to the Mapi Research Trust, thus allowing the researcher the use of the PedsQL4.0 (Appendix 10).

Rationale for using the measure:
The PedsQL was chosen as it is has similar dimensions to the EQ-5D-Y. It is longer than the EQ-5D-Y as each dimension is made up of a number of items, creating a sub-score per dimension for comparison with the similar dimension on the EQ-5D-Y. A total score is also generated by summing the sub-scores giving an overall HRQoL score, which could be compared to the EQ-5D-Y VAS. . It has been used reliably in conjunction with the EQ-5D-Y in other studies to assess HRQoL in children (151)(89)(1)(123).

3.4.6 The Paediatric Functional Independence Measure (WeeFIM)

The WeeFIM measure (Appendix 11) is an observational instrument, developed to assess daily functioning skills in children and was derived from the adult FIM (105)(140)(141)(120). It measures functional ability and independence in a variety of tasks and the level of assistance a child may need to achieve the task.

Functional performance is measured in three dimensions, namely self-care, mobility and cognition. It is not a self-report measure, but requires direct observation or interview or a combination of both by a professional who understands the rating system. Translation was, therefore, not required as the researcher studied the accompanying user manual and administered the forms. The instrument comprised of 18 items, each rated on an ordinal scale from 1-7, ranging from complete dependence (rated as “1”) through modified dependence from a person or device to complete independence (rated as “7”). The higher the dimension sub-score, the more independently the person is able to perform functions in that dimension. A maximum total score of 126 would indicate complete independence in all dimensions (105) (152). The scoring levels were:

No helper

- 7 = Complete Independence
- 6 = Modified

Helper – Modified Dependence

- 5= Supervision
- 4= Minimal assistance
- 3= Moderate assistance

Helper – Complete Dependence

- **2** = Maximal assistance
- **1** = Total assistance

In order to use the WeeFim, the researcher obtained an International Research License Agreement from Uniform Data System for Medical Rehabilitation (UDSMR) (

COMMUNICATION			
14.Comprehension		Mode: A – Auditory V - Visual C - Both	Yes No
15.Expression		Mode: V – Vocal N - Non-vocal B - Both	Yes No
SOCIAL COGNITION			
16.Social interaction			Yes No
17.Problem solving			Yes No
18.Memory			Yes No
Cognition total			
FIM total			

Appendix 12).

3.4.6.1 Reliability and validity studies

The measure had been found reliable and valid in detecting change in functional independence in children under eight years and in children with developmental disorders over eight years and has been used reliably in a variety of children up to sixteen years (143)(140)(105).

Rationale for using the measure:

The WeeFIM was chosen as a measure of functional ability and independence because it allows for a functional assessment of bedridden or wheelchair bound children, as well as children who are able to climb stairs independently. Therefore, this measure could be used for acutely-ill children in a hospital bed, children confined to a wheelchair and ambulant children (130). It was used to determine functional ability in the children with a health condition or disability, in order to explore the relationship between the EQ-5D-Y Mobility dimension and their actual functional independence.

3.4.7 The Faces Pain Scale-Revised (FPS-R)

The FPS-R (Appendix 13) is a self-report measure of pain intensity developed for children. It was developed using a series of facial expressions depicting pain intensity. The scoring ranges from 0-10 and can be used to self-rate pain intensity in children over eight years of age (153).

A horizontal line of six faces, showing progressively worsening pain expressions, are used to rate pain.

- The first face shows “no pain” = 0
- The next four faces show increasing pain, rated 2, 4, 6, 8 respectively; and
- The last face shows “very much pain” = 10

The child marks the face indicating his/her pain level. A rating of 10 would indicate severe pain.

The administrator of the measure needs to be mindful that their the explanation to the children should not to use words such as “happy” or “sad” because the child needs to measure how much pain they feel, not how the faces look (153).

As the instrument relies on pictures, only the instructions needed to be translated into Afrikaans and isiXhosa. No specific translation procedure was required by the developers of the instrument.

3.4.7.1 Reliability and validity study

The measure has been found to be reliable and valid in measuring pain levels in a clinical sample of children from five years and upwards (153).

No specific user agreement was required as the instrument could be photocopied for non-commercial clinical, educational or research purposes, directly from the International Association for the Study of pain (IASP) online at www.iasp-P/D.org/FPS-R.

Rationale for using the measure:

The FPS was easy to understand as it used pictures and allowed for a detailed rating of pain, which could then be compared to the pain dimension in the EQ-5D-Y measure.

3.5 Summary of Pilot study

Prior to the main research study a pilot study was undertaken to assess the reliability of the EQ-5D-Y in a smaller group of children from the same institutions as would be used in the main study. The pilot study is described in detail in Appendix 14.

The specific objectives of the pilot study were to:

1. Establish test-retest reliability of the EQ-5D-Y in children with a chronic or acute health condition and MS children by re-administering the measure the day after the initial assessment.

2. Determine the feasibility of administering the EQ-5-D-Y by assessing the time taken by the children to complete the form.

3. Determine the usefulness of obtaining contextual information from the medical files, using the self-designed contextual information sheet by assessing some of the demographic and medical information of participants.

In summary, the EQ-5D-Y was tested on 38 children and retested the next day. Nine children (23.7%) were drawn from the MS school, five (13.2%) from the SS, nine (23.7%) from the CI and 15 (39%) were from the acute hospital. The time taken for each child to complete the form was recorded. The contextual information sheet was used to obtain the extra information regarding the health status of each child.

Following analysis of the collected data, it was found that using the test-retest scores, kappa values for each dimension mostly fell within the moderately agreeable range (>0.50), except for UA dimension which was poorly agreed (<0.20). The ICC for VAS was 0.765 (95% Confidence intervals (CIs) =0.594-0.870), indicating good reliability. The children completed the measure in a mean time of six minutes, ensuring feasibility as it did not require much time to complete. It was possible to obtain useful information on each participant's demographics, medical condition and other factors which could influence HRQoL, using the self-designed contextual information form. This information was obtained from the medical file, nursing staff, therapists and from the children themselves.

As a result of the pilot study, the inclusion criteria needed to be amended to include typically developing children in grade seven who fulfilled the other criteria, as there were too few MS children in the 12 year age group. It was also found necessary to allow for more time in which to obtain consent from parents with children at the SS, as they were not in regular contact with their parents.

3.6 Procedure followed for main study

This section will describe the steps taken, after completion of the pilot study, in carrying out the main study.

3.6.1 Ethical approval

Ethical approval was obtained from the Human Research Ethics Committee of the University of Cape Town (HREC REF 354/2013) (Appendix 15). Permission was then sought from the various institutions at which the study took place (Appendix 16).

3.6.2 Research personnel and their role in the study

All clinical physiotherapists at the institutions were invited to participate in the study component related to the use of the EQ-5D-Y. Those who expressed interest were informed of the purpose of the study, their role in the study and were trained in the use of the EQ-5D-Y by the researcher. Their role included identifying children from their institution who fulfilled the inclusion criteria, giving the information packs to parents, explaining the purpose of the study and obtaining informed consent from the parents if they agreed to allow their child to take part in the study. The therapists were required to provide some of the information on each child's contextual information sheet. They were also initially required to administer the EQ-5D-Y to each child, but this was amended to whenever their time restraints allowed for this, as they felt that they would not always have the time to perform this task.

Using the additional contextual information sheet (Appendix 3), the clinical therapists at the institutions were required to report whether there had been any change in health status or management of the child since admission and subsequent administrations of the EQ-5D-Y. This information was in the medical files of each child, to which the therapists had ready access. Nursing staff, parents and the child themselves were questioned with regard to any life event that might have occurred in the last six months at home or in the family of the child, such as divorce or a death, moving home or a new baby, which could impact on the child's psychological well-being and the therapist recorded this on the information sheet. In the case of the MS children, the researcher examined the school records, questioned the class teachers and children themselves for this information.

Ultimately, the research personnel consisted of the principle researcher and clinical physiotherapists from the SS (three), from the CI (two) and AI (four), who all signed consent to their part in the study (Appendix 17).

3.6.1 Recruitment and requirements from participants

All 201 children identified as fulfilling the inclusion criteria from the MS school records were given informed consent forms (Appendix 18) in their homework books for parents to sign. After three verbal reminders, during a two week period, only 105 children returned legible, signed consent forms. The purpose of the study was then explained to these children, as well as their role in the study, which was to complete two HRQoL outcome measures (EQ-5D-Y and PedsQL), a pain outcome measure (FPS) and to be observed and questioned about functional independence (recorded by researcher on the WeeFIM). They were required to sign an assent form if they were willing to participate. Their right to withdraw from the study at any time was also explained to them. Subsequently, all 105 MS school children were included in the study.

The therapists examined the records at the SS and the CI and identified 45 and 47 children, respectively, fulfilling the inclusion criteria. They were all diagnosed with a chronic physical disability

(SS) or chronic health condition (CI). Identified children were given an informed consent form to be signed by a parent when they returned home for the weekend/holidays or when the parent visited the institution. Extra time was allocated for this process, prior to starting data collection as most of the children were not in regular contact with their parents. One month was assigned to obtaining signed consent forms, with weekly reminders from the researcher. Only 35 children at the SS and 32 children from the CI returned signed consent and assent forms and were included in the study. Their role was explained to them, as in the case of the MS children. They were also afforded the right to withdraw at any time.

At the AI, all children admitted during the six week data collection period allocated to the institution, and fulfilling the inclusion criteria, were recruited two days post admission by the therapist taking part in the study. Most children did have a parent with them in hospital and they were asked to sign an informed consent form after the therapists or researcher explained the process to both parent and child. A total of 52 children with signed consent and assent forms were included in the study.

A total of seven parents did refuse consent without stating a specific reason; two from the MS school, one form the SS, two from the CI and two from the Acute Care Hospital. A large number of consent forms (78) were not returned at the MS school.

Table 8 gives an outline of the research personal's duties and what measurements were taken at the different time intervals.

Table 8: Outline of procedure

INSTITUTION	Consent from:	Obtained once by:	Data collection timing	Instruments	Administered at the institution on the same day by:
MS	Institution Parent Child assent	Researcher Researcher Researcher	Baseline, three and six months	Contextual info	Teacher Child
				EQ-5D-Y PedsQL FPS WeeFIM (if necessary)	All measures administered by researcher
				EQ-5D-Y Proxy	Completed once, at three months by a parent, on the same day as self-report
SS	Institution Parent Child assent	Researcher School therapist Researcher	Baseline, three months and six months	Contextual info	School therapist Child
				EQ-5D-Y	Clinical physiotherapist (when possible, otherwise by researcher)
				PedsQL FPS WeeFIM	Researcher Researcher Researcher
				EQ-5D-Y Proxy	Completed by clinical physiotherapist, on same day as self-report
CI	Institution Parent Child assent	Researcher School therapist Researcher	Baseline, three months and six months	Contextual info	Clinical physiotherapist Child
				EQ-5D-Y	Clinical physiotherapist (when possible, otherwise researcher)

				PedsQL FPS WeeFIM EQ-5D-Y Proxy	Researcher Researcher Researcher Completed by clinical physiotherapist on same day as self-report
AI	Institution Parent Child assent	Researcher Ward therapist or researcher Researcher	Baseline and just prior to discharge OR at two weeks and one month later if still hospitalised	Contextual info EQ-5D-Y PedsQL FPS WeeFIM EQ-5D-Y Proxy	Ward therapist Child Ward therapist (when possible, otherwise researcher) Researcher Researcher Researcher Completed by ward physiotherapist, on same day as self-report

3.6.2 Baseline measures

Following recruitment and once all necessary informed consent was obtained, baseline measures were taken. At the institutions the therapists were asked to complete the EQ-5D-Y proxy before administering the EQ-5D-Y to the child to complete, but on the same day as the child completed the self-report. They were asked to complete the proxy as they felt the child would complete it (Proxy version 2). The therapists were not always available to administer the self-report to the children, it which case it was done by the researcher. At the MS school there were no therapists and the researcher administered the EQ-5D-Y to the children, but did not study the EQ-5D-Y self-reports of the children (in order to limit bias as much as possible). The researcher only ensured that there were no missing values, which was seldom. The parents of MS children were asked to complete EQ-5D-Y proxy measure (Proxy version 2).

On the same day as the EQ-5D-Y was administered by either the therapist or the researcher, the researcher independently administered the other self-reported HRQoL measures, the PedsQL and FPS to each child and completed the observational/ interview WeeFIM instrument. Small groups of eight children at a time were taken out of class, to a quiet room, to complete the outcome measures. As each small group was from the same class and at the same educational standard, the explanations of the use of the different measures were set appropriately for their level of understanding.

In the acutely-ill children this was performed at the bedside, usually without a parent present as the parents were often not available.

The use of each measure was first explained to the group and they were then monitored to ensure that all items on the two self-reported HRQoL measures were completed. Missing values occurred slightly more often on the longer PedsQL outcome measure. The possibility of inadvertently introducing bias will be discussed later. Following this, each child completed the self-reported FPS. The researcher then observed and questioned each child individually with regard to their functional ability and scored each item on the WeeFIM sheet. Children were then interviewed individually again to gain information needed on the contextual information sheet. After the children were returned to their class, the teacher was also interviewed for any additional contextual information of which they might be aware.

The same protocol was applied at all institutions at which the research took place, with some omissions at the MS school, which will be explained. The outcome measures were always administered in this order, with a short break in between measures, and under the same conditions in repeated assessments.

For the purposes of the study, it was assumed that the children at the MS school were TD children, with no functional problems. They would, therefore, show a high ceiling effect in baseline WeeFIM test scores and hence no change over time. The WeeFIM was, therefore, administered only to the MS children who reported a problem with "Mobility" on the EQ-5D-Y, in order to assess whether or not they actually did have a functional problem. This was done to assess the specificity of the test in accurately identifying individuals who do not have a problem with function. If after assessment any MS child was found to have a health related functional problem, they were to be referred to an appropriate health care provider for further assessment.

3.6.3 Repeated measures

Repeated measures were taken three months post baseline and again three months after that. Contextual information was also updated. Each child was therefore assessed three times over a seven month period, under the same circumstances as described for baseline. Some slight variations were needed at the AI.

At the AI, the child's stay was short, usually no longer than seven days. Baseline measures were taken two days post admission, thus allowing the children time to settle in, and repeated just prior to discharge, which was approximately five days post baseline. A third assessment was taken at two weeks or one month later, if the child was still hospitalised. Only the longer stay, hospitalised children were therefore assessments three times.

3.6.4 Proxy measures

The clinical therapists at each of the three institutions were asked to complete an EQ-5D-Y proxy for each child. This was done at baseline and also repeated at three monthly intervals. The parents with children at the MS school were asked to complete EQ-5D-Y proxy forms in the same way, but only did this at the three month post baseline assessment, as it was anticipated that there would not be much change in the proxy version over a seven month period in these children. The parents were instructed not to consult with their child when completing the proxy measure.

3.6.5 Usefulness and feasibility questionnaire

At the end of the data collection period, all therapists participating in the study, whether using the EQ-5D-Y measure or only filling in proxy measures, were asked to complete a questionnaire drawn up by the researcher, in order to assess feasibility and usefulness of including the measure in routine patient assessments (Appendix 4). This was done anonymously.

3.7 Data management, safety and monitoring

All participants completed paper versions of the instruments. The data were not anonymous as it was a longitudinal study and the responses of each child needed to be tracked. The completed paper versions were stored in a file in a locked cupboard in each institution's physiotherapy department. Identifying features were removed once data collection was completed and the data had been entered into an Excel spreadsheet for electronic analysis. This data was stored on a laptop which was password protected and will be kept for a maximum of four years. Secure backups were stored on a separate hard drive, in a password protected file

3.7.1 Data analysis

Statistical analysis was performed using Statistica 12 and IBM SPSS Statistics version 22. Both descriptive and inferential statistics were used. The way in which the data was analysed to address each research objective is described in Table 9.

The QALY weightings suggested by Craig et al, 2015, (66) were used in this study to produce a single composite score, the Index Score. The weighting produced by Craig et al (66) describe the values of losses in child HRQoL, assuming that each lasted 1 year. It is noted that the weightings are actually a decrement in QALY's and the Index Score should be calculated using 1-(the weighting). This was not done in this study, so the higher Index Score do indicate worse HRQoL. Where the Index Score is more than 1, it is assumed that the health state is worse than death. This EQ-5D-Y Index Score was used to compare against different measures which do sum dimension scores.

Table 9: Data analysis

Research Objective	**Description**	**Statistical Tests**
Describing the sample	Frequency tables were used to indicate how many participants were: • of a particular gender, • age, • educational level, • drawn from each institution • diagnosed with a specific health condition and • whether the health condition was chronic or acute.	• A One Way ANOVA was used to determine whether there were significant differences in mean values. • If the p value was =/<0.05, then post-hoc Tukey tests were performed to determine where the significant differences lay.
	The participants' length of stay at each institution, at baseline assessment, was also indicated.	• Minimum, maximum and Std. Dev. of length of stay.
Describing baseline assessments of each outcome, as listed below.	Baseline scores for each outcome measure were tested for distribution.	If not normally distributed, as in the case of EQ-5D-Y Index Score and VAS, then non-parametric analysis was

		used.
1. EQ-5D-Y dimensions	The level of problem reported in each dimension, across institutions.	Pearson Chi-square, P value.
2. EQ-5D-Y Index Score	The difference in mean rank and medians of Index Scores, across institutions.	Median score, mean rank.
3. EQ-5D-Y VAS	Comparing VAS across institutions.	Median score, mean rank, Kruskal-Wallis H test.
4. EQ-5D-Y Proxy dimension scores	Level of problems reported in each dimension across institutions.	Percentage of proxy vs. self-reporting problems on the different level.
5. EQ-5D-Y Proxy Index Scores	Comparing mean ranking of proxy index score across institution.	Median score, mean rank, Kruskal-Wallis H test
6. **Baseline PedsQL dimensions** scores and total score	Level of problems reported in each dimension across institutions.	Percentage of children from each institution with similar level of problems.
7. **Baseline WeeFIM dimensions** scores and total score	Level of problems reported in each dimension across institutions.	Percentage of children from each institution with similar level of problems.
8. Baseline FPS	Level of pain reported across institutions.	Kruskal-Wallis ANOVA by ranks.
To determine the performance of the EQ-5D-Y on children with different health statuses, by assessing the psychometric properties of the measure when used on these different groups of children:		
Reliability was assessed by: • **Test-retest reliability** of EQ-5D-Y was assessed in the pilot study (Appendix 14:).	Individual items of the EQ-5D-Y dimensions and the VAS were tested to establish consistency between test and retest scores, at each institution.	This was calculated using calculating Cohen's kappa coefficient for agreement between the two sets of dimension scores. ICC was used for correlations between test and retest VAS scores.
• **Intra-rater reliability of EQ-5D-Y** was assessed in the main study to determine stability of the measure, in the groups which were not expected to show a change in HRQoL (MS and SS).		
❖ Stability of EQ-5D-Y dimension scores over time.	The stability of dimension scores over time (from baseline to three months and to seven months) was examined at the institutions where the child's	Cohen's kappa coefficient for agreement between the dimension scores at different times was calculated for the repeated measures.

	health status was not expected to change (MS and SS). Correlations between scores were also given for children who were expected to show a change (CI and AI).	
❖ Stability of EQ-5D-Y Index Scores over time.	Consistency of index scores between baseline and second assessment.	This was calculated using calculating Cohen's kappa coefficient for agreement between the two scores.
❖ Stability of EQ-5D-Y VAS Scores over time.	Consistency of VAS scores between baseline and second assessment.	This was calculated using Spearman's correlation coefficient (Rho) and (p) and Friedman ANOVA.
○ **Intra-rater** reliability of **PedsQL** total.	PedsQL total scores were examined for consistency at the same time intervals as the EQ-5D-Y.	This was calculated using Spearman's correlation coefficient (Rho) and (p).
○ **Inter-rater reliability** of EQ-5D-Y was examined using proxy and self –reports.		
○ Comparing proxy and self-report dimension scores.	Level of agreement between dimension scores.	This was calculated using calculating Cohen's kappa coefficient.
○ Comparing proxy and self-report VAS scores.	Correlations between proxy and self-report VAS scores.	This was calculated using Spearman's correlation coefficient (Rho) and (p), across institutions. And ICC for all groups.
Validity was assessed by:		
• **Construct validity** of EQ-5D-Y was determined by examining both **discriminant** and **convergent** validity.		
❖ **Discriminant validity of the EQ-5D-Y,** was examined by comparing the profiles of the HRQoL of children with different health statuses, using dimension scores.	It was hypothesised that the measure would be able to discriminate between MS school children (with no health problems and therefore a high HRQoL), chronically disabled or ill children (with some problems on the different dimensions, caused by their health condition) and acutely ill children (with a lot of problems on all dimensions and therefore low HRQoL).	Kruskal-Wallis ANOVA by ranks, median scores and mean rankings were used to compare dimensions across groups.

❖ **Discriminant validity based on VAS** across institution.		
• Comparing VAS against the three levels of problems on the dimensions, across institutions	Comparing VAS score against the ranking of the different levels of dimensions.	Mean ranking, Kruskal-Wallis H
❖ **Discriminant validity of PedsQL** was also examined, to compare against discriminant validity of EQ-5D-Y.	Dimension sub-scores and total scores were compared across the different groups of children.	Box- Whisker graphs to demonstrate median scores, followed by Kruskal-Wallis H test to determine where differences lay.
❖ **Discriminant validity of the WeeFIM** was examined, to compare against EQ-5D-Y.	Dimension sub-scores and total scores were compared across the different groups of children.	Box- Whisker graphs to demonstrate median scores, followed by Kruskal-Wallis H test to determine where differences lay.
• **Convergent validity** was examined by comparing dimensions of EQ-5D-Y to similar dimensions on:		
❖ The PedsQL and ❖ The WeeFIM and ❖ The Faces Pain Scale	Correlations between baseline dimension scores were used. Similar dimensions on each measure should produce similar results, in each institution. Correlations between EQ-5D-Y and PedsQL dimensions and FPS are expected to be positive. Negative correlations are expected between EQ-5D-Y and WeeFIM dimensions.	Kruskal-Wallis H value or Mann-Whitney U z value, were used depending on the number of levels assessed, as well as p values.
• **Convergent validity** was further examined by comparing total scores.	Correlations between EQ-5D-Y Index Score and VAS are expected to be negatively correlated, as are Index Scores and WeeFIM. EQ-5D-Y Index Scores are expected to correlate positively with PedsQL total and FPS.	Spearman Rho is significant for correlations at ***p<.05***
Responsiveness of the different outcome measures over time was examined.	Repeat measures taken at three months post baseline were used to investigate the treatment effect, during this time period.	Effect size (r) was based on Wilcoxon signed rank test (Z) and was calculated by $(r=Z/Sq.Rt.N)^2$
Responsiveness of: • EQ-5D-Y dimension		

scores • VAS • Index score • FPS • PedsQL total • WeeFIM total Were all determined and compared over time, across institutions.		
To establish whether **changes in HRQoL are related to life events**:		
Changes or incidents in home life, surgery or change in management of condition was examined.	Assessing whether there was a significant change in VAS score or change in WSU dimension score, after a life event.	A t-test comparing mean "yes" with mean "no" answers to a life event was used.
To determine the **clinical feasibility and usefulness o**f using the EQ-5D-Y.	The time taken to complete the measure compared to the recommended time.	Mean, minimum, maximum and SD of time taken.
	Assessing the answers to usefulness questionnaire completed by the participating therapists.	Frequency of positive responses was analysed.

3.8 Ethical considerations

Ethical principles of autonomy, confidentiality, beneficence/non-maleficence and justice applied in the study were based on the Helsinki Declaration(154) and outlined below. Autonomy refers to the respect given to study participants to make their own decisions by ensuring that they have been given all the necessary information with which to make the decision. It includes confidentiality and maintaining the participant's privacy. Beneficence/non-maleficence refers to actions performed to the benefit of the participants and to prevent and remove any harm that may occur. Justice refers to equal distribution of burdens and benefits to all members of a group.

3.8.1 Autonomy

Prior to commencement of the study, consent was obtained from the various institutions at which the study was to take place (Appendix 16). To ensure autonomy, all participants in the study were asked to sign informed consent forms after the purpose of the study and their role was fully explained to them, in their home language if needed (a isiXhosa translator was available and the researcher could give the Afrikaans explanation). They were also made aware that they could refuse consent to taking part at any time in the study without negative consequences.

The role of the clinical physiotherapists interested in participating in the study was explained to them, after which they were asked to complete forms consenting to their participation in the study (Appendix 17).

Parents were required to consent to their child's participation in the study by signing an informed consent form in their home language, English, Afrikaans or isiXhosa, prior to collecting any data (English version Appendix 18).

The purpose of the study, the forms and information being used were described in the information pack each parent received. The participating therapists or the researcher also verbally explained the process to the parents at the initial contact or whenever the parent visited the child at each facility. An interpreter was used when necessary. When it was not possible to see the parents face-to-face, as in

the case of the MS children, information packs and informed consent forms were sent home with the children for their parents to sign. Permission to access contextual information from the nurse, teacher and school files was included in informed consent forms.

The children of the seven parents who refused consent were withdrawn from the study, without any consequences.

3.8.2 Confidentiality

Participants' information was kept confidential and secure in a locked cupboard and identifying names were removed once data capture was completed. The computer files were password protected and kept on secure computers. No participants were identified in the write up of the research.

3.8.3 Beneficence/non-maleficence

The study did not influence or affect any treatment the child was receiving and no extra costs were involved on the part of the participants; therefore, there was no need for monetary reimbursements. The study could potentially benefit children with a health condition if the information gained in HRQoL dimensions was used to assist the clinical physiotherapists in planning their management of each child. Regular assessment of HRQoL might, in fact, improve the quality of treatment.

As mentioned earlier, if any unidentified health related problem had been determined in a child at the mainstream school from the responses to the HRQol measures, this child would have been referred to the appropriate health provider by the researcher. This scenario, however, did not arise.

The study was minimal risk as there were no known risks to the participants and therefore no insurance was required for research-related injuries. In the unlikely event that a child became distressed during the assessment, the interview was to be terminated immediately and the treating therapist would counsel the child. This did not occur. Acutely ill children with abnormally high or low and unstable vital statistics were not asked to participate as it is likely that they and their parents would have found it too stressful.

If, by the end of this study, the EQ-5D-Y is found to be useful, a recommendation will be put forward to incorporate this measure into the undergraduate physiotherapy programme as a routine outcome measure to be used in paediatric practice.

3.8.4 Justice

Every child who was eligible for the study was recruited from the participating institutions.

It is hoped that findings from this study will help to stimulate research into HRQoL in children in South Africa. Therefore, justice would have been served as this is an important area of research in paediatrics, which is often overlooked. The paucity of local research articles attests to this.

As the EQ-5D-Y and other paediatric HRQoL measures take the child's perspective into account, it may be used as more than an outcome measure. In particular, it can be used to develop a value set for use in children, which is presently lacking.

4 CHAPTER 4: RESULTS

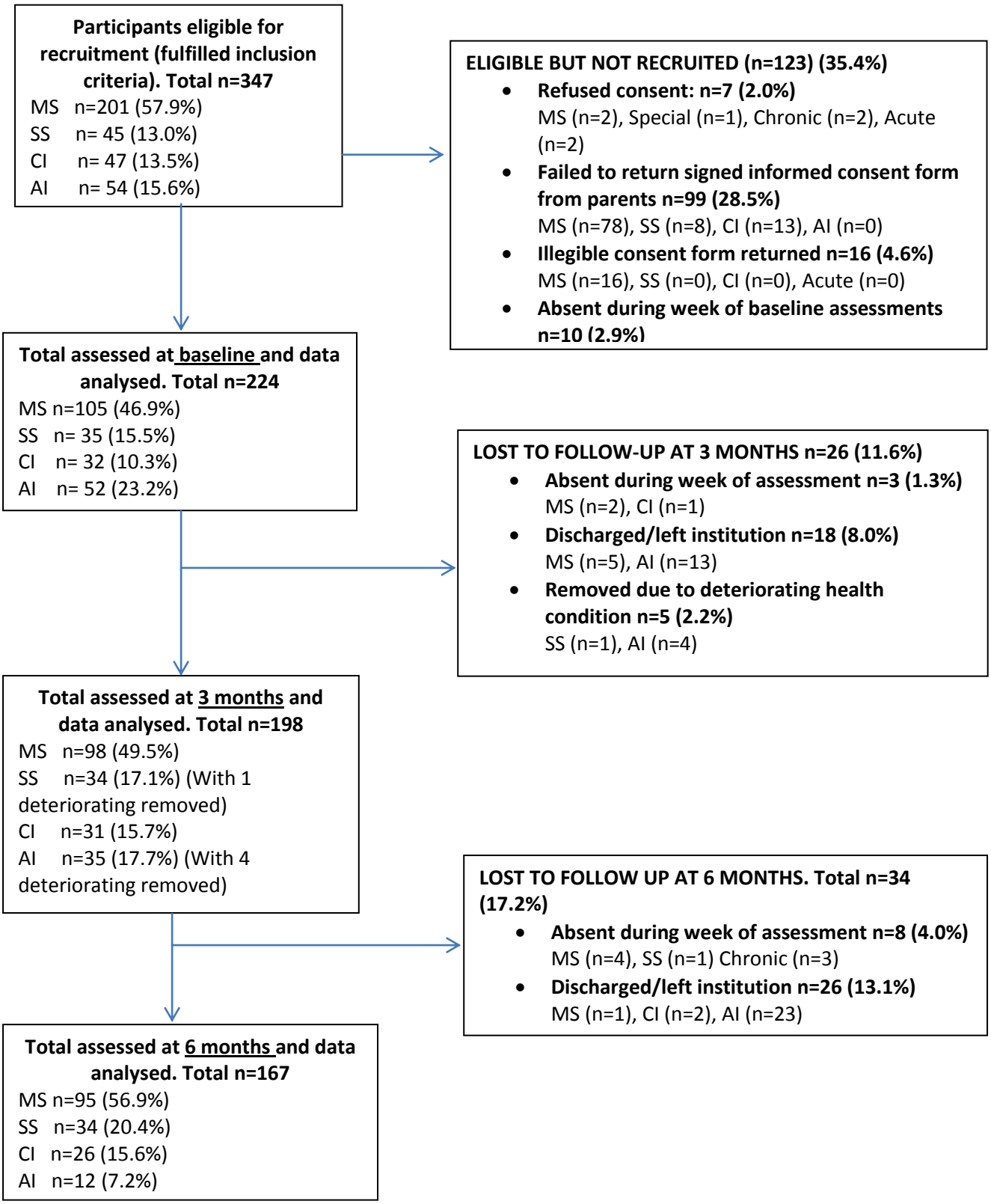
4.1 Sample

A total of 347 children from all the institutions fell within the identified age range and had successfully completed a minimum of two years of schooling. Figure 4 gives an overview of the number of participants at each three monthly stage of assessment and data analysis. Two hundred and one children from the MS school were recruited, while there were a total of 45 and 47 children with a health condition that were recruited at the SS and CI, respectively. Fifty four children were recruited from the AI where they fulfilled the inclusion criteria during the six week data collection period allocated to that institution. The parents of seven participants refused consent, while 99 children (78 of whom were at the MS school) failed to return signed consent forms. Another 16 returned forms that were illegible and those participants were excluded, as were the 10 children who were absent during the first assessment period.

A total of 224 children were, therefore, assessed at baseline. Of the 52 children from the AI, 12 were discharged before they were re-assessed at three days post admission and, of the remaining 40, 28 children were discharged before the third assessment, six days later.

At the second assessment stage, a total of 204 children were assessed and, of these, a further 39 were lost to follow-up at the third assessment period, resulting in a total of 167 children being assessed at the third stage. The 52 children at the AI were not assessed at three monthly intervals. The median for their length of stay was ten days, with a minimum of one day and a maximum of 131 days. Most children were assessed initially two days post admission and just prior to discharge which was approximately three to four days later. Only 12 children remained in hospital long enough to be assessed a third time.

Figure 4: Flow diagram indicating the number of participants at each stage of the study



4.1.1 Demographic details of the respondents

There were 119 males and 105 females (Table 10). Apart from the CI, there were more males in each of the institutions. Gender distribution was not significantly associated with institution (Chi-sq=1.43; p=.698)

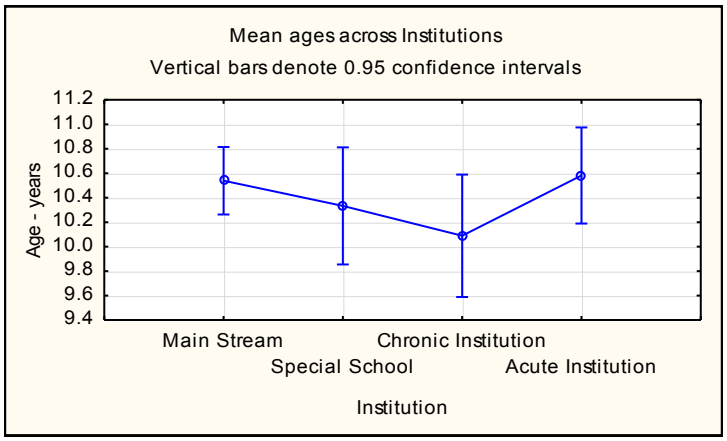
Table 10: Gender distribution in each institution

Institution	Gender		Totals (N)
	Male	Female	
MS	54	51	105
	51.4%	48.6%	
SS	21	14	35
	60.0%	40.0%	
CI	15	17	32
	46.9%	53.1%	
AI	29	23	52
	55.8%	44.2%	
Totals	119	105	224

N=224

Chi-sq=1.43, p=.698

The mean age of the children at the various institutions was 10.5 years (Std.Dev. = 1.45), with a minimum age of 7.0 years and a maximum of 13.8 years. A one-way ANOVA indicated that there was no significant difference in the mean ages of the children between the four institutions (p=0.379), although those in the CI were generally younger (Figure 5). The children from each institution were therefore similar in terms of age and gender.



N=224

F=(3, 220) = 1.03, p=0.379

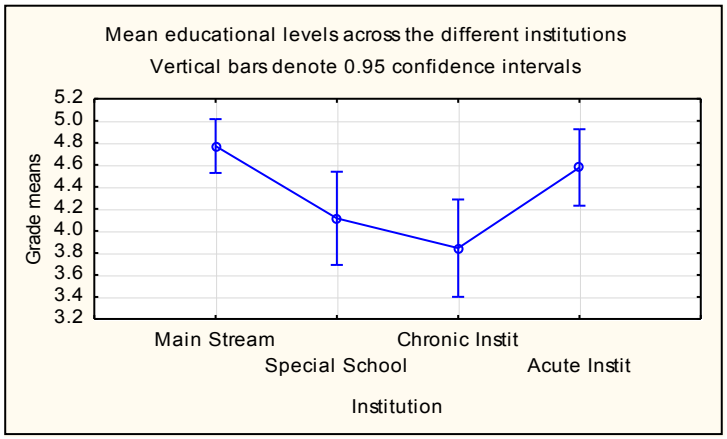
Figure 5: Mean age per institution

The distribution of children in each grade is given in Table 11. A one-way ANOVA (Figure 6) indicated that the mean educational grade of the children differed significantly between the institutions (p=0.001), with more MS children in higher grades. There were generally fewer children with health conditions in the higher grades and the children at the CI were mostly in the lower grades.

Table 11: Distribution (%) of educational grade at each institution

Institution	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Totals
MS	0	16	36	22	18	13	105
SS	1	10	12	8	4	0	35
CI	2	15	7	3	4	1	32
AI	0	17	10	9	10	6	52
All Groups	3	58	65	42	36	20	224
Percentage	1.3	25.9	29.0	18.8	16.1	8.9	100

N=224



N=224

F(3, 220)=5.581, ***p=.001***

Figure 6: Mean educational level across institutions

A post-hoc Tukey test (Table 12) indicated that the MS children were in higher grades than the SS or CI children. This difference was significant. In addition, the educational level of the children at CI was less than the AI and this difference was also significant.

Table 12: Post-hoc Tukey test determining where the difference in educational level lay

Institution	Mean	MS	SS	CI	AI
MS (n=105)	4.8		***0.040***	***0.002***	0.803
SS (n=35)	4.1	***0.040***		0.820	0.342
CI (n=32)	3.8	***0.002***	0.820		***0.050***
AI (n=52)	4.6	0.803	0.342	0.050	

Bold and italic indicate a significant difference (***p<.05***)

4.1.2 Health status of participants

Seventy eight participants reported a chronic condition which was either a condition with a functional limitation or a chronic health condition. The majority (67) of these participants were at either the SS or the CI (Table 13). Amongst the 105 children at the MS, 8 participants reported a minor health problem such as asthma and other allergic responses, and three of the participants at the acute care facility had chronic as well as acute health problems. All 52 participants at the acute care facility were being treated for their acute health problems.

Table 13: Chronic or acute health status of participants at each institution

Institution	Health status chronic	Health status acute	Health status none	Totals (N)
MS	8 (7.6%)	0	97 (92.4%)	105
SS	35 (100%)	0	0	35
CI	32 (100%)	0	0	32
AI	*3 (5.8%)	49 (94.2%)	0	52
All Groups	78 (34.8%)	49 (21.9%)	97 (43.3%)	224

N=224

* These three children were treated for an acute health problem, unrelated to their chronic condition.

Table 14 lists the health conditions reported. As was expected, the majority of the children at the MS school did not have a health condition, although eight of the total 105 reported minor ailments and three were diagnosed with mild learning difficulties. Cerebral palsy was the most common health condition seen overall with a total of 18 participants, 12 of whom were at the SS. There were ten children with spina bifida, again with the majority (9) being at the SS. The majority of the SS children were diagnosed with a disability which limits their functional independence. Ten children were diagnosed with Human Immunodeficiency Virus (HIV), the majority (8) of whom were from the CI. Diabetes and muscle diseases were the next most common health conditions and they were mostly from the CI. These children were admitted for management of their chronic health condition but did not experience many limitations in their functional independence. The most common conditions seen at the AI were neoplasms (6), appendicitis (7) and joint injuries (7) and all of the AI children were confined to bed. Three AI children were diagnosed with an acute condition superimposed on an existing chronic condition.

Table 14: Number and type of general health conditions of participants at each facility

General Health Conditions	MS	SS	CI	AI	Totals	%
None	94	0	0	0	94	42.0
Cerebral Palsy	0	12	4	2	18	8.0
Spina Bifida	0	9	0	1	10	4.5
Human Immunodeficiency Virus *(*HIV)	0	0	8	2	10	4.5
Diabetes	0	0	8	0	8	3.6
Muscle diseases	0	7	0	1	8	3.6
Spinal Cord Injuries	0	5	1	1	7	3.1
Neoplasms	0	0	1	6	7	3.1
Appendicitis	0	0	0	7	7	3.1
Joint injuries	0	0	0	7	7	3.1
Asthma and allergic responses	5	0	0	1	6	2.7
Cardiac defects, newly diagnosed	0	0	0	5	5	2.2
Burns	0	0	0	5	5	2.2
Infections	0	1	0	3	5	2.2
Traumatic fractures	0	0	0	3	3	1.3
Learning difficulty	3	0	0	0	3	1.3
Failure to thrive	0	0	3	0	3	1.3
Skeletal abnormalities, fractures/amputations	0	1	0	2	3	1.3
Syndromes	0	0	3	0	3	1.3
Traumatic Brain Injuries	0	0	2	0	2	0.9

Hepatitis	0	0	1	1	2	0.9
Other	3	0	1	5	8	3.6
All Groups	105	35	32	52	224	99.8

N=224

In summary the MS children were from the general population group with no disability. The children with a chronic disability and resulting limited mobility were from the SS, and the CI group were mostly diagnosed with a chronic health condition, but relatively functionally independent. The children from the AI were mostly previously TD, but were now limited to bed due to a short lived, acute condition.

For the full list of health conditions see Appendix 2.

4.1.3 Length of stay

The majority of MS and SS children had attended the respective school since the first grade of schooling, with a stay of approximately four years. The CI children were only admitted to the facility (and attended school there) once the family could no longer manage their chronic health condition at home or it had deteriorated and required special care. They therefore had spent less time at the institution at baseline assessment, approximately two years. The AI children were only admitted at the institution for a short period while their acute problem was managed, which was approximately seven days (Table 15).

Table 15: Length of stay of participants at each institution at baseline

Institution	Mean length stay at baseline in months	Minimum	Maximum	Std.Dev.
MS (n= 105)	50.9	0.1	84.1	20.4
SS (n=35)	46.8	9.6	96.0	19.2
CI (n=32)	27.8	0.5	97.0	25.4
AI (n=52)	0.4	0.0	4.4	0.8

N=224

4.2 Test-retest reliability of EQ-5D-Y dimension scores

This was performed in the pilot study (Appendix 14), on a small sample of convenience of children from each of the four institutions being used for the main study. A sample of convenience was used in that the children most likely to return a signed consent form were recruited. Cohen's kappa coefficient of agreement was used to determine agreement between the test and retest dimension scores (Table 16). According to Landis and Koch's interpretation of kappa (110), Mobility, LAM (Looking After Myself) and WSU (Worried, Sad or Unhappy) dimensions all fell within the moderately agreed range, while P/D (Pain/Discomfort) was fairly well agreed. Only UA (Usual Activities) dimension fell with the slightly agreed category.

Table 16: Agreement between first and second EQ-5D-Y dimension scores for pilot study.

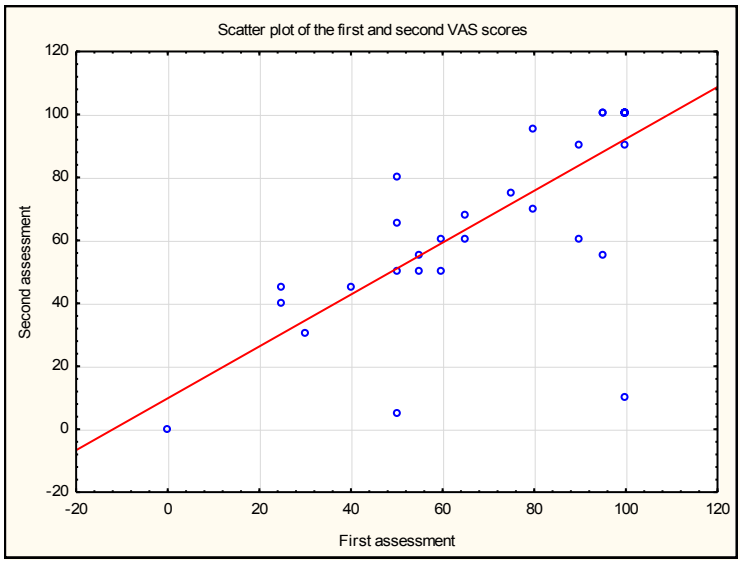
	Kappa	p
Mobility	0.546	***p<.001***
LAM*	0.653	***p<.001***
UA*	0.199	p<.127
P/D*	0.365	p<.08
WSU*	0.551	***p<.001***

N=38

*LAM (Looking After Myself)
UA (Usual Activities)
P/D (Pain/Discomfort)
WSU (Worried, Sad or Unhappy)

4.3 Test-retest reliability of VAS scores

This was also performed in the pilot study (Appendix 14) using the intraclass correlation coefficient, ICC and found to be 0.765 (95% Confidence intervals (CIs) =0.594-0.870), which is interpreted as strong agreement between the two scores according to Rankin and Stokes (1998) (155). As can be seen in Figure 7, apart from two outliers, the first and second VAS scores of the children were similar.



N=38

Figure 7: Test-retest EQ-5D-Y VAS scores for pilot study

4.4 Baseline assessment of outcome measures across all institutions

For the main study the baseline EQ-5D-Y, PedsQL, WeeFIM and FPS assessment results were described in frequency tables to determine the performance of the various measures when used on children with different health statuses (grouped by the different institutions). This was followed by statistical analyses to assess some psychometric properties of the EQ-5D-Y, using analyses of the other measures for comparison.

There were no missing responses on the EQ-5D-Y or PedsQL, as the researcher asked the child whether the missing response was due to the child not wanting to provide a score for that item or if the child had forgotten to provide a score for the item. All children with missing responses did so inadvertently and were willing to provide the missing score, without coercion. There were only two missing responses on the FPS out of a total of 224 responses, which would not affect the interpretation of results.

Each of the five EQ-5D-Y dimensions (Mobility, LAM, UA, P/D and WSU) was described per institution, using histograms and as the data were ordinal, non-parametric statistics was used. The VAS scores were not normally distributed (K-S d=.29, $p<0.01$), therefore non-parametric statistics were used to determine whether there were significant differences in overall HRQoL (VAS) between the children with different health states (as indicated by institution).

The EQ-5D-Y dimensions and their scores were described individually as they do not share the same underlying latent construct (they are multi-dimensional) and the intervals between each problem level within a dimension are not equal.

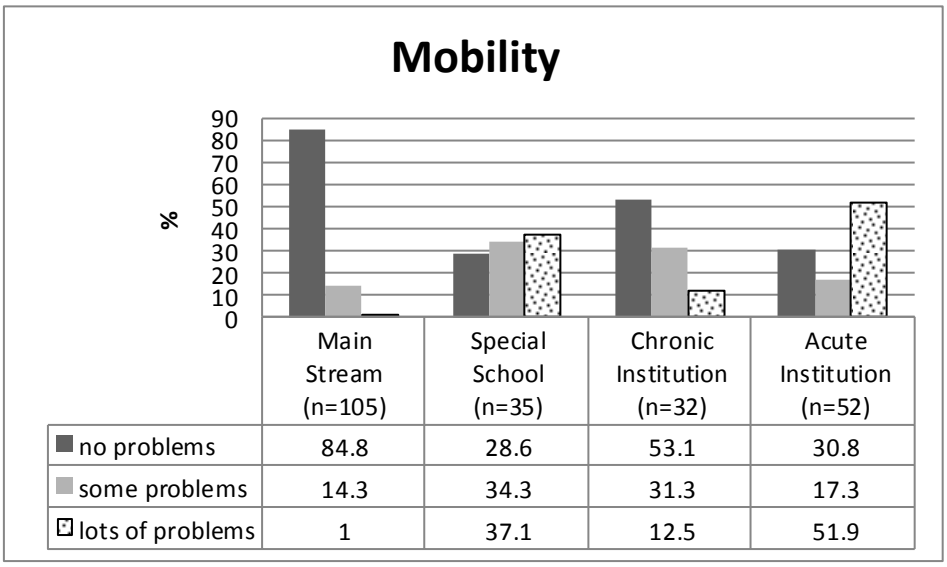
However, a composite score was developed (Index Score), using the QALY values suggested by Craig et al, 2015, (66). This recently published study is the first child health valuation study to use adult preferences on child health outcomes, to produce a composite score, derived from EQ-5D-Y dimension scores, based on a QALY scale. It should be noted that the results obtained by Craig et al have not yet been formally implemented by the EuroQoL Group.

4.4.1 Baseline EQ-5D-Y Dimension scores:

In total, 54.3% (57) of the MS children reported no problems on any dimension, compared to 11.8% (14) of the children with a health condition.

The level of problems reported in each dimension was associated with the institution, as indicated by the Pearson Chi-square statistic and p value ($p<0.001$ in all cases).

Just over 50% of the children at the AI reported "a lot of problems" (level 3) with Mobility, followed by 37% of the children at the SS. Only 13% of children at the CI reported "a lot of problems" with Mobility and as expected the MS children reported the least problems in this dimension, 85% with no Mobility problems (level 1) (Figure 8). The level of problems (i.e. no problems, some problems or lots of problems) reported was significantly different across institutions (Chi-Sq= 82.8, p<.001).



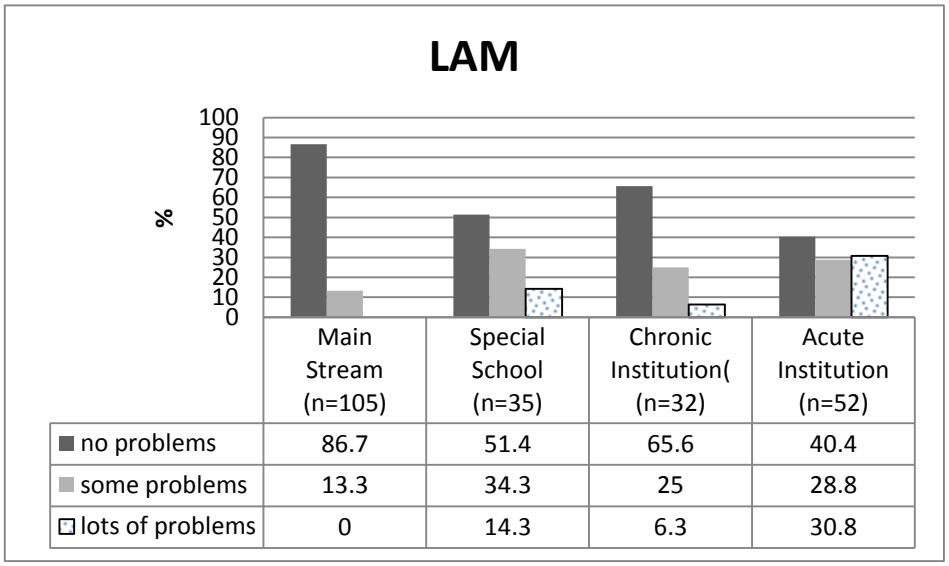
	Main Stream (n=105)	Special School (n=35)	Chronic Institution (n=32)	Acute Institution (n=52)
no problems	84.8	28.6	53.1	30.8
some problems	14.3	34.3	31.3	17.3
lots of problems	1	37.1	12.5	51.9

N=224

Chi-Sq= 82.8, ***p<0.001***

Figure 8: Percentage of Mobility dimension at each institution

At the AI 29% of children reported "some problems" (level 2) and 31% reported "a lot of problems" (level 3) in the Looking After Myself dimension (LAM). Similarly, 34% of children at the SS reported "some problems" with LAM, but only 14% reported "a lot of problems". Sixty-six percent of the children at the CI and 87% at the MS school reported "no problems" with LAM (Figure 9).



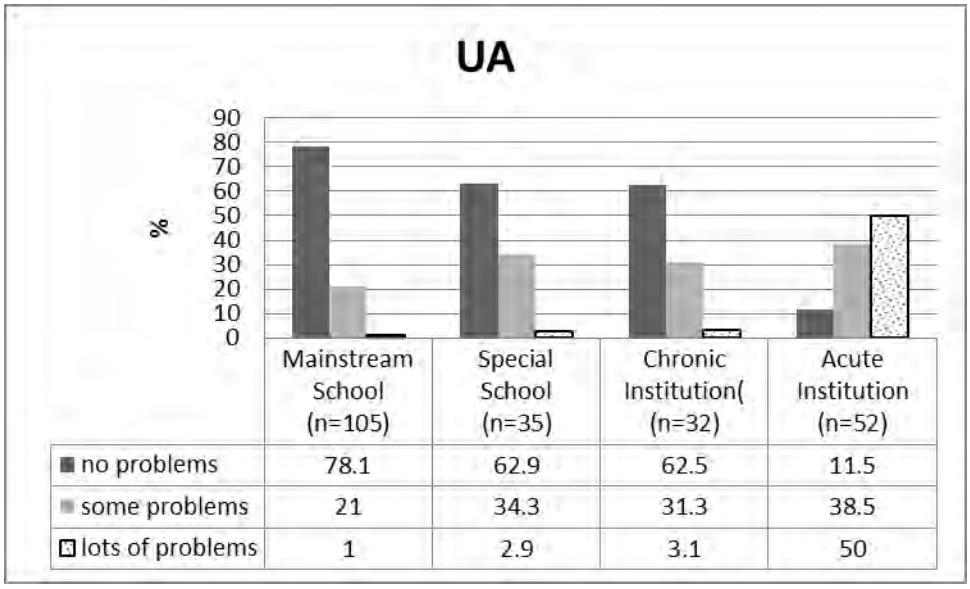
	Main Stream (n=105)	Special School (n=35)	Chronic Institution((n=32)	Acute Institution (n=52)
no problems	86.7	51.4	65.6	40.4
some problems	13.3	34.3	25	28.8
lots of problems	0	14.3	6.3	30.8

N=224

Chi-Sq= 53.0, ***p<0.001***

Figure 9: Percentage of Self-Care dimension (LAM) at each institution

Fifty percent of the children at the AI reported having" lots of problems" (level 3) with doing their Usual Activities (UA), whereas the majority of the children at the other three institutions reported" no problem" in this dimension (Figure 10).



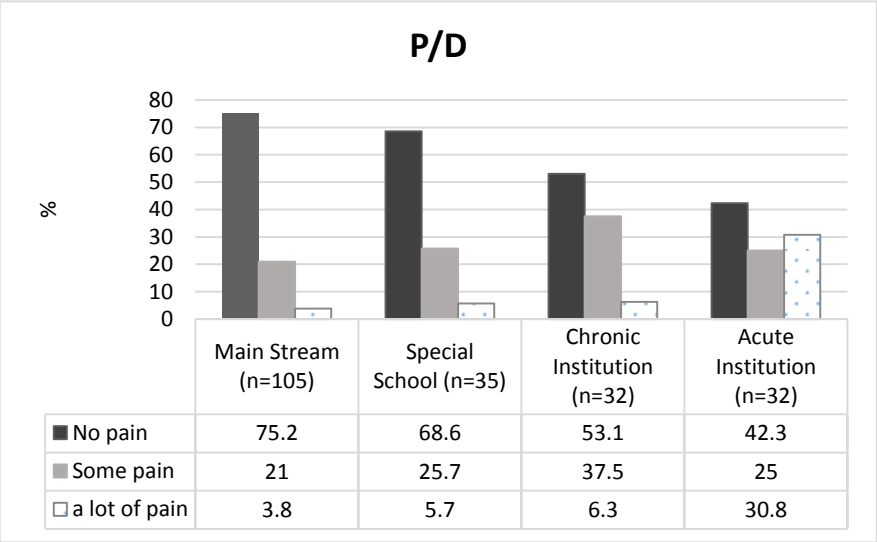
	Mainstream School (n=105)	Special School (n=35)	Chronic Institution((n=32)	Acute Institution (n=52)
no problems	78.1	62.9	62.5	11.5
some problems	21	34.3	31.3	38.5
lots of problems	1	2.9	3.1	50

N=224

Chi-Sq= 103.2, ***p<0.001***

Figure 10: Percentage of UA dimension at each institution

As expected 75% of the MS children reported "no" Pain /Discomfort (P/D) (level 1), whereas 31% of the children at the AI reported "a lot of" P/D (level 3). The majority of the children at the SS and at the CI reported "no" P/D (Figure 11).



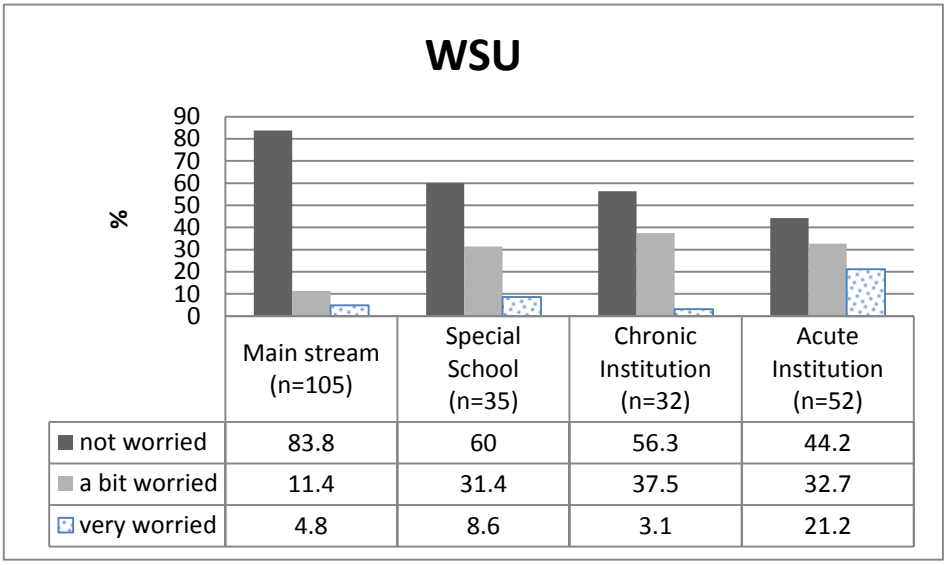
	Main Stream (n=105)	Special School (n=35)	Chronic Institution (n=32)	Acute Institution (n=32)
No pain	75.2	68.6	53.1	42.3
Some pain	21	25.7	37.5	25
a lot of pain	3.8	5.7	6.3	30.8

N=224

Chi-Sq= 33.9, ***p<0.001***

Figure 11: Percentage of P/D dimension at each institution

Eighty four percent of the MS children were “not” Worried, Sad or Unhappy (WSU), as opposed to 60%, 56% and 44% of the children at the SS, CI and AI respectively report “not” being WSU. About a third of the children at the SS, CI and AI all reported being “a bit” WSU and 21% of the AI children reported being “very” WSU (Figure 12).



	Main stream (n=105)	Special School (n=35)	Chronic Institution (n=32)	Acute Institution (n=52)
not worried	83.8	60	56.3	44.2
a bit worried	11.4	31.4	37.5	32.7
very worried	4.8	8.6	3.1	21.2

N=224

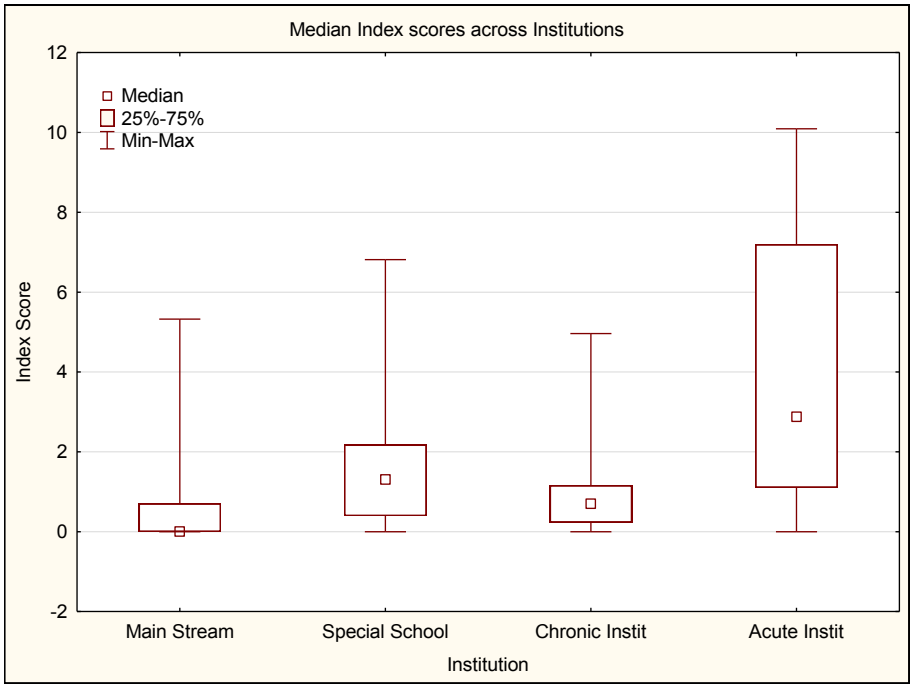
Chi-Sq= 32.3, ***p<0.001***

Figure 12: Percentage of WSU dimension at each institution

In summary, the children from the MS school had the least problems in all dimensions and on most levels of the EQ-5D-Y, the exception being the WSU dimension, in which 4.8% indicated problems on level 3 (lots of problems), compared to only 3.1% of children at the CI. The AI children reported the most problems on level 3 for all dimensions.

4.4.2 Index scores across institutions

The index scores were not normally distributed and non-parametric tests were used. In Table 17 it can be seen that the median Index Score of the MS children was 0.15, indicating minimal problems on any dimension. The AI children had the highest Index Score (2.8), followed by the SS (1.4) and the CC (0.75). Note that the higher scores indicate worse HRQoL.



N= 224

Figure 13: Median Index Score per institution

There was a significant difference in the mean rank of Index Scores across institutions (H=72.86, p<0.001) and Table 17 indicated that the MS children scored significantly differently from the other three groups, as did the AI children. There was however no difference between the SS and CI Index Scores.

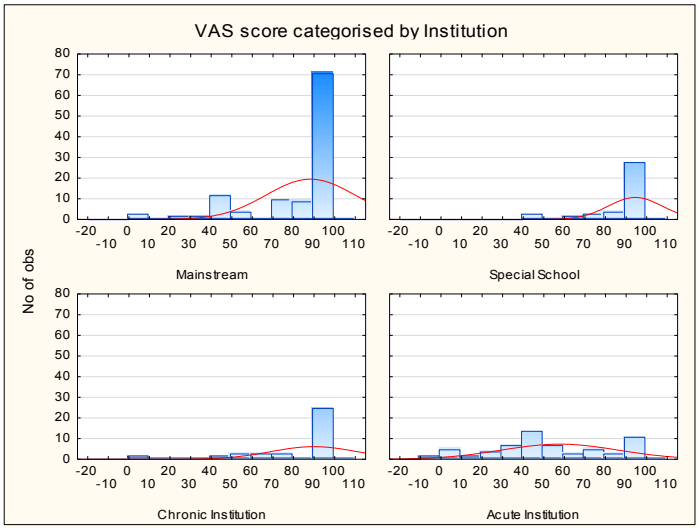
Table 17: Comparing Index Scores across institutions

Institution	Mean Rank	Median Score	MS	SS	CI	AI
MS (n= 105)	78.62	0.15		***p<0.001***	***0.045***	***p<0.001***
SS (n=35)	129.20	1.4	***p<0.001***		1.000	***0.030***
CI (n=32)	113.61	0.75	***.045***	1.000		***p<0.001***
AI (n=52)	168.98	2.8	***<.001***	***0.030***	***p<0.001***	

Kruskal-Wallis test: H (3, N= 224) =72.86 ***p<0.001***

4.4.3 Baseline EQ-5D-Y VAS scores

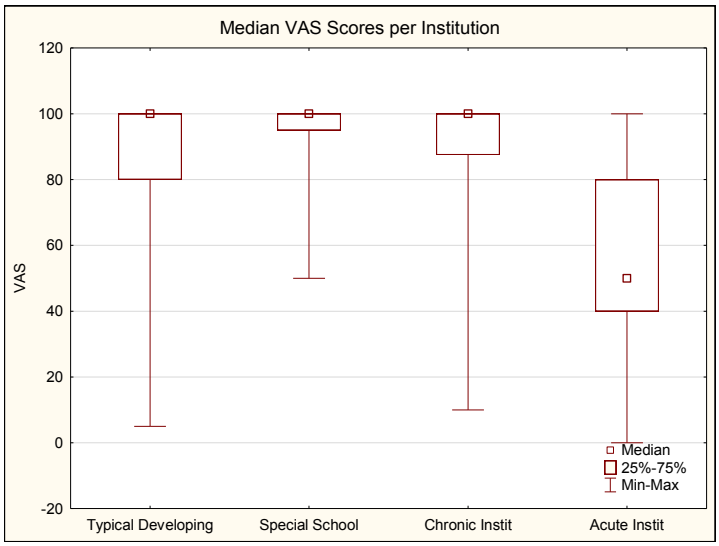
As can be seen in Figure 14, the scores of the VAS in all Institutions apart from the AI were not normally distributed (KS<0.01 throughout). Consequently, non-parametric statistics were used. There was no correlation between age and VAS (rho=-0.04, p=0.515)



N=224

Figure 14: VAS scores categorised by Institution

A higher VAS indicates better HRQoL. The median VAS was 100 for the MS (range 5-100), SS (range 50-100) and CI (range 10-100), indicating a strong ceiling effect. The median VAS of the AI was 50 (range 0-100) (Figure 15)



N=224

Figure 15: Median VAS per Institution

The Kruskal Wallace test indicated that there was a significant difference in the mean rank of VAS between the institutions ($p<0.001$). The mean ranking of the SS VAS was the highest, indicating better HRQoL, followed by CI and then MS. The AI children had the lowest ranked VAS, indicting worst HRQoL and it was significantly different from the other three groups.

Table 18: Comparison of ranking of VAS across institutions

Institution	N	Mean Rank	Median (range)	MS	SS	CI	AI
MS	105	125.9	100 (5-100)		1.00	1.00	***p<0.001***
SS	35	138.9	100 (50-100)	1.00		1.00	***p<0.001***
CI	32	132.3	100 (10-100)	1.00	1.00		***p<0.001***
AI	52	55.6	50 (0-100)	***p<0.001***	***p<0.001***	***p<0.001***	

Kruskal Wallis (H (3, N= 224) =62.81 ***p<0.001***)

4.4.4 Proxy scores per EQ-5D-Y dimension

Proxy forms were sent to all parents with children at the MS school and to all therapists managing children at the various institutions. Table 19 indicates that only 66 out of 98 parents with children at the MS school completed proxy forms. All the therapists managing children at the SS (34) and the CI (31) completed a proxy report. Only 11 proxy measures were completed by the therapists at the AI, as not all the children were receiving therapy.

Table 19: Proxy EQ-5D-Y dimension scores per institution

		MS (n=66)		SS (n=34)		CI (n=31)		AI (n=11)		Total (n=142)	
		count	%	count	%	count	%	count	%	count	%
Proxy Mobility	1	66	100	8	22.9	22	73.3	3	27.3	99	69.7
	2	0	0	11	31.4	5	16.7	3	27.3	30	21.1
	3	0	0	16	45.7	3	10	5	45.5	24	16.9
Proxy Self-care	1	64	97	14	40	22	73.3	2	18.2	102	71.8
	2	2	3	9	25.7	7	23.3	5	45.5	23	16.2
	3	0	0	12	34.3	1	3.3	4	36.4	17	12.0
Proxy UA	1	62	93.9	11	31.4	24	80	1	9.1	98	69.0
	2	3	4.5	12	34.3	5	16.7	3	27.3	23	16.2
	3	1	1.5	12	34.3	1	3.3	7	63.6	21	14.8
Proxy P/D	1	49	74.2	31	88.6	28	93.3	4	36.4	112	78.9
	2	15	22.7	4	11.4	2	6.7	6	54.5	27	19.0
	3	2	3	0	0	0	0	1	9.1	3	2.1
Proxy WSU	1	54	81.8	35	100	28	93.3	6	54.5	123	86.6
	2	10	15.2	0	0	2	6.7	3	27.3	15	10.6
	3	2	3	0	0	0	0	2	18.2	4	2.8

1=no problems, 2=some problems, 3=lots of problems

N=142

4.4.4.1 Comparing proxy and self-report dimension scores

Cohen's kappa coefficient of agreement was used to determine whether the dimension scores differed between children with a health condition and their therapists and MS children and their parents (Table 20). There was moderate to good agreement for the Mobility domain at all institutions and Fair to Moderate in the UA domain for groups except SS. Each institution had two to three domains which demonstrated fair to good agreement.

Table 20: Kappa values comparing Proxy EQ-5D-Y dimensions scores and child self- report dimensions scores

Institution	Kappa value for Mobility	Kappa value for LAM	Kappa value for UA	Kappa value P/D	Kappa value for WSU
MS (n=66)	*.000	-0.048 Slight	**0.363 Fair**	**0.222 Fair**	0.074 Slight
SS (n=34)	**0.551 Moderate**	0.198 Slight	0.082 Slight	-0.161 Slight	*.000
CI (n=31)	**0.835 Good**	**0.508 Moderate**	**0.420 Moderate**	0.082 Slight	0.007 slight
AI (n=11)	**0.725 Substantial**	0.061 Slight	**0.405 Moderate**	**0.457 Moderate**	**0.353 Fair**

N=142

*No kappa was computed because there was no variance in the proxy reporting for mobility at MS and WSU at SS. All proxies reported 'no problems' for mobility at MS and 'no problems' for WSU at SS.

Of the children with a health condition, the proxies reported fewer problems in all dimensions than the children from the CI. The proxies at the SS reported more problems in the Mobility, LAM and UA dimensions than the SS children did, but fewer problems with P/D and no WSU problems. The proxies at the AI reported more problems in all dimensions, except for WSU, than the AI children themselves reported (Table 21).

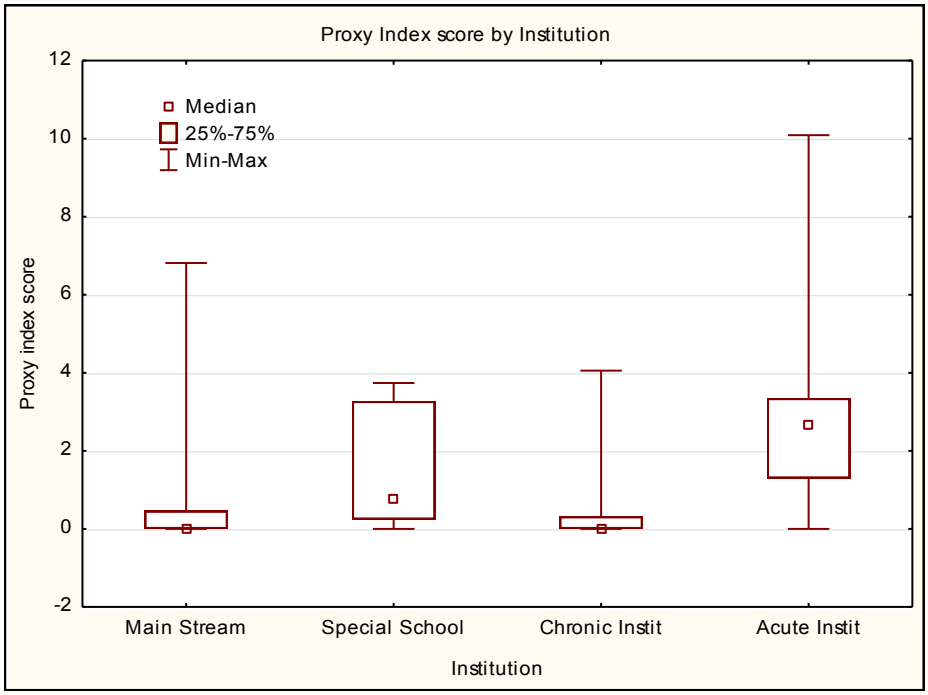
Table 21: Percentage of children and proxies reporting some or severe problems in each domain

	MS (n=66)		SS (n=34)		CI (n=31)		AI (n=11)	
Dimension	**% Self**	**% Proxy**	**% Self**	**% Proxy**	**% Self**	**% Proxy**	**% Self**	**% Proxy**
Mobility	15.2	0	71.4	77.1	41.9	25.8	63.6	72.7
LAM	13.3	3	48.6	60	34.3	19.3	63.6	81.8
UA	21.9	6.1	37.1	68.6	35.5	22.6	72.7	90.9
P/D	24.8	25.8	31.4	14.3	45.2	6.7	45.4	63.6
WSU	16.2	18.2	40	0	38.7	9.7	63.6	45.5

Shaded cells denote *more* problems reported

4.4.4.2 Index proxy scores across institutions

Similar to the self-report, there was a significant difference in the mean ranking of the proxy Index Scores between institutions (Kruskal-Wallis H =39.466 p <0.001) (Figure 16).



N=142

Figure 16: Comparing proxy Index Scores across institutions.

The significant differences in ranking of proxy Index Scores are depicted in Table 22. There were no differences in ranking between the MS and CI and between the SS and AI.

Table 22: Mean rank of proxy Index Scores across institutions

Institution	N	Mean rank	MS	*SS*	CI	*AI*
MS	66	59.17		***p<.001***	1.000	***p<.001***
SS	34	95.53	***P<.001***		***0.001***	0.890
CI	31	57.05	1.000	***0.001***		***p<.001***
AI	11	116.23	***p<.001***	0.890	***p<.001***	

Kruskal-Wallis test: H (3, N= 142) =39.466 ***p<0.001***

There was no significant difference in ranking of proxy VAS scores between MS and SS and between CI and AI (Table 23)

Table 23: Comparison of ranking of proxy VAS across institutions

	N	Mean Rank	Median (range)	MS	SS	CI	AI
MS	65	85.18	100 (0-100)		1	***0.001***	***p<0.001***
SS	35	82.4	100 (80-100)	1		***0.011***	***p<.0001***
CI	31	50.77	95 (70-98)	***0.001***	***0.011***		0.071
AI	11	14.41	70 (6-95)	***p<0.001***	***p<0.001***	0.071	

Kruskal-Wallis test: H (3, N= 142) =43.664 ***p<0.001***

4.4.4.3 *Comparing Proxy and self-report VAS scores*

The ICC for the child and proxy reports of the VAS of all respondents was .58 (n=142), which is considered moderate. However, at an institutional level, the proxy and child self-report VAS was only correlated significantly (p=0.015) at the MS School (r=0.30, a weak relationship) (Table 24). In contrast the correlation between child and proxy VAS, was moderate (r=0.50), but was not significant, possibly due to small sample size (n=11).

Table 24: P value and Spearman's Rho for Proxy: Child self-report for VAS

	r	p
MS	0.297	***0.016***
SS	0.080	0.653
CI	0.201	0.288
AI	0.503	0.115

N=142

The scatterplot in Figure 17 demonstrates the weak positive correlation between self- and proxy-reported VAS, across institutions.

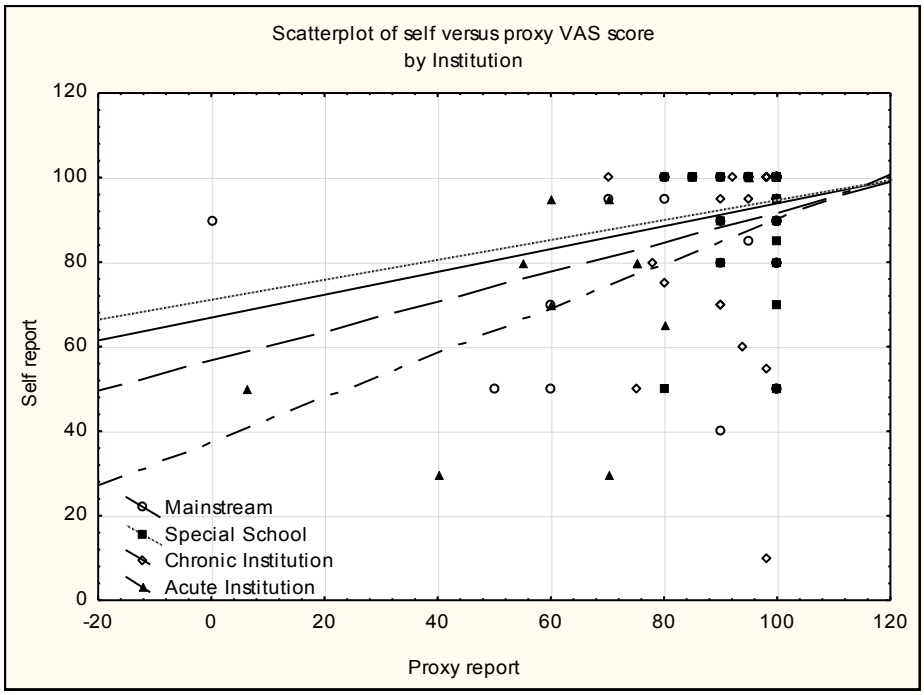


Figure 17: Self-report VAS versus proxy VAS, per institution

4.4.5 Discriminant validity of EQ-5D-Y at baseline

It was possible to examine the discriminant validity of the EQ-5D-Y by comparing the profiles of HRQoL of the different groups of children and determining whether the measure could discriminate between the groups.

4.4.5.1 *Examining discriminant validity by comparing the ranking of the different levels on each dimension across institutions*

Mobility Dimension

Using the Kruskal-Wallis ANOVA by ranks for Mobility, across the institutions, it was evident that there was a significant difference between the different groups (H=71.058 p<0.001) (Table 25). Multiple comparisons of Mobility rankings across institutions indicate that the significant differences were between the MS and SS and between the MS and the AI. The CI was only significantly different from the AI. There was no significant difference evident between the CI and MS and between the CI and the

SS, on the mobility dimension, indicating poor discriminant validity of this dimension when used on these children (Table 26).

Table 25: Kruskal-Wallis ANOVA by Ranks for Mobility, per institution

Mobility	N	Sum of Ranks	Mean Rank
MS	105	8512.5	81.07
SS	35	5174.0	147.83
CI	32	3579.0	111.84
AI	52	7934.5	152.59

Kruskal-Wallis test: H (3, N= 224) =71.058 ***p<.001***

Table 26: Multiple comparisons of mean ranking of Mobility problems, per institution

Mobility	MS	SS	CI	AI
MS		***p<0.001***	0.112	***p<0.001***
SS	***p<0.001***		0.139	1.000
CI	0.112	0.139		***0.031***
AI	***p<0.001***	1.000	***0.031***	

LAM Dimension

There was a significant difference in rankings for LAM, across the institutions (H=45.349 p<0.001) (Table 27). Multiple comparisons of mean rankings indicated that the significant differences were between the MS and SS and between MS and AI. There was again no significant difference evident between the CI and MS or between the CI and SS, on the LAM dimension, indicating poor discriminant validity when used on these children (Table 28).

Table 27: Kruskal-Wallis ANOVA by Ranks for LAM problems, per institution

LAM	N	Sum of Ranks	Mean Rank
MS	105	9439.5	89.90
SS	35	4566.0	130.46
CI	32	3525.0	110.16
AI	52	7669.5	147.49

Kruskal-Wallis test: H (3, N= 224) =45.349 ***p<0.001***

Table 28: Multiple comparisons of mean ranking of LAM problems, per institution

LAM	MS	SS	CI	AI
MS		***0.008***	0.730	***p<0.001***
SS	***0.008***		1.000	1.000
CI	0.730	1.000		0.062
AI	***p<0.001***	1.000	0.062	

UA Dimension

There was a significant difference in mean ranking of UA problems across institutions (H=85.311 p<.001) (Table 29). Multiple comparisons of mean rankings indicated that the significant differences were between the AI and the other three groups, which were not significantly different from each other (Table 30).

Table 29: Kruskal-Wallis ANOVA by Ranks for UA, per institution

UA	N	Sum of Ranks	Mean Rank
MS	105	9219.0	87.80
SS	35	3624.0	103.54
CI	32	3231.0	100.97
AI	52	9126.0	175.50

Kruskal-Wallis test: H (3, N= 224) =85.311 ***p<0.001***

Table 30: Multiple comparisons of mean ranking of UA, per institution

UA	MS	SS	CI	AI
MS		1.000	1.000	***p<0.001***
SS	1.000		1.000	***p<0.001***
CI	1.000	1.000		***p<0.001***
AI	***p<0.001***	***p<0.001***	***p<0.001***	

P/D Dimension

There was a significant difference in mean ranking of P/D problems across institutions (H=21.030 p<0.001) (Table 31). The only significant difference was between MS and AI. There was poor discriminant validity between the other groups (Table 32).

Table 31: Kruskal-Wallis ANOVA by Ranks for P/D, per institution

P/D	N	Sum of Ranks	Mean Rank
MS	105	10372.5	98.786
SS	35	3717.5	106.214
CI	32	3800.0	118.750
AI	52	7310.0	140.577

Kruskal-Wallis test: H (3, N= 224) =21.030 ***p<0.001***

Table 32: Multiple comparisons of mean ranking of P/D, per institution

P/D	MS	SS	CI	AI
MS		1.000	0.763	***0.001***
SS	1.000		1.000	0.092
CI	0.763	1.000		0.803
AI	***0.001***	0.092	0.803	

WSU Dimension

There was a significant difference in mean ranking of WSU problems across institutions (H=25.895 p<0.001).The only significant difference was again between MS and AI. There was poor discriminant validity between the other groups for WSU (Table 33 and Table 34).

Table 33: Kruskal-Wallis ANOVA by Ranks for WSU dimension, per institution

WSU	N	Sum of Ranks	Mean Rank
MS	105	9946.5	94.729
SS	35	4213.5	120.386
CI	32	3810.0	119.063
AI	52	7230.0	139.038

Kruskal-Wallis test: H (3, N= 224) =25.895 **p<0.001**

Table 34: Multiple comparisons of mean ranking of WSU, per institution

WSU	MS	SS	CI	AI
MS		0.255	0.378	***p<0.001***
SS	0.255		1.000	1.000
CI	0.378	1.000		1.000
AI	***p<0.001***	1.000	1.000	

The difference between the Index Scores across institutions was presented in 4.4.2 and indicated a significant difference in ranking between MS and the other three groups and between AI and the other three groups. There was however no difference between the SS and CI Index Scores.

4.4.5.2 Examining discriminant validity by comparing VAS across Institutions

The EQ-5D-Y VAS only indicated a significant difference in the mean rank of VAS scores between the AI and the other three institutions (p<0.001 for all three) and therefore showed poor discriminant validity across the other three groups (Table 18).

4.4.5.3 Summary of discriminant validity of EQ-5D-Y

Discriminant validity was evident between the AI and the MS school, only, on all dimensions. There was particularly poor discriminant validity between the SS children and the CI children.

In the Mobility dimension, the EQ-5D-Y could discriminate between MS and SS and between the MS and the AI, as well as between AI and CI, all of whom did have different levels of mobility. However no difference in Mobility depicted between CI and MS and between CI and SS, despite differences in mobility between these groups.

In the LAM dimension, differences were depicted between MS and SS between MS and AI. There was no difference depicted between the AI and SS and between AI and CI. The discriminant validity was particularly poor for the CI on this dimension, with no significant difference between it and any other group.

On the UA dimension there was a significant difference between the AI and the other three groups, which were not significantly different from each other.

On the P/D and WSU dimensions the only significant difference was again between MS and AI. There was poor discriminant validity between the other groups for WSU.

The Index Score indicated a significant difference in ranking between all institutions apart from between the SS and CI.

The EQ-5D-Y VAS could only discriminate between the AI children with significantly lower VAS (poorer HRQoL) and the other three groups.

4.4.6 Comparison of VAS against the three levels of the dimensions, across institutions.

The VAS score was compared against the ranking of different levels of the dimensions, across institutions using Kruskal Wallis H statistic and p value and was found to be significant on all dimensions at the AI (Table 18). A significant difference was also evident for WSU dimension, at the SS, but was not detected with multiple comparisons of mean ranking (Table 34).

When comparing VAS across the mean ranking of the three levels on each dimension, it was evident that this was significant at the AI only (Table 35).

Table 35: Comparison of VAS against the three levels of the dimensions, across institutions.

	MS	**SS**	**CI**	**AI**
Mobility	H=2.135, NS	H=1.530, NS	H=1.870, NS	***H=11.227, p=0.004***
LAM	H=.00, NS	H=2.538, NS	H=.813, NS	***H=8.958, p=0.011***
UA	H=1.56, NS	H=.429, NS	H=2.36, NS	***H=5.22, p<0.001***
P/D	H=4.440, NS	H=1.166, NS	H=1.722, NS	***H=23.334, p<0.001***
WSU	H=4.295, NS	***H=6.412, p=0.041***	H=2.110, NS	***H16.977, p<0.001***

NS, Not Significant
N=224

4.5 Baseline PedsQL measure across all institutions

Children who “never had a problem” in each item of the four PedsQL dimensions (Activity, Feelings, Socialising and Schooling) were described per institution in a frequency table (Table 36). The difference in ranking for PedsQL dimension subtotals of the different health conditions was analysed using Kruskal-Wallis (as the data was ordinal) to determine where the difference between institutions lay. This was then illustrated graphically with Box-Whisker graphs.

The full baseline PedsQL can be seen in (Appendix 21).

Table 36: Baseline PedsQL of children who "never had a problem" in the various items in each dimension

		MS (n=105)		**SS** (n=35)		**CI** (n=32)		**AI** (n=52)	
		count	%	count	%	count	%	count	%
ACTIVITY: Hard to walk more than one block	Never a problem	85	**81.0**	11	31.4	17	53.1	18	34.6
Hard to run	Never a problem	62	**59.0**	6	17.1	17	53.1	13	25.0
Hard to do sport or exercise	Never a problem	72	**68.6**	9	25.7	12	37.5	12	23.1
Hard to lift something heavy	Never a problem	38	**36.2**	7	20.0	6	18.8	11	21.2
Hard to bath or shower by myself	Never a problem	94	**89.5**	22	62.9	24	75.0	28	53.8
Hard to do chores around house	Never a problem	74	**70.5**	15	42.9	14	43.8	19	36.5
I hurt	Never a problem	35	33.3	15	**42.9**	6	18.8	19	36.5
Have low energy	Never a problem	46	**43.8**	14	40.0	12	37.5	17	32.7
FEELINGS: Feel afraid or scared	Never a problem	44	41.9	18	**51.4**	7	21.9	22	42.3
Feel sad	Never a problem	43	41.0	14	40.0	9	28.1	25	**48.1**
Feel angry	Never a problem	27	25.7	10	28.6	3	9.4	35	**67.3**
Have trouble sleeping	Never a problem	42	40.0	21	**60.0**	12	37.5	30	57.7
Worry what will happen to me	Never a problem	32	30.5	15	**42.9**	7	21.9	15	28.8
SOCIALISING: Have trouble getting along with other kids	Never a problem	44	41.9	27	**77.1**	13	40.6	28	53.8
Other kids do not want to be my friends	Never a problem	53	50.5	16	45.7	16	50.0	30	**57.7**
Other kids tease me	Never a problem	39	37.1	10	28.6	8	25.0	26	**50.0**
Cannot do things other kids my age can do	Never a problem	71	**67.6**	12	34.3	19	59.4	20	38.5
Hard to keep up when I play with other kids	Never a problem	55	52.4	10	28.6	18	**56.3**	18	34.6
SCHOOLING: Hard to pay attention in class	Never a problem	74	**70.5**	28	80.0	16	50.0	31	59.6
I forget things	Never a problem	36	**34.3**	12	**34.3**	11	**34.4**	18	**34.6**
Have trouble keeping up with schoolwork	Never a problem	71	**67.6**	23	65.7	15	46.9	24	46.2
Miss school because not feeling well	Never a problem	52	**49.5**	10	28.6	8	25.0	4	7.7
Miss school to go to Dr or hospital	Never a problem	46	**43.8**	11	31.4	4	12.5	0	0.0

N=224

Largest % highlighted in **bold**

In the Activity and Schooling dimension, the MS children reported having the least problems. However, in the Feelings and Socialising dimensions, children in the other institutions reported less problems in most of the items e.g. 77% of children at the SS children reported no problems with "Have trouble getting along with other kids", compared to 42% of MS children.

4.5.1 PedsQL dimension scores across all institutions

4.5.1.1 Discriminant validity

The ability of the PedsQL to depict significant differences between the different groups of children was examined for discriminant validity. Kruskal-Wallis analysis was used to determine whether there was a difference in ranking for dimension subtotals of the different health conditions at each institution and graphically illustrated with Box-Whisker graphs.

4.5.1.2 PedsQL dimension subtotals, across all groups

There was a significant difference in ranking for PedsQL Activity, Feelings and Schooling dimensions across all the institutions, but not for the Socialising dimension (Table 37), indicating fair discriminant validity.

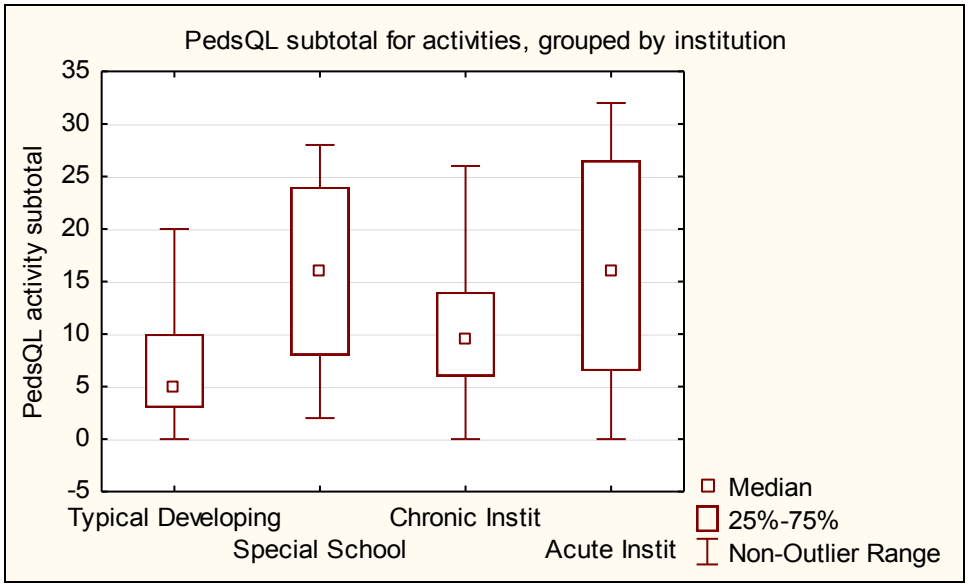
Table 37: Kruskal-Wallis H and p values for each PedsQL dimensions subtotal, across all groups

Institution	PedsQL dimension sub scores	N	Kruskal-Wallis H value	p value
All groups	Activity	224	35.31	***p<0.001***
All groups	Feelings	224	14.38	***0.002***
All groups	Socialising	224	6.23	0.1
All groups	Schooling	224	20.86	***p<0.001***

N=224

4.5.1.3 PedsQL subtotal score for Activity dimension, per institution

Figure 18 illustrates the medians for Activity score across institutions. The MS Activity scores were ranked significantly lower (fewer problems) than those of the SS and AI (Table 38).



N=224

Figure 18: PedsQL median for Activity subtotal per institution

Table 38: PedsQL Activity mean ranking across institutions

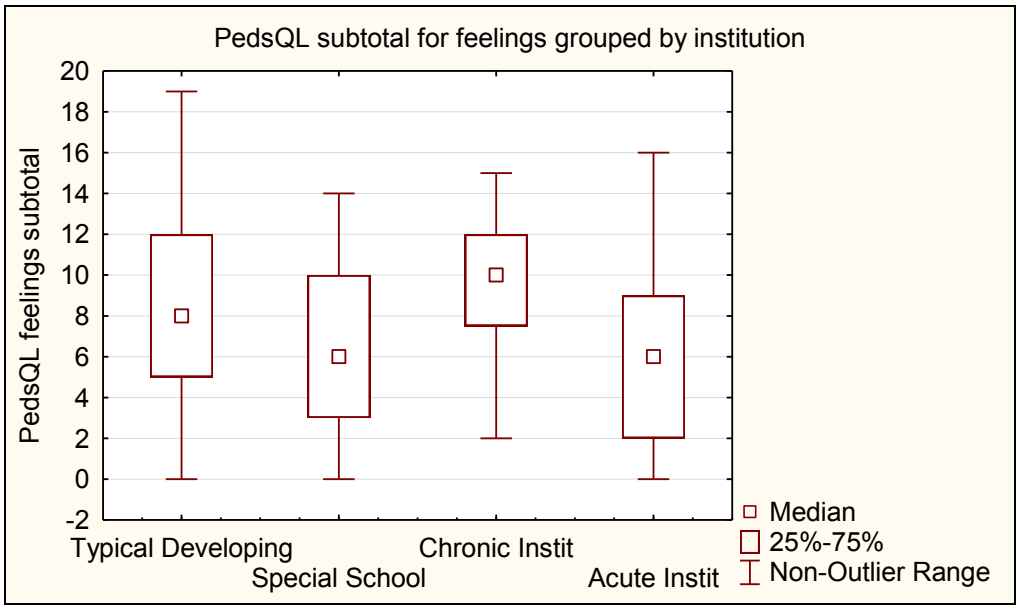
	Mean rank	N	MS	SS	CI	AI
MS	86.4	105		***p<0.001***	0.081	***p<0.001***
SS	139.1	35	***p<0.001***		1.000	1.000
CI	118.7	32	0.081	1.000		0.528
AI	143.5	52	***p<0.001***	1.000	0.528	

Kruskal-Wallis test: H (3, N= 224) =35.305 ***p <0.001***

4.5.1.4 PedsQL subtotal score for Feelings dimension, per institution

Figure 19 illustrates the medians for Feelings score across institutions.

Table 39 indicated that there was a significant difference in ranking for PedsQL Feelings subgroup, at the different institutions (p=0.002). The Feelings scores of the children in the CI were ranked significantly higher (more problems) than the SS and the AI and between the AI and CI.



N=224

Figure 19: PedsQL median for Feelings across institution

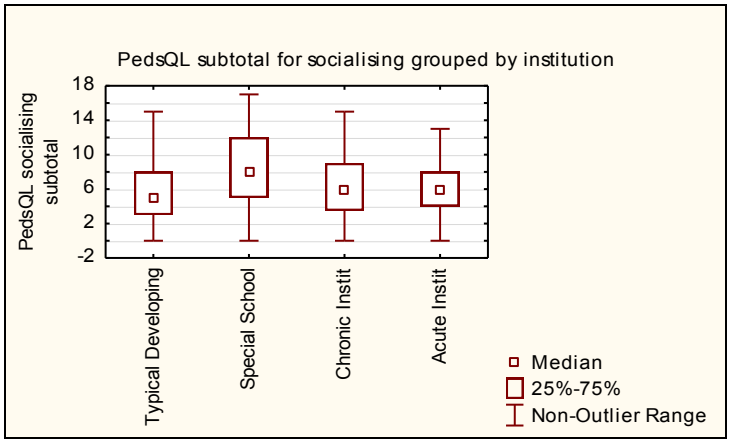
Table 39: PedsQL Feelings mean ranking

	Mean rank	N	MS	SS	CI	AI
MS	119.8	105		0.389	0.762	0.071
SS	97.4	35	0.389		***0.038***	1.000
CI	139.7	32	0.762	***0.038***		***0.006***
AI	92.1	52	0.071	1.000	***0.006***	

Kruskal-Wallis test: H (3, N= 224) =14.378 ***p =0.002***

4.5.1.5 PedsQL subtotal score for Socialising dimension, per institution

Figure 20 illustrates the medians for Socialising score across institutions. There was no significant difference in the ranking for PedsQL socialising dimension across the institutions (p=0.1).



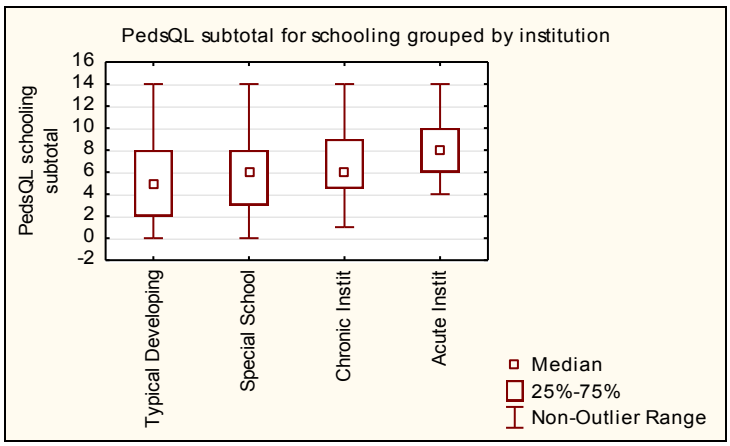
N=224

Figure 20: PedsQL median for socialising across the institutions

4.5.1.6 *PedsQL subtotal score for Schooling dimension, per institution*

Figure 21 illustrates the medians for Socialising score across institutions.

Table 40 indicated that there was a significant difference in ranking for PedsQL schooling subgroup, at the different institutions ($p<0.001$). The Schooling score of the children in the AI was ranked significantly higher (more problems) than the MS and SS children.



N=224

Figure 21: PedsQL Schooling median at each institution

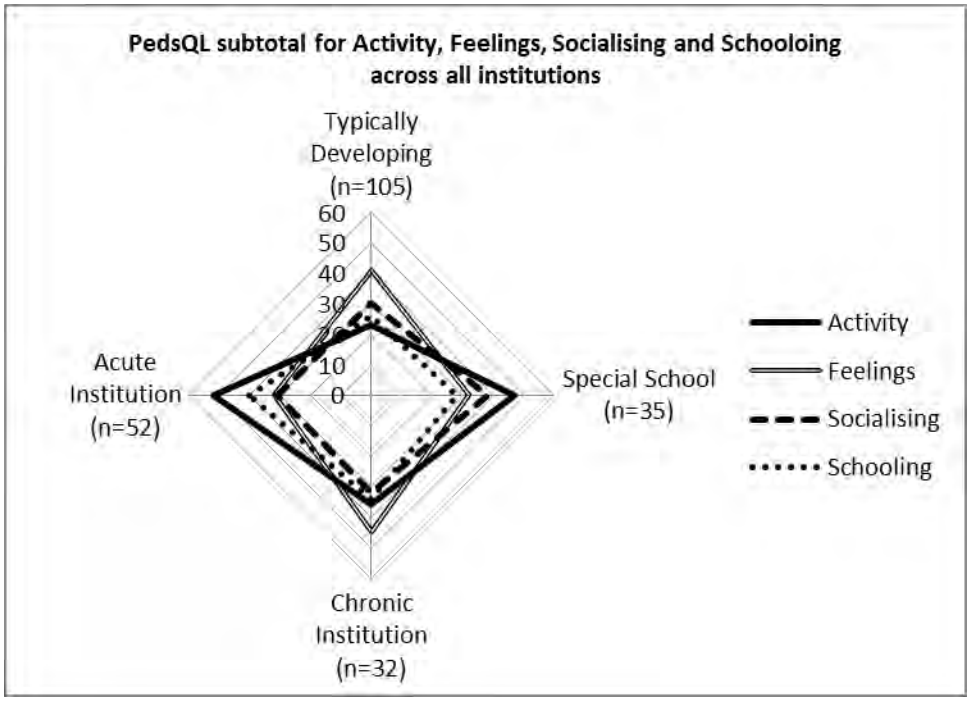
Table 40: PedsQL mean ranking for schooling

	Mean ranking	N	MS	SS	CI	AI
MS	97.0	105		1	0.172	***p<0.001***
SS	100.6	35	1		0.684	***0.014***
CI	125.6	32	0.172	0.684		1
AI	143.7	52	***p<0.001***	***0.014***	1	

Kruskal-Wallis test: H (3, N= 224) =20.860 ***p <.001***

4.5.2 PedsQL dimensions sub scores across institutions

Figure 22, a radar graph, illustrates the weighted scores for the four dimensions assessed in the PedQL. A lower score indicted fewer problems in that dimension. The MS children had the fewest problems with Activities, Socialising and Schooling, whereas the AI children had the most problems with Activity and Schooling. The SS children had the most problems with Socialising and the CI children experienced the most problems in the Feelings dimension.



N=224

Figure 22: PedsQL subtotal for activity, feelings, socialising and schooling at each institution

4.5.3 PedsQL total scores across institutions

The PedsQL total scores were not normally distributed Figure 23.

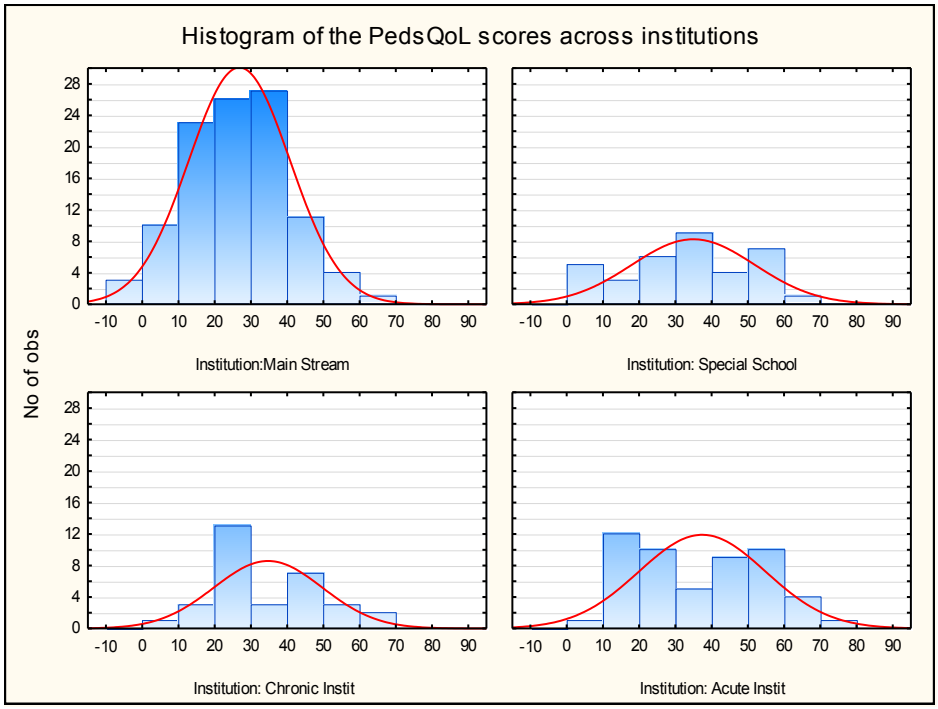


Figure 23: Histogram indicating distribution of PedsQL total scores

A higher PedsQL score (obtained from a sum of the dimension scores) indicates worst HRQoL. The AI demonstrated the highest score and therefore worst HRQoL (median37, range 9-73), followed by the

SS (median 35, range 4-65) and CI (median 30, range 7-64). The MS children scored the lowest (median 27, range 0-66), indicting the best HRQoL.

The Kruskal Wallace test (Table 41) indicated that there was a significant difference in the mean rank of PedsQL total score between the institutions (p=.001) and that the ranking of the MS children was significantly lower than the AI children. There were no other significant differences in ranking, indicating poor discriminant validity when used on children with chronic conditions or disabilities.

Table 41: Comparison of ranking of the PedsQL total across institutions

	Mean Rank	N	MS	SS	CI	AI
MS	94.4	105		0.072	0.123	***0.003***
SS	126.2	35	0.072		1.000	1.000
CI	124.7	32	0.123	1.000		1.000
AI	132.3	52	***0.003***	1.000	1.000	

Kruskal-Wallis test: H (3, N= 224) =15.740 ***p =.001***

4.5.4 Summary of performance of PedsQL across the different health conditions, indicating discriminant validity

Three of the PedsQL dimensions, Activity, Feelings and Schooling indicated a significant difference in ranking across the institutions. This did not apply to the Socialising dimension.

The PedsQL total score ranking was significantly different (lower indicating better HRQoL) for the MS children compared to the ranking for the AI children. It was not possible to differentiate between the MS, SS, and CI children using the PedsQL total score ranking.

4.6 Baseline WeeFIM measure across all institutions

The WeeFIM, a measure of functional independence, consisting of three dimensions (self-care, mobility and cognition), is described for “complete independence” per institution in Table 42. This was followed by Kruskal Wallis ANOVA to determine the impact of different health conditions on functional independence of children at each institution.

Table 42: Baseline WeeFIM for "complete independence" of all items in each dimension

	MS (n=16)		**SS** (n=35)		**CI** (n=32)		**AI** (n=52)	
	Count	%	Count	%	Count	%	Count	%
Self-care: Eat	16	100.0	34	97.1	31	96.9	33	63.5
Groom	12	75.0	30	85.7	27	84.4	34	65.4
Bath	15	93.8	28	80.0	28	87.5	16	30.8
Dressing upper body	16	100.0	28	80.0	28	87.5	24	46.2
Dressing lower body	16	100.0	21	60.0	27	84.4	17	32.7
Toileting	16	100.0	23	65.7	28	87.5	36	69.2
Bladder management	16	100.0	19	54.3	29	90.6	23	44.2
Bowel management	16	100.0	21	60.0	29	90.6	23	44.2
Mobility: Transfer - chair/wheelchair	16	100.0	12	34.3	26	81.3	23	44.2
Transfer – Toilet	16	100.0	11	31.4	27	84.4	22	42.3
Transfer – Bath tub	16	100.0	11	31.4	27	84.4	20	38.5
Walk / wheel Chair	16	100.0	11	31.4	26	81.3	20	38.5
Stairs	15	93.8	9	25.7	21	65.6	3	5.8
Cognition: Comprehension	4	25.0	19	54.3	16	50.0	40	76.9
Expression	11	68.8	23	65.7	23	71.9	40	76.9
Social interaction	15	93.8	26	74.3	21	65.6	42	80.8
Problem solving	10	62.5	15	42.9	18	56.3	43	82.7
Memory	6	37.5	13	37.1	15	46.9	43	82.7

N=135

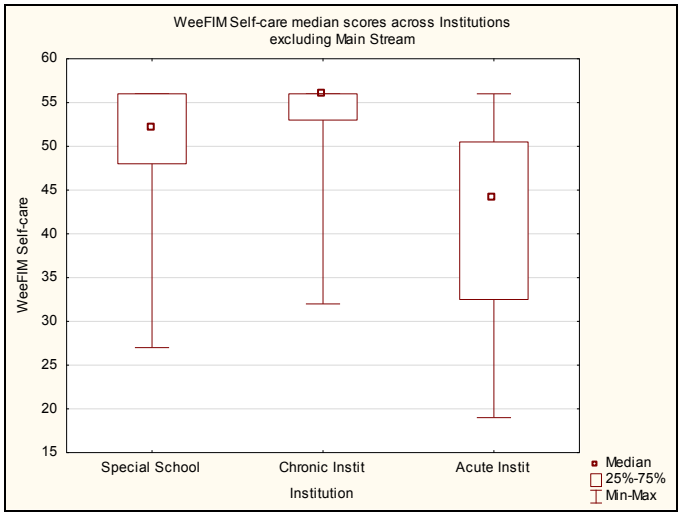
For full baseline WeeFIM assessment, see Appendix 22

The WeeFIM, as a measure of functional independence, would have demonstrated a ceiling effect in most of the MS children (all being functionally independent). Since these children would have scored maximally, they were not assessed using this measure. There were, however, 16 MS children who indicated on the EQ-5D-Y measure that they had a problem with Mobility and were subsequently assessed with the WeeFIM instrument to determine whether they did in fact have a functional problem. As can be seen from Table 42 above, they did not have a problem with Mobility on assessment with WeeFIM. Therefore, a total of 135 children were assessed with a WeeFIM measure at baseline; that is, 16 MS children and 119 children from SS, CI and AI were analysed.

4.6.1 WeeFIM dimension scores across institutions

4.6.1.1 WeeFIM Self-care across institutions

The median scores for Self-care across institutions are illustrated in Figure 24. There was a significant difference in the WeeFIM Self-care score rankings across the institutions ($p<.001$). A Kruskal Wallis ANOVA (Table 43) indicated that the AI children had significantly lower ranked Self-care scores, indicating lower independence, than the children at the SS and the CI, which were also significantly different in ranking, from each other.



N=119 Note: The higher the score, the greater the independence.

Figure 24: WeeFIM median Self-care score per institution, excluding MS

Table 43: Differences in WeeFIM Self-care ranking across Institutions

	Mean Rank	N	SS	CI	AI
SS	64.4	35		***0.045***	***0.008***
CI	84.9	32	***0.045***		***0.000***
AI	41.7	52	***0.008***	***0.000***	

N=119

Kruskal-Wallis test: H (2, N= 119) =32.95 ***p<.001.***

4.6.1.2 WeeFIM Mobility across institutions

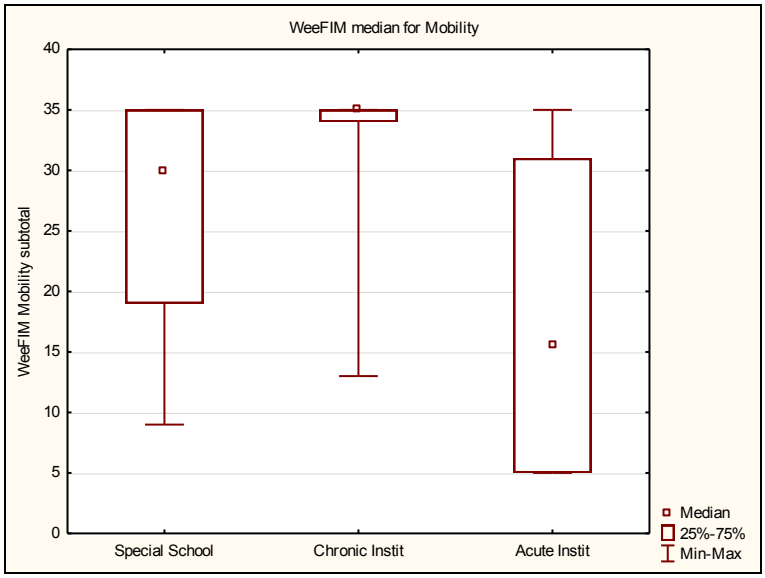


Figure 25: WeeFIM median Mobility scores per institution, excluding MS

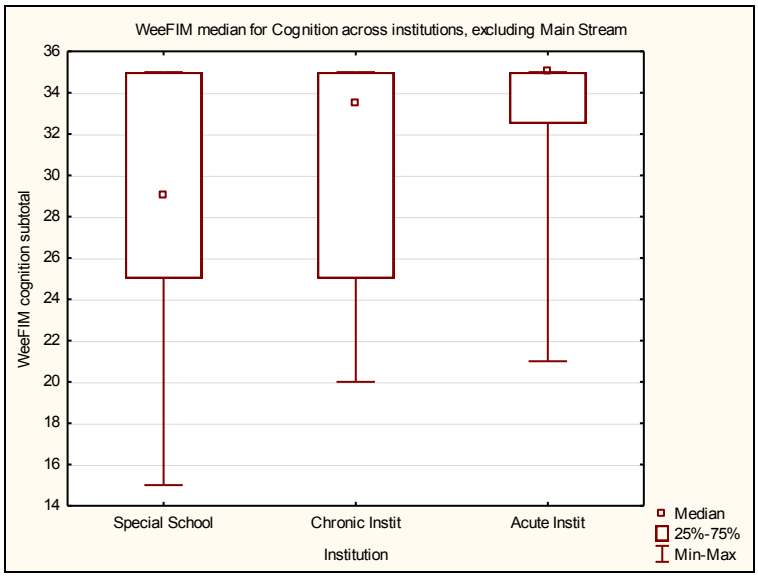
The median scores for Mobility across institutions are illustrated in Figure 25. There was a significant difference in Mobility ranking across the institutions p<.001). A Kruskal Wallis ANOVA (Table 44) revealed that the AI children had significantly lower ranked Mobility scores, indicating lower independence, than the SS and CI children, whose rankings were also significantly different from each other.

Table 44: WeeFIM Mobility ranking across institutions

	Mean rank	N	SS	CI	AI
SS	62.6	35		***0.007***	***0.011***
CI	88.3	32	***0.007***		***p<.001***
AI	40.8	52	***0.011***	***p<.001***	

Kruskal-Wallis test: H (2, N= 119) =38.8 ***p<.001***

4.6.1.3 WeeFIM Cognition across institutions



N=119

Figure 26: WeeFIM Cognition median score per institution

The median scores for Cognition across institutions are illustrated in Figure 26. There was a significant difference in ranking of cognition scores across the institutions (p<.001).). A Kruskal Wallis ANOVA (Table 45) indicates that the ranking of Cognition subtotal of the AI children was significantly higher (fewer problems) than the ranking of cognition of the SS and CI children. The ranking of Cognition scores was not significantly different between the CI and SS children.

Table 45: Differences in mean ranking of WeeFIM cognition score between the institutions

	Mean Rank	N	SS	CI	AI
SS	45.9	35		0.965	***0.001***
CI	54.2	32	0.965		***0.045***
AI	73.1	52	***0.001***	***0.045***	

N=119

Kruskal-Wallis test: H (2, N= 119) =16.60 **p<.001**

4.6.1.4 WeeFIM dimension totals across institutions

As expected, the AI participants had the most problems (lowest percentage score) with Mobility (51%) and Self-care (74%), but the fewest problems overall with Cognition (94%). The participants at the SS experienced more problems with Self-care (89%), Mobility (74%) and Cognition (83%) compared to the participants at the CI. Out of all CI children, 95%, 93% and 87% experienced problems in each of the three dimensions, respectively (Table 46).

Table 46: Weighted percentage scores for WeeFIM self-care, mobility and cognition

	Self-care subtotal %	Mobility subtotal %	Cognition subtotal %
SS (n=35)	89	74	83
CI (n=32)	95	93	87
AI (n=52)	74	51	94

4.6.2 WeeFIM total score per institution

Figure 27 demonstrates the non-normal distribution of WeeFIM total scores across institutions.

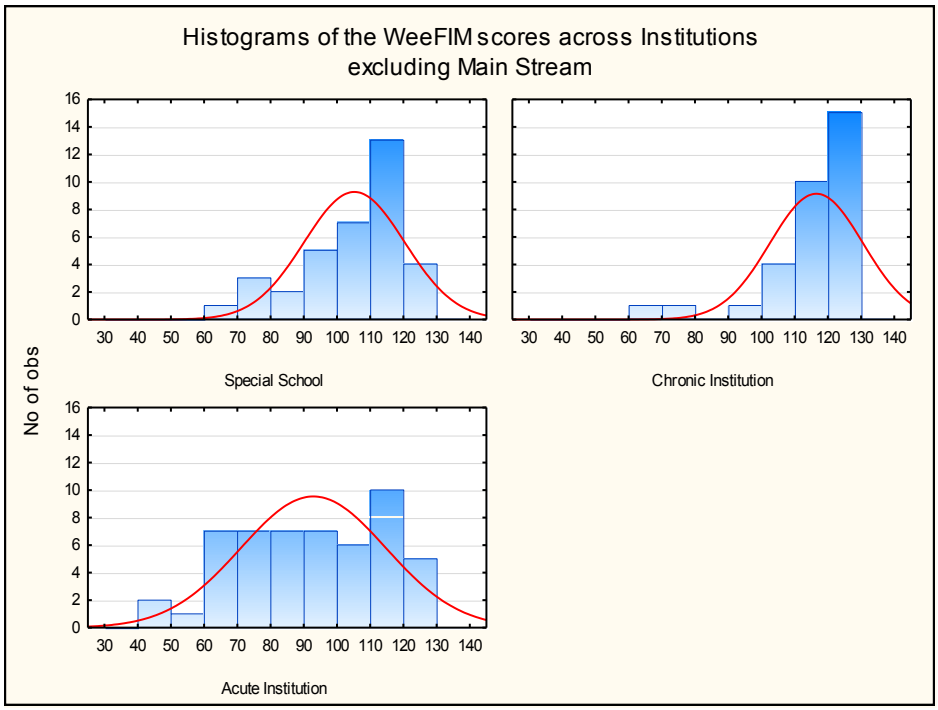
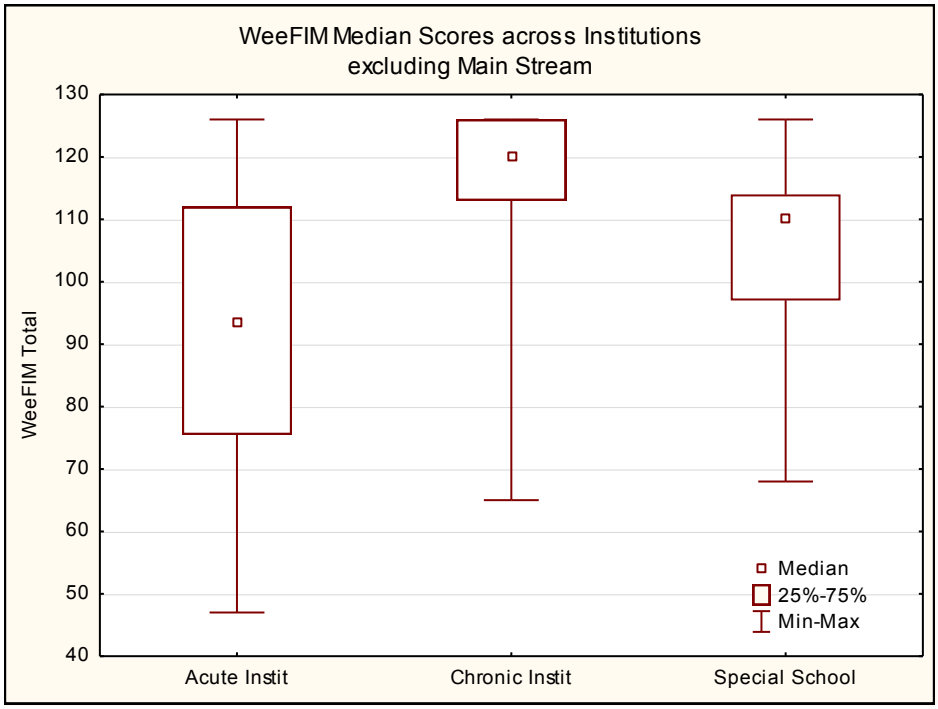


Figure 27: Distribution of WeeFIM total scores across institutions, excluding MS



N =119

Figure 28: WeeFIM median score per institution, excluding MS, at baseline

The median scores for WeeFIM total across institutions are illustrated in Figure 28. There was a significant difference in the WeeFIM total scores ranking across three institutions (p<.001), excluding MS. A Kruskal Wallis ANOVA (Table 47) revealed that the children from the CI had significantly higher ranked WeeFIM total scores (median 120, range 65-126), indicating greater functional independence than the participants at the SS (median 110, range 68-126) and the AI (median 93.5, range 47-126). However, according to the WeeFIM total scores, the AI and SS children experienced a similar degree of functional limitation.

Table 47: Differences in WeeFIM total score ranking between the institutions, excluding MS

	Mean Rank	N	SS	CI	AI
SS	60.0	35		***0.005***	0.090
CI	86.6	32	***0.005***		***p>0.001***
AI	43.6	52	0.09	***P<0.001***	

N=119

Kruskal-Wallis test: H (2, N= 119) =30.84 ***p<.001***

4.6.3 Summary of performance of WeeFIM across the different health conditions at the different institutions, indicating discriminant validity.

The inclusion criterion for using the WeeFIM was a problem with mobility, as assessed on the EQ-5D-Y Mobility dimension. Sixteen of the MS children indicated a Mobility problem. These children were then assessed for functional independence using the WeeFIM and were found to not have a problem on the Mobility dimension. However, the WeeFIM Mobility dimension did not depict a significant difference between the MS children with no mobility problems and children from the CI, some of whom did have mobility problems.

The Self-care dimension on the WeeFIM indicated a significant difference between the AI children and the other two groups only.

The Cognition dimension showed a significant difference between the acutely ill children and the SS children, but not between these and the other groups.

The WeeFIM total score depicted a significant difference in functional independence between the children at the CI and the SS and AI. However, it was unable to depict a difference in functional independence between the AI and SS, indicating poor discriminate validity in these groups.

4.7 Baseline Faces Pain Scale measure across all institutions

Using six pictures to indicate increasing pain intensity from 0-10, the FPS was measured at each institution (Table 48). A total of 222 children completed the FPS at baseline. There were missing scores from two MS children. The Kruskal-Wallis ANOVA was used to determine the difference in pain ranking across institutions (Table 49).

Table 48: Baseline Faces Pain Scale

Faces Pain Scale Score	**MS** (n=103)		**SS** (n=35)		**CI** (n=32)		**AI** (n= 52)		**Total** (n=222)	
	count	%	count	%	count	%	count	%	count	%
0	51	49.5	24	68.6	14	43.8	16	30.8	105	47.3
2	29	28.2	6	17.1	11	34.4	16	30.8	62	27.9
4	12	11.7	1	2.9	3	9.4	8	15.4	24	10.8
6	2	1.9	0	0.0	3	9.4	1	1.9	6	2.7
8	3	2.9	1	2.9	0	0.0	4	7.7	8	3.6
10	6	5.8	3	8.6	1	3.1	7	13.5	17	7.7
All Groups	103		35		32		52		222	100.0

N=222; 2 missing

Table 49: Kruskal-Wallis ANOVA by Ranks; Faces Pain Scale across institutions

	Mean Rank	**Median (range)**	**Std.Dev.**	**MS**	**SS**	**CI**	**AI**
MS (n=103)	107.70	2 (0-10)	2.76		0.817	1	0.097
SS (n=35)	88.97	0 (0-10)	3.07	0.817		0.871	***0.008***
CI (n=32)	111.86	2 (0-10	2.41	1	0.871		0.753
AI (n=52)	133.97	2 (0-10)	3.47	0.097	***0.008***	0.753	

N=222

Kruskal-Wallis test: H (N= 222) =12.67 ***p =.005***

There was a significant difference in pain ranking across the institutions. Table 49 shows that the mean ranking for pain was the highest at the AI, indicating that these children experienced the most pain, followed by the children at the CI, then the MS children. The SS children's ranking for pain was the lowest. The only significant difference in pain ranking was between the AI children (most pain) and the SS children (least pain).

4.8 Convergent validity

Convergent validity of the EQ-5D-Y was examined by correlating the dimension scores for children with different health states with their scores on similar dimensions of the PedsQL, WeeFIM and the FPS. Significant correlations between VAS and PedsQL and WeeFIM total scores, assessed at the same time, were also examined for concurrent validity.

Of the two self-reporting HRQoL measures, the EQ-5D-Y has five dimensions and the PedsQL has four. Some dimensions are similar between the two measures. Pain is included as single items in the PedsQL Activity dimension, so the "I hurt" item was compared to EQ-5D-Y P/D dimension. EQ-5D-Y LAM was compared with Self-care subtotal on the WeeFIM. As all the PedsQL dimensions contain some aspect of UA, the PedsQL total was used to compare with EQ-5D-Y UA dimension. The EQ-5D-Y P/D dimension was compared to the FPS. A summary of the dimensions that were compared is given in Table 50.

Table 50: Comparing similar dimensions of EQ-5D-Y, PedsQL, WeeFIM and Faces Pain Scale

EQ-5D-Y dimensions	PedsQL dimensions	WeeFIM	Faces Pain Scale
Mobility	Health and Activity subtotal	Mobility subtotal	
LAM		Self-care subtotal	
UA	PedsQL total		
P/D	"I hurt" item		FPS
WSU	"Feelings" subtotal		

The mean ranking of the relevant subtotal scores of the PedsQL and WeeFIM dimensions were compared across the different levels of similar EQ-5D-Y dimensions as described in Table 50.
The scores were not normally distributed and as the EQ-5D-Y dimensions have three levels of problems, the Kruskal-Wallis H test was conducted to determine if the ranking of the PedsQL or WeeFIM scores were different for the three levels of the EQ-5D-Y. When there were five or fewer scores for a particular problem level (1, 2 or 3) on the independent EQ-5D-Y variable, this level was excluded and the Mann-Whitney U Test was used to compare the remaining two levels.

Convergent validity was demonstrated when the Kruskall Wallis ANOVAs indicated that the scores between the two measures were significantly different across all the levels ($p<.05$ in every case).

4.8.1 Convergent validity of the EQ-5D-Y Mobility and PedsQL Health and Activities dimensions across institutions

In the MS children, there were no scores on level three of EQ-5D-Y Mobility dimension so this level was excluded. The Mann-Whitney U Test was used to compare the remaining two levels in these children (Table 51).

Table 51: Comparing EQ-5D-Y Mobility and PedsQL Activities dimensions

Institution	N	Kruskal-Wallis H value	p value	Mann-Whitney U z value	p value
MS	105			-1.42	0.156
SS	35	16.7	***p<0.001***		
CI	32	3.73	0.15		
AI	52	15.81	***p<0.001***		

N=224

Based on the Mann-Whitney U test, there was no significant difference in ranking of PedsQL Activity scores across two levels of the EQ-5D-Y Mobility in the MS children (Table 51). The Kruskal-Wallis test

showed that, in the children at the CI, there was no significant difference in ranking of the PedsQL Activity score across the different levels of the EQ-5D-Y Mobility dimension. However, there was a significant difference in ranking of PedsQL Activity score across the three EQ-5D-Y levels for the SS and AI children, with p values of <0.001 in both cases, thus indicating good convergent validity between the measures for these health conditions.

The institutions at which there was a significant difference in ranking (SS and AI) are shown in the Box-Whisker graph, in Appendix 23.

4.8.2 Convergent validity of the EQ-5D-Y UA and PedsQL total

All of the PedsQL dimensions contain items of Usual Activities, so the PedsQL total was compared against EQ-5D-Y AU dimension.

Table 52: Comparing EQ-5D-Y UA dimension with PedsQL total across institutions

Institution	N	Kruskal-Wallis H value	p value	Mann-Whitney U z value	p value
MS	105			-2.682	***0.007***
SS	35			-0.847	0.397
CI	32			-3.043	***0.002***
AI	52	10.335	***0.006***		

In the case in the MS children, SS and CI, level 3 (a lot of problems) on the EQ-5D-Y UA dimension was excluded as there was only one score for this level and the Mann-Whitney U Test was used to compare two levels.
The Mann-Whitney U in Table 52 indicated a significant difference in ranking of PedsQL total across two levels of the EQ-5D-Y UA dimension, for MS and CI children. The Kruskal-Wallis test showed there was a significant difference in ranking of PedsQL total across the different levels of the EQ-5D-Y UA dimension at the AI.

The institutions at which there was a significant difference in ranking (MS, CI and AI) are shown in the Box-Whisker graph, in Appendix 23.

4.8.3 Convergent validity of the EQ-5D-Y P/D dimension with PedsQL "I hurt" item across institutions

As there were only 2 cases of level 3 P/D at SS and CI and only 4 cases at the MS, this level was excluded and these were analysed using Mann-Whitney U test.

Table 53: Comparing EQ-5D-Y P/D dimension with PedsQL "I hurt" item

Institution	N	Kruskal-Wallis H value	p value	Mann-Whitney U z value	p value
MS	103			0.753	0.452
SS	35			1.678	0.093
CI	32			1.566	0.117
AI	52	26.78	***p<0.001***		

N=222

The Mann-Whitney U test indicated no significant difference in ranking of PedsQL "I hurt" across two levels of EQ-5D-Y P/D dimension the MS children, the SS and CI

The Kruskal-Wallis test indicated a significant difference in ranking of PedsQl "I hurt" across the three levels of EQ-5D-Y P/D dimension at the AI only (Table 53).

This is shown in Box-Whisker graph in Appendix 23.

4.8.4 Convergent validity of the EQ-5D-Y WSU dimension with PedsQL Feelings dimension, at each institution

In the MS children, SS and CI, level 3 on the EQ-5D- Y WSU was excluded as there were less than five reporting this level. The Mann-Whitney U Test was used to compare two levels.

Table 54: Comparing EQ-5D-Y WSU dimension and PedsQL Feelings dimension across institutions

Institution	N	Kruskal-Wallis H value	p value	Mann-Whitney U z value	p value
MS	105			-2.318	***0.020***
SS	35			-1.210	0.226
CI	32			0.872	0.383
AI	52	12.14	***0.007***		

N=224

The Mann-Whitney U test in (Table 54) indicated a significant difference in ranking of PedsQL Feelings score across two levels of the EQ-5D-Y WSU dimension for MS children, but no difference in ranking at the SS and CI.

The Kruskal-Wallis test, however, indicted a significant difference in ranking of PedsQL feelings score across all three levels of EQ-5D-Y WSU at the AI.

The institutions at which there was a significant difference in ranking (MS and AI) are shown in the Box-Whisker graphs, in Appendix 23.

4.8.5 Convergent validity of the EQ-5D-Y Mobility and WeeFIM Mobility dimensions, at each institution.

All 16 MS children scored maximally (indicating no problem with mobility and a ceiling effect) on the WeeFIM, so they were excluded from this analysis.

Table 55: Comparing EQ-5D-Y Mobility and WeeFIM Mobility dimension, across institutions

Institution	N	Kruskal-Wallis H value	p value
SS	35	22.12	***p<0.001***
CI	32	9.19	***0.01***
AI	52	21.75	***p<0.001***

N=135

The Kruskal-Wallis test in Table 55 indicated that there was a significant difference in ranking of the WeeFIM Mobility score across the three levels of the EQ-5D-Y Mobility dimension, for the SS, CI and AI. This indicated concurrent validity for the EQ-5D-Y Mobility dimension and WeeFIM mobility, for these three institutions.

This is shown in the Box-Whisker graphs, for each institution, in Appendix 23.

4.8.6 Convergent validity of the EQ-5D-Y LAM and WeeFIM Self-care dimension, at each institution

In the MS children there were no scores for level 3 (a lot of problems) on the EQ-5D-Y self-care dimension so the Mann-Whitney U Test was used to compare the other two levels.

Table 56: Comparing EQ-5D-Y LAM and WeeFIM Self-care dimensions, across institutions

Institution	N	Kruskal-Wallis H value	p value
SS	35	14.19	***p<0.001***
CI	32	8.69	***0.013***
AI	52	15.57	***p<0.001***

N=135

In MS children, the Mann-Whitney U indicated no significant difference in ranking of WeeFIM subtotals across two levels of the EQ-5D-Y for the Self-care dimension. In the children at the SS, CI and AI, the Kruskal-Wallis test showed a significant difference in ranking of WeeFIM Self-care subtotal across the different levels of the EQ-5D-Y LAM dimension (Table 56).

This is shown in the Box-Whisker graphs, in Appendix 23.

4.8.7 Convergent validity of the EQ-5D-Y P/D dimension with Faces Pain Scale, at each of the institutions

Table 57: Comparing Faces Pain Scale with the independent variable, EQ-5D-Y P/D

Institution	N	Kruskal-Wallis H value	p value	Mann-Whitney U z value	p value
MS	103			0.150	0.881
SS	35			0.788	0.431
CI	32			-0.360	0.719
AI	52	29.76	***p<0.001***		

N=222

In the MS, SS and CI children, the Mann-Whitney U test was used as there were only two levels in Table 57 indicated no significant difference in ranking of Faces Pain Scale (FPS) across two levels of EQ-5D-Y P/D dimension.

The Kruskal-Wallis test indicated a significant difference in ranking of FPS across the three levels of EQ-5D-Y P/D dimension at the AI only.

This is shown in Box-Whisker graph in Appendix 23.

4.8.8 Summary of Convergent validity between similar dimensions for all outcome measures, across institutions

The institutions at which convergent validity was evident on similar dimensions of EQ-5D-Y and the other outcome measures are illustrated in bold in Table 58.

Table 58: Convergent validity between similar dimensions

EQ-5D-Y dimensions	PedsQL dimensions and institutions	WeeFIM dimensions and institutions	Faces Pain Scale and institutions
Mobility	Health and Activity subtotal: Convergent validity evident at **SS** and **AI**	Mobility subtotal: Convergent validity evident at **SS, CI** and **AI**	
LAM		Self-care subtotal: Convergent validity evident at **SS, CI** and **AI**	
UA	PedsQL total: Convergent validity evident at **MS, Chronic** and **AIs**		
P/D	"I hurt" item: Convergent validity evident at **AI** only		FPS: Convergent validity evident at **AI** only
WSU	Feelings subtotal: Convergent validity evident at **MS** and **AI**		

It would seem that the PedsQL and EQ-5D-Y demonstrated convergent validity at the AI for all dimensions compared. The WeeFIM and EQ-5D-Y demonstrated convergent validity at SS, CI and AI for all dimensions compared. The FPS and EQ-5D-Y P/D dimension showed convergent validity at the AI only.

4.8.9 Convergent validity between the sets of total scores

Convergent validity was indicated by significant correlations between the EQ-5D-Y VAS, Index Score and total scores of PedsQL and WeeFIM ($p<.05$ in all cases) across institutions (Table 59). The highest correlation was between the PedsQL and WeeFIM (-.48), followed closely by The EQ-5D-Y and PedsQL (.47) and the lowest between the PedsQL and the VAS (-.28).

Table 59: Correlations between the different instruments total scores, all groups

	EQ-5D-Y VAS	EQ-5D-Y Index Score	PedsQL total	WeeFIM total
EQ-5D-Y VAS		-0.34	-0.28	0.39
EQ-5D-Y Index Score	-0.34		0.46	-0.59
PedQl total	-028	0.46		-0.48
WeeFIM total	0.39	-0.59	-0.48	

N=224, apart from WeeFIM where N=135
Spearman Rho is significant at ***p<.05***

Note: High scores on VAS and WeeFIM indicate high HRQoL and high scores on the PedsQL and Index indicate poor HRQoL.

Table 60: Summary of significant correlations between the different instruments per institution

Institution	Significant correlations were evident between	Rho	p
MS	EQ-5D-Y Index Score and PedsQL total	0.199	***0.041***
SS	All similar EQ-5D-Y and WeeFIM dimensions		
	EQ-5D-Y Index Score and PedsQL total	0.441	***0.009***
	PedsQL and WeeFIM total	0.557	***p<0.001***
CI	All similar EQ-5D-Y and WeeFIM dimensions		
	EQ-5D-Y VAS and PedsQL total	-0.523	***0.002***
	EQ-5D-Y index Score and WeeFIM total	-0.398	***0.024***
	PedsQL and WeeFIM total	-0.558	***p<0.001***
AI	All similar EQ-5D-Y and PedsQL dimensions		
	All similar EQ-5D-Y and WeeFIM dimensions		
	EQ-5D-Y P/D dimension and FPS		
	EQ-5D-Y VAS and Index Score	-0.767	***p<0.001***
	EQ-5D-y Index Score and PedsQL total	0.635	***p<0.001***
	EQ-5D-Y Index Score and WeeFIM total	-0.659	***p<0.001***
	EQ-5D-Y VAS and PedsQL total	-0.564	***p<0.001***
	EQ-5D-Y VAS and WeeFIM total	0.525	***p<0.001***
	PedsQL and WeeFIM total	-0.529	***p<0.001***

Table 61: Summary of discriminant ability of the outcome measures, between the different groups

Outcome Measure	Institutions between which the outcome measures discriminated significantly.
EQ-5D-Y Dimensions	Between the AI and the MS school on all dimensions.
EQ-5D-Y Index Score	Between AI all other groups Between MS and all other groups Between SS and MS and between SS and AI Between CI and MS and between CI and AI
EQ-5D-Y VAS	Between the AI and the other groups
PedsQL	Between the AI and the MS
WeeFIM	Between the CI and SS and between CI and the AI, but not between AI and SS

4.9 Repeated measures over the study period

The outcome measures were repeated at three months and six months, post baseline assessment to assess stability of scores, in the children whose HRQoL was not expected to change (MS and SS).

The responsiveness of the two HRQoL measures, the EQ-5D-Y and PedsQL, and the measure of functional independence the WeeFIM, to change over time was also assessed. As noted previously, it was anticipated that the MS children would not show a change in HRQoL over time. The SS children, with a stable chronic disability, might show no change or a slight improvement as a result of continued treatment. The children at the CI would show some improvement with better management and treatment of their chronic health condition. The AI children should show a significant improvement in HRQoL over time with treatment and pain medication for their acute condition.

Table 62: Length of time between assessment periods

	Between baseline and 2nd assessment period			Between 2nd and 3rd assessment periods			Between baseline and 3rd assessment period		
	N	Days since baseline (Mean)	Days since baseline Std.Dev.	N	Days from 2nd assessment (Mean)	Days from 2nd assessment Std.Dev.	N	Days from baseline (Mean)	Days from Baseline Std.Dev
Main Stream	98	93.4	5.52	93	87.1	3.29	95	180.7	6.14
SS	34	86.9	7.58	34	114.2	6.63	34	201.0	4.67
CI	31	88.7	5.17	26	99.3	2.99	26	186.8	3.49
AI	35	3.3	1.49	12	6.8	5.59	12	11.0	5.61

Table 62 showed that the mean length of time between baseline and the second assessment period was approximately 90 days (three months), while the time between the second and third assessment periods was approximately 100 days (just over three months) for the MS, SS and CI participants. There were approximately 190 days between baseline and third assessment (just under seven months). At the AI, the time period between baseline and second assessment was about three days and almost one week between second and third assessment periods. Only a few acutely-ill children were still in hospital for a third assessment, which was approximately 11 days post baseline.

Table 63: Number of participants with a deteriorating health condition

	Between baseline and 2nd assessment period			Between 2nd and 3rd assessment periods		
	Worsening health condition*	Indicated by	Causative factor	Worsening health condition	Indicated by	Causative factor
MS	N=0			N=0		
SS	*N=1	Decreased activity	Superimposed acute illness	N=1	Decreased activity	Superimposed Acute illness
CI	N=0			N=0		
AI	*N=4	Decreased activity	Raised temperature in 4 participants Surgical intervention since baseline in 3 participants	N=1	Decreased activity	Surgical intervention

*Excluded from data analysis

A number of participants with a deteriorating health condition were identified, as indicated in Table 63

Table 63, indicates that a total of seven participants with deteriorating health, were removed from the data base after baseline data analysis to fit in with the hypothesis described above.

4.10 The EQ-5D-Y scores over time, across institutions

4.10.1 Stability of dimension scores over time, per institution

The ability of the EQ-5D-Y dimension scores and VAS to produce consistent results between baseline and second assessment (three month interval) and between baseline and third assessment (seven month interval) whose HRQoL was not expected to change significantly, was examined for stability in the MS and SS children. Moderate agreement between scores over time, was expected, allowing for some variability in HRQoL, which would normally occur.

Agreement between these values was also examined for the Chronic and AIs (about six days apart), but was expected to be weak, as the HRQoL of these children did change over time.

Cross tabulations and Cohen's kappa coefficient were used to measure observed agreement between scores at the different time intervals. Kappa values indicating level of agreement between scores (Table 64) were interpreted using Landis and Koch labelling (110).

Table 64: Kappa values for agreement in EQ-5D-Y dimensions at all time intervals

		MS	SS	CI	AI
Kappa value for Mobility	Between baseline and 2nd assessment	0.061 Slight (n=98)	***0.516*** ***Moderate*** (n=34)	***0.248*** ***Fair*** (n=31)	0.105 Slight (n=35)
	Between 2nd and 3rd assessment	0.00 No statistics were computed because 3rd assessment Mobility was a constant (n=95)	***0.593*** ***Moderate*** (n=34)	***0.568*** ***Moderate*** (n=26)	0.***429*** ***Moderate*** (n=12)
	Between baseline and 3rd assessment	0.00 No statistics were computed because 3rd assessment Mobility was a constant (n=95)	***0.688*** ***Substantial*** (n=34)	0.133 Slight (n=26)	***0.250*** ***Fair*** (n=12)
Kappa value for LAM	Between baseline and 2nd assessment	0.125 Slight	***0.341*** ***Fair***	***0.336*** ***Fair***	0.173 Slight
	Between 2nd and 3rd assessment	-.019 Slight	***0.470*** ***Moderate***	***0.540*** ***Moderate***	***0.351*** ***Fair***
	Between baseline and third assessment	-.020 Slight	***0.421*** ***Moderate***	***0.224*** ***Fair***	-0.200 Slight
Kappa value for UA	Between baseline and 2nd assessment	0.172 Slight	0.118 Slight	***0.389*** ***Fair***	0.202 Slight
	Between 2nd and 3rd assessment	.057 Slight	0.114 Slight	***0.381*** ***Fair***	0.030 slight

	Between baseline and third assessment	-.010 Slight	***0.339 Fair***	0.177 Slight	-0.032 slight
Kappa value P/D	Between baseline and 2nd assessment	0.152 Slight	-0.12 Slight	0.069 Slight	0.186 Slight
	Between 2nd and 3rd assessment	.086 Slight	0.122 Slight	-0.014 Slight	***0.238 Fair***
	Between baseline and third assessment	.145 Slight	0.148 Slight	***0.273 Fair***	0.167 Slight
Kappa value for WSU	Between baseline and 2nd assessment	.172 Slight	***0.205 Fair***	0.082 Slight	0.111 Slight
	Between 2nd and 3rd assessment	***.279 Fair***	0.106 Slight	0.084 Slight	***0.600 Moderate***
	Between baseline and third assessment	***.235 Fair***	0.050 Slight	-0.174 Slight	0.077 Slight
Index score ICC for absolute agreement	Between baseline and 2nd assessment	0.409	0.488	0.395	0.155

Interpreted according to Landis, Koch (110)

The Kappa values for all the EQ-5D-Y dimensions for MS children indicated slight agreement, all <0.20, indicating a change despite no change in health status. This indicated poor stability (which was not as hypothesised) between baseline and three months. Over the longer seven month period, similar poor agreement between scores was evident.

At the SS, the kappa values indicated there was fair to moderate agreement for Mobility, LAM and WSU (some intra-rater reliability and stability in these dimensions, as hypothesised), but slight agreement for UA and P/D which demonstrated poor stability and changes in these dimensions. It was hypothesised that these children would show better stability in all dimensions as the state of their chronic disability did not change. Over the longer seven month interval, there was fair to substantial agreement in all dimensions expect P/D, indicating improved reliability and stability, which might indicate that over the longer period, transient variability in HRQoL was not as evident.

The kappa values for Mobility, LAM and UA at the CI all indicated fair agreement between baseline and second assessment, indicating limited reliability and, therefore, some changes in the dimensions, as expected. P/D and WSU only slightly agreed (poor reliability), indicating a change which was hypothesised. P/D and LAM were the only dimensions with fair agreement over the seven month interval. The agreement between all other dimensions was poor, as expected.

At the AI, the kappa values for all dimensions indicated slight agreement between baseline and second assessment, indicating a change had taken place, as was expected. The only difference over the seven month interval was fair agreement evident in the mobility dimension.

4.10.1.1 *Summary of stability of EQ-5D-Y Index Scores on repeated testing*

Table 65: Correlations between baseline and second assessment Index Scores across institutions

	ICC	95% Confidence Interval		P value
		Lower Bound	Upper Bound	
MS	0.409	0.124	0.603	***0.005***
SS	0.488	-0.036	0.745	***0.031***
CI	0.395	-0.169	0.697	0.069
AI	0.155	-0.478	0.543	0.285

The ICC indicated that the absolute agreement between the baseline and second assessment Index Sores was significant in the MS and SS, demonstrating no significant change (and stability of score over time). The ICC in the CI and AI was not significant indicating that there was a change (and scores were not stable) (Table 65).

4.10.2 Correlations of EQ-5D-Y VAS scores at different time intervals

Table 66: Spearman's Rho and p value for correlations in VAS at different time intervals

Institution	Between 1st and 2nd assessment			Between 2nd and 3rd assessment			Between 1st and 3rd assessment		
	N	r	P value	N	r	P value	N	r	P value
MS	98	0.235	***0.020***	95	0.282	***0.006***	95	0.346	***0.001***
SS	34	-0.118	0.507	34	0.110	0.542	34	0.516	***0.002***
CI	31	-0.183	0.361	26	0.238	0.242	26	0.201	0.324
AI	35	0.359	0.043	12	-0.174	0.681	12	-0.300	0.344

Table 66 indicated that the VAS of the MS school was significantly correlated, but weakly so, over the seven months from baseline to 3rd assessment, as was expected. The SS VAS indicated a strongly significant correlation over the longer period only; that is, from baseline to third assessment. The VAS for CI and AI was not correlated at all times, which were expected as their HRQoL did change over time.

4.10.3 Stability of VAS over baseline, 2nd assessment and 3rd assessment

The stability of the VAS over all three assessment periods was assessed for all institutions

Table 67: Friedman ANOVA showing differences between VAS over time, at all institutions

	Average Rank	Mean	Std.Dev.
Baseline VAS	1.952	89.97	17.91
2nd assessment VAS	1.957	91.99	14.49
3rd assessment VAS	2.091	94.30	12.13

ANOVA Chi Sqr. (N = 160, df = 2) = 9.581 ***p = 0.008***

Table 67 indicated that there was a significant difference in VAS scores over time (p=0.008). The differences were evident at the AI between baseline and 2nd assessment (p=0.003) and between baseline and 3rd assessment (p=0.011).

The institutions at which the VAS scores were stable, e.g. MS and SS, are shown below (Table 68 and Table 69).

Table 68: Friedman ANOVA showing differences between VAS scores at MS over time.

	Average Rank	Mean	Std.Dev.
Baseline VAS	1.95	89.97	17.91
2nd Assessment VAS	1.96	91.99	14.49
3rd assessment VAS	2.09	94.30	12.13

ANOVA Chi Sqr. (N = 93, df = 2) = 2.237 p = 0.327

The VAS scores were stable over time, at the MS School (p=0.327)

Table 69: Friedman ANOVA showing differences between VAS at SS over time

	Average Rank	Mean	Std.Dev.
Baseline VAS	1.92	95.30	10.96
2nd assessment VAS	1.95	96.82	6.23
3rd assessment VAS	2.12	98.79	3.31

ANOVA Chi Sqr. (N = 33, df = 2) = 1.922 p =0 .383

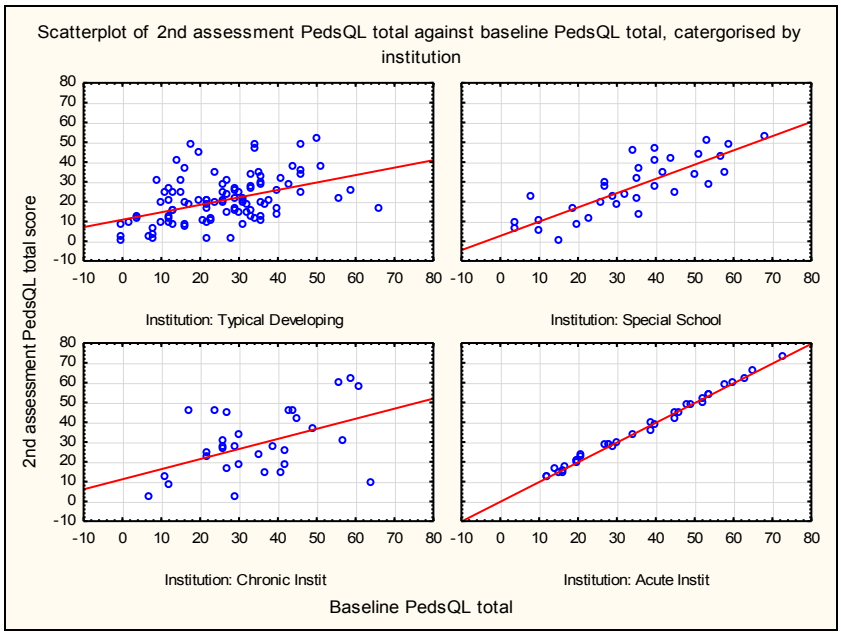
The VAS scores were stable over time, at the SS (p=0.383)

4.11 The PedsQL scores over time, across institutions

4.11.1 Stability of PedsQL scores at different time intervals

The stability of the PedsQL total score over time was examined in order to determine which measure of HRQoL, EQ-5D-Y or PedsQL showed better stability.

Correlations between PedsQL total scores were also examined for the CI and AI, but, again, it was expected to be weak because these children's HRQoL did change over time.



N=198

Figure 29: Scatter plot showing correlations between 2nd assessment PedsQL total and baseline PedsQL total; categorised by institution

Table 70: Spearman's Rho and p value for correlations between PedsQL totals at different time intervals

Institution	Between 1st and 2nd assessment			Between 2nd and 3rd assessment		
	N	r	P value	N	r	P value
MS	98	0.433	***p<0.001***	95	0.554	***p<0.001***
SS	34	0.819	***p<0.001***	34	0.747	***p<0.001***
CI	31	0.464	***0.009***	26	0.845	***p<0.001***
AI	35	0.998	***p<0.001***	12	0.968	***p<0.001***

Table 70 showed that the baseline PedsQL total scores were significantly correlated at all time intervals and across all institutions, despite that the children's HRQoL at the CI and AI changed over this time period. The strongest correlations were at the AI and SS and the weakest at the MS.

4.12 Responsiveness of the different outcome measures

The ability of the EQ-5D-Y, PedsQL and WeeFIM to depict a change in HRQoL over a three month period was examined to determine the responsiveness of each outcome measure. The EQ-5D-Y dimensions were reduced to a binary categorical variable (no problems / any problems).

Responsiveness was described by examining the effect size (r) of Wilcoxon Signed-rank test (Z) and was calculated by $(r=Z/Sq.Rt.\ N)$[2] where N is the total number of the samples; that is, the number of responses before and after, not the number of participants (Table 71).

[2] http://yatani.jp/teaching/doku.php?id=hcistats:wilcoxonsigned

Table 71: The Wilcoxon Signed rank test and effect size of each outcome measure per institution, between baseline and 2nd assessment dimension scores

Institution		No. of non-ties	T value	Z	p-value	Effect size =(Z/Sq.Rt. N)	Size
MS	Mobility	19	47.5	1.912	0.056	0.31	**Medium**
	LAM	16	48.0	1.034	0.301	0.18	Small
	UA	23	84.0	1.642	0.101	0.24	Small
	P/D	*37*	*204.0*	2.225	***0.026***	0.26	Small
	WSU	28	147.0	1.275	0.202	0.17	Small
SS	Mobility	11	30.5	0.222	0.824	0.05	Small
	LAM	*12*	*11.0*	*2.197*	***0.028***	0.45	**Medium**
	UA	15	59.5	0.028	0.977	0.01	Small
	P/D	17	72.0	0.213	0.831	0.04	Small
	WSU	14	42.0	0.659	0.510	0.12	Small
CI	Mobility	*12*	*12.0*	*2.118*	***0.034***	0.43	**Medium**
	LAM	9	15.0	0.889	0.374	0.21	Small
	UA	8	8.0	1.4	0.161	0.50	**Medium**
	P/D	14	32.5	1.256	0.209	0.24	Small
	WSU	14	39.5	0.816	0.414	0.15	Small
AI	Mobility	*21*	*45.5*	*2.433*	***0.015***	0.38	**Medium**
	LAM	*19*	*45.0*	*2.012*	***0.044***	0.33	**Medium**
	UA	17	42.5	1.609	0.108	0.28	Small
	P/D	18	48.0	1.633	0.102	0.27	Small
	WSU	*17*	*27.5*	*2.320*	***0.020***	0.40	**Medium**

Effect size interpretation - Small: 0.1, Medium: 0.3, Large: 0.5 [3]

The EQ-5D-Y dimension scores depicted a small effect size between baseline and second assessment in most dimensions at the MS school, except for mobility which was medium, but not significant. The only significant effect was in the P/D dimension.

There was a small effect size in most dimensions at the SS, as expected, as these children's chronic disabilities did not change over time (the exception being LAM dimension, which was significant).

At the CI, a small effect size was evident in most dimensions, except for Mobility which was medium and significant, as a result of improved function due to better management of the chronic condition.

At the AI, P/D and UA effect size was small, while Mobility, LAM and WSU dimensions demonstrated a significant, medium effect size due to management of the acute condition.

[3] http://yatani.jp/teaching/doku.php?id=hcistats:wilcoxonsigned

Table 72: The Wilcoxon Signed rank test and effect size based on each outcome measure's total score

Institution		No. of non-ties	T value	Z	p-value	Effect size =(Z/Sq.Rt. N)	Size
MS	VAS	48	442.0	1.497	0.134	0.15	Small
	Index Score	60	889.50	0.19	0.85	0.02	Small
	Faces Pain Scale	*53*	*403.50*	*2.76*	***0.01***	0.27	Small
	PedsQL	*94*	*1235.0*	*3.762*	***p<0.001***	0.27	Small
	WeeFim*						
SS	VAS	16	59.0	.465	0.642	0.08	Small
	Index Score	25	123.50	1.05	0.29	0.15	Small
	Faces Pain Scale	15	39.00	1.19	0.23	0.22	Small
	PedsQL	*32*	*78.0*	*3.478*	***0.001***	0.43	**Medium**
	WeeFim	*34*	*44.5*	*4.325*	***p<0.001***	0.52	**Large**
CI	VAS	13	19.5	1.817	0.069	0.36	**Medium**
	Index Score	*21*	*45.00*	*2.45*	***0.01***	**0.38**	**Medium**
	Faces Pain Scale	15	52.50	0.43	0.67	0.08	Small
	PedsQL	20	74.0	1.157	0.247	0.18	Small
	WeeFim	*15*	*18.5*	*2.357*	***0.018***	0.43	**Medium**
AI	VAS	*28*	*62.0*	*3.211*	***0.001***	0.43	**Medium**
	Index Score	*31*	*84.50*	*3.20*	***p<0.001***	0.41	**Medium**
	Faces Pain Scale	*24*	*60.50*	*2.56*	***0.01***	0.37	**Medium**
	PedsQL	20	74.0	1.157	0.247	0.18	Small
	WeeFim	*34*	*44.5*	*4.325*	***p<0.001***	0.52	**Large**

*Not tested

The EQ-5D-Y VAS detected a medium difference in the AI and CI and were therefore the most responsive to change in these children over time.

The Index Score was also was the most responsive to change in the AI and CI children, with a medium difference.

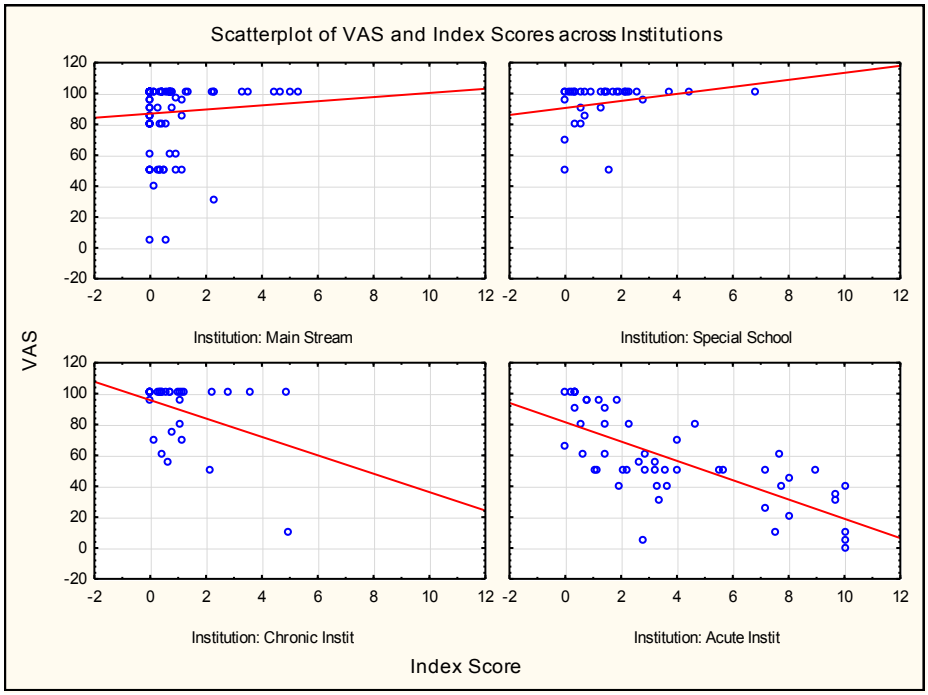
The FPS detected a significant small change in pain at the MS school and a significant medium change in the AI group.

The PedsQL only detected a medium size change in the SS children whose condition might have been expected to be more stable and not in the children whose condition was expected to change at the CI and AI.

The WeeFIM detected a medium to large improvement in all groups tested, even though the SS children had fairly stable health conditions with minimal changes. The CI children were expected to change slightly. The AI was the only institution at which a large change was expected.

4.13 Correlations between dimension scores and VAS

The dimension profiles as summarised by the Index Score were compared with the self-perceived global perception of health, VAS. There was no correlation between the VAS and Index Score or the change in VAS and the change in Index Score over time in any group apart from the AI children. (Figure 30 and Table 73).



One outlier removed from the CI.

Figure 30: Scatterplot of VAS versus Index Score

Table 73: Spearman Rank Order Correlations between VAS and Index Scores across institutions

	N	Spearman Rho	P value
MS	105	-0.047	0.638
SS	35	0.304	0.075
CI	32	-0.202	0.268
AI	52	-0.786	***p<0.001***

Similarly, when comparing the *change* in VAS scores against the *change* in Index Scores between baseline and second assessment (Figure 31 and Table 74), it was evident that the only significant correlations were, again, in the AI group.

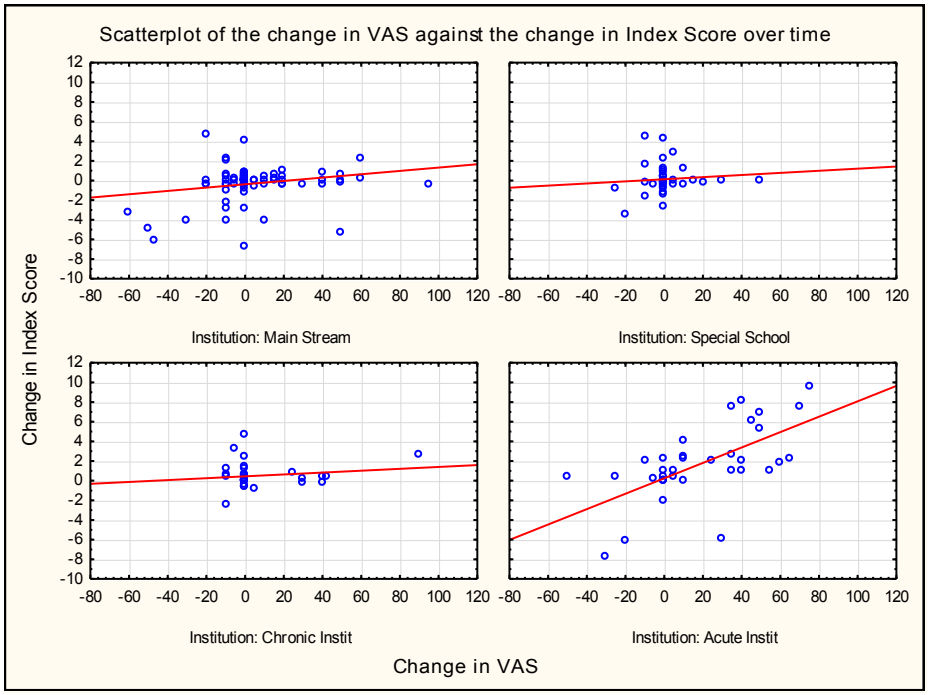


Figure 31: Scatterplot of Change in VAS versus Change in Index Score

Table 74: Spearman Rank Order Correlations between Change in VAS and Change in Index Scores across institutions

	N	Spearman Rho	P value
MS	97	0.168	0.100
SS	34	0.193	0.275
CI	31	-0.084	0.652
AI	35	0.692	***p<0.001***

Despite reporting a lot of problems in the various dimensions (observable changes - alpha changes), (Figure 8-Figure 12), the CI and SS children reported the better overall global HRQoL on the VAS (changes in internal standards of overall HRQoL - beta changes) compared to the MS children (Table 18), with no significant correlation between the scores.

4.14 Life events and changes in HRQoL

In order to establish whether changes in HRQoL are related to life events, such as any change or incident in home life, surgery or change in management of condition, a Chi-Square was used to determine whether a "yes" response to a life event had an effect on WSU dimension scores. A Mann Whitney U test was used to determine whether a life event had an effect on VAS score.

Table 75: Number of participants who experienced an additional life event between baseline and 2nd assessment

	Life event between baseline and 2nd assessment
MS	14 (14.3%), varying from a death in the family, moving house, divorce and a new baby
SS	0 reported
CI	1 (3.2%), bullying 2, changes in management of condition
AI	2 (5.7%), depressed due to no visits from family 4 (11.4%), varying from surgery, to changes in P/D medication, changes in management of condition

A total of 23 (11.6%) out of 198 children reported a life event occurring between baseline and 2nd assessment (Table 75).

Table 76: Effect of life event on EQ-5D-Y WSU dimension

	Number of "no"	**Number of "yes"**	**Chi-Square**	**df**	**P value**
2nd assessment EQ-5D-Y WSU	175 (88.4%)	23 (11.6%)	6	22	0.999

Table 76 showed a Chi-Square test comparing "yes" responses to a life event (23; 11.6%) and WSU dimension scores, at second assessment. A life event did not significantly affect the dimension score.

Table 77: Effect of life event on VAS

	Rank Sum no	**Rank Sum yes**	**U**	**Z**	**P value**
Difference in VAS between baseline and 2nd assessment	17207.0	1708.0	1455.0	1.760	0.078

As seen in Table 77, the Mann Whitney U test indicated that a life event did not have an effect on VAS either.

4.15 Feasibility and Usefulness of EQ-5D-Y outcome measure

The clinical feasibility of using the EQ-5D-Y was determined by assessing the time taken to complete the measure compared to the recommended time of five minutes. As can be seen in Table 78, it took the children approximately five minutes to complete the form.

4.15.1 Time taken to complete EQ-5D-Y measure

Table 78: Time taken in minutes to complete EQ-5D-Y

	N	Mean	Minimum	Maximum	Std.Dev.
Time taken in minutes to complete EQ-5D-Y form at baseline.	224	5.1	3	10	1.428
Time taken in minutes to complete EQ-5D-Y form at 2nd assessment.	198	4.6	3	8	0.804
Time taken in minutes to complete EQ-5D-Y form at 3rd assessment.	167	4.6	3	5	0.539

4.15.2 Therapists responses to usefulness questionnaire

The usefulness of the measure was assessed by analysing the frequency of positive responses in the questionnaire completed by the participating therapists, and assessing whether the measure was utilised by the therapists to inform management decisions and whether they would continue to use it in the future.

A summary of the participating therapists' responses is in documented in Table 79

Table 79: Responses on usefulness of EQ-5D-Y

		CI (n*=2)	**SS** (n*=3)	**AI** (n*=4)	**Agree Count** (out of 9)
	Number of self-reports administered	9		9	
	Number of proxy report completed	30	35	7	
Ease of administration	Very easy	0	2	4	6 (66.7%)
	Moderately easy	2	1		3 (33.4%)
	Difficult				0
Reason if not very easy	Time constraints	2	1		3 (33.4%)
	Child not understanding			2	2 (22.3%)
	Age group of child not understanding			8-9 years	
Dimensions which children found difficult to understand	Mobility			1	1
	LAM				
	UA			2	2
	P/D				0
	WSU				0
	VAS				0
A relationship was noticed between child's response and clinical signs in:	Mobility	1	3	4	8 (88.9%)
	LAM	1	2	2	5 (55.6%)
	UA		1		1 (11.1%)
	P/D		1	2	3 (33.3%)
	WSU	1		1	2 (22.2%)
	VAS			1	1 (11.1%)
Outcome measure assisted with planning management of child		2	3	4	9 (100%)
Most useful dimension when planning management of child	Mobility	2	1	1	4 (44.4%)
	LAM	2	1	1	4 (44.4%)
	UA	1	1	2	4 (44.4%)
	P/D	2	2	3	7 (77.8%)
	WSU	1	2	3	6 (66.7%)
	VAS	1		3	4 (44.4%)
Measure provided additional information on child's health status			3	4	7 (77.8%)
Most useful dimension in providing additional information	Mobility		1	1	2 (22.2%)
	LAM		1	1	2 (22.2%)
	UA		1		1 (11.1%)
	P/D		3	2	5 (55.6%)
	WSU		1	2	3 (33.3%)
	VAS			3	3 (33.3%)
Will continue to use measure routinely			3	3	6 (66.7%)

*n=number of clinical physiotherapists

As can be seen in Table 79, there were nine clinical physiotherapists who administered the EQ-5D-Y to 18 children and completed a total of 72 proxy reports. Six of the therapists found the measure very easy to use. The reason for three respondents finding it only moderately easy to use was mostly due to time restraints. In addition two therapists found 8-9 year- old acutely-ill children had some difficulties understanding the UA dimension. A relationship between responses and clinical signs was mostly noticed in the Mobility dimension (eight reported positively on observing this relationship), followed by P/D and WSU. All therapists found the measure useful in planning the management of the child, especially the information on P/D and WSU. Six of the therapist agreed that they would continue to use the instrument to assist the planning of management and as an outcome measure of HRQoL.

4.16 Summary of the psychometric properties of the EQ-5D-Y, PedsQL and WeeFIM as assessed on the different health profiles of the children.

Table 80: Summary of the psychometric properties of the outcome measures

Test retest Reliability: Pilot study (n=38)			
EQ-5D-Y Dimensions	**kappa**	**P value**	**Agreement interpreted according to Landis and Koch's guidelines**
Mobility	0.55	***p<0.001***	Moderate
LAM	0.65	***p<0.001***	Substantial
UA	0.12	p<0.127	Poor
P/D	0.37	p<0.08	Fair
WSU	0.55	***p<0.001***	Moderate

Agreement between **EQ-5D-Y VAS** scores	ICC=0.765	Good agreement

Main Study		**MS** (no serious health condition)	**SS** (chronic physically disabling condition)	**CI** (chronic health condition)	**AI** (acute health condition)	**All Groups**
Stability of EQ-5D-Y dimensions over a three monthly interval (three to five days at AI).	Agreement between **EQ-5D-Y dimension** scores on repeated assessment (kappa coefficient).	All dimensions (<0.20)slight	Mobility **(0.51) moderate** LAM **(0.34) fair** UA (0.12) slight P/D (-0.12) slight WSU **(0.21) fair**	Mobility **(0.25) fair** LAM **(0.35) fair** UA **(0.39) fair** P/D **(0.39) fair** WSU (0.08) slight	All dimensions (<0.20) slight	Poor at MS and slight to moderate at SS, where there was no change in health status of the participants and greater stability would have been expected.
	Correlations between **VAS** scores on repeated assessment (Spearman Rho and P).	Weak correlation significant (r=0.24) ***(p=0.020)***	Weak, NS	Weak, NS	Weak, NS	Weak correlations expected at CI and AI only.
	Correlations between **Index Scores** on repeated assessment (ICC).	Fair correlation, significant (ICC= 0.408) ***(p=0.005)***	Fair correlation, significant (ICC= 0.488) ***(p=0.031)***	Fair correlation, NS (ICC= 0.395) (p=0.069)	Poor correlation, NS (ICC= 0.155) (p=0.285)	Stability of Index Score at MS and SS, no change. AI and CI were not stable, did change.
Stability of EQ-5D-Y over a seven month interval (seven days at AI) using:	Agreement between **EQ-5D-Y dimension** scores (kappa coefficient).	Mobility (0.00) No statistics was computed because 3rd assessment Mobility was a constant LAM (-0.02) slight UA (-0.01) slight P/D (0.15) slight WSU **(0.24) fair**	Mobility **(0.69) substantial** LAM **(0.42) moderate** UA **(0.34) fair** P/D (0.15) slight WSU (0.05) slight	Mobility (0.13) slight LAM **(0.22) fair** UA (0.18) slight P/D **(0.27) fair** WSU (-0.17) slight	Mobility **(0.25) fair** LAM (-0.20) slight UA (-0.03) slight P/D (0.17) slight WSU (0.08) slight	Poor reliability and stability at MS school over seven months. Fair to substantial for physical activities, but not for P/D or WSU at SS where stability and reliability was expected in all dimensions.
	Correlations between **VAS** (Spearman's Rho and P).	(r<0.35) ***(p<0.001***) Weak correlation significant	Weak, NS	Strong, significant ***(p=0.002***)	Weak, NS	Significant correlations were expected at MS and SS, but were only evident at MS and CI.
Stability of PedsQL scores over a three monthly interval (three to five days at AI).	Correlation between **PedsQL total** scores (Spearman's Rho and P value)	Moderate, significant (r = 0.43) (***p<0.001***)	Strong, significant (r=0.82) (***p<0.001***)	Moderate, significant (r=0.46) ***(p=0.009)***	Strong, significant (r>0.99) ***(p<0.001)***	Moderate to strong correlations were not expected at Chronic or AIs
Construct validity of EQ-5D-Y by examining **discriminate validity** of:	**EQ-5D-Y dimensions** Multiple comparisons of **Mobility**	Significantly different from SS ***(p<0.001)*** and AI (***p<0.001***)	Significantly different from MS (***p<0.001***)only	Significantly different from AI (***p=0.031***) only	Significantly different from MS (***p<0.001***) and CI (***p=0.031***)	All EQ-5D-Y dimensions were only significantly different between the MS and AI

	rankings (Krushkal-Wallis H and p value).					
	Multiple comparison of ***LAM*** rankings.	Significantly different from SS (***p=0.008***) and AI (***p<0.001***)	Significantly different from MS (***p=0.008***) only.	NS different from any other group.	Significantly different from MS (***p<0.001***) only.	
	Multiple comparison of **UA** rankings.	Significantly different from AI (***p<0.001***) only.	Significantly different from AI (***p<0.001)*** only.	Significantly different from AI (***p<0.001***) only.	Significantly different from all other groups (***p<0.001***).	
	Multiple comparisons of **P/D** rankings.	Significantly different from AI (***p<0.001***) only.	NS different from any other group.	NS different from any other group.	Significantly different from MS (***p<0.001***) only.	
	Multiple comparison of **WSU** rankings.	Significantly different from AI (***p<0.001***) only.	NS different from any other group.	NS different from any other group.	Significantly different from MS (***p<0.001***) only.	
	EQ-5D-Y Index Score (Krushkal-Wallis H and p value).	Significantly different from SS ***(p<0.001)*** CI ***(p=0.045)*** AI ***(p<0.001)***	Significantly different from MS ***(p<0.001)*** AI **(p=0*.030)***	Significantly different from MS ***(p=0.045)*** AI ***(p<0.001)***	Significantly different from MS ***(p<0.001)*** SS **(p=0*.030)*** CI ***(p<0.001)***	MS and AI scored significantly differently from the other three groups, there was no difference between the SS and CI.
	EQ-5D-Y VAS (Krushkal-Wallis H and p value)	NS difference from any other group.	NS difference from any other group.	NS difference from any other group.	Significantly different from other three groups ***(p<0.001***)	The EQ-5D-Y VAS only discriminated between the acutely ill children and the other groups, but not between the other groups.
Discriminant validity of:	**PedsQL total** (median – range) (mean ranking) MS	 27 (0-66) 94.41	 NS difference from any other group.	 NS difference from any other group.	 Significantly different from MS (0***.003)*** only	The PedsQL only discriminated between the MS and AI.
	SS	NS difference from any other group.	35 (4-68) 126.19	NS difference from any other group.	NS difference from any other group.	
	CI	NS difference from any other group	NS difference from any other group.	30(7-64) 124.73	NS difference from any other group.	
	AI	Significantly different from AI (0***.003)*** only	NS difference from any other group.	NS difference from any other group.	39 (9-73) 132.29	
Discriminant validity of: MS	**WeeFIM**	Not tested.				WeeFIM total score was able to discriminate between the children at the CI, SS and AI. However it was unable to depict a difference in functional

						independence between the AI and SS, indicating poor discriminate validity in these groups.
	SS		110(68-126) 61.94	Significantly different from CI (***p=0.005)***	NS difference from any other group	
	CI		Significantly different from SS ***(p=0.005)***	120(65-126) 94.17	Significantly different from AI (***p>0.001)***	
	AI		NS difference from any other group	Significantly different from CI (***p<0.001)***	93.5(47-126) 45.24	
Construct validity continued - by examining **convergent validity** between similar dimensions of EQ-5D-Y and PedsQL.	**EQ-5D-Y Mobility** with **PedsQL Activity** dimension (Kruskal-Wallis and p value if all three levels of problems) (Mann-Whitney U if only two levels of problems).	NS, poor convergent validity	Significant difference in PedsQL Activity score across all levels of EQ-5D-Y mobility score. Good convergent validity ***(p<0.001)***	NS, poor convergent validity	Significant difference in ranking of PedsQL Activity score across all EQ-5D-Y Mobility levels ***(p<0.001)***	Convergent validity between EQ-5D-Y and PedsQL was evident at the AI only for all similar dimensions.
	EQ-5D-Y UA and **PedsQL total**	Significant difference in ranking of PedsQL total across different levels of the EQ-5D-Y Usual Activity dimension compared to CI ***(p=0.002)*** and AI ***(p=0.006)*** Good convergent validity.	NS, poor convergent validity.	Significant difference in ranking of PedsQL total across different levels of the EQ-5D-Y Usual Activity dimension compared to MS ***(p=0.007)*** and AI ***(p=0.006)*** Good convergent validity.	Significant difference in ranking of PedsQL total across different levels of the EQ-5D-Y Usual Activity dimension compared to MS ***(p=0.007)*** and CI ***(p=0.002)*** Good convergent validity.	
	EQ-5D-Y P/D dimension with **PedsQL "I hurt"** item	NS, poor convergent validity.	NS, poor convergent validity.	NS, poor convergent validity.	Significant difference in ranking of PedsQl "I hurt" across the three levels of EQ-5D-Y P/D ***(p=0.001)*** Good convergent validity.	
	EQ-5D-Y WSU dimension with **PedsQL Feelings** dimension	Significant difference in ranking of PedsQL Feelings score across two levels of	NS, poor convergent validity.	NS, poor convergent validity.	Significant difference in ranking of PedsQL feelings score across all three levels of	

		the EQ-5D-Y WSU ***(p=0.020***) Good convergent validity			EQ-5D-Y WSU ***(p=0.007)*** Good convergent validity	
Convergent validity between similar dimensions of EQ-5D-Y and WeeFIM.	**EQ-5D-Y Mobility** and **WeeFIM Mobility** dimensions.	Not tested. All scored maximally.	Significant difference in ranking of the WeeFIM Mobility across the three levels of the EQ-5D-Y Mobility ***(p<0.001)*** Good convergent validity.	Significant difference in ranking of the WeeFIM Mobility across the three levels of the EQ-5D-Y Mobility ***(p=0.01)*** Good convergent validity.	Significant difference in ranking of the WeeFIM Mobility across the three levels of the EQ-5D-Y Mobility ***(p<0.001)*** Good convergent validity.	Convergent validity between EQ-5D-Y and WeeFIM dimensions was evident at the SS, CI and AI.
	EQ-5D-Y LAM and **WeeFIM Self-care** dimension.	Not tested.	Significant difference in ranking ***(p<0.001)*** Good convergent validity.	Significant difference in ranking ***(p=0.001)*** Good convergent validity.	Significant difference in ranking ***(p<0.001)*** Good convergent validity.	
Convergent validity between **EQ-5D-Y P/D** dimension and **Faces Pain Scale**		NS, poor convergent validity.	NS, poor convergent validity.	NS, poor convergent validity.	Significant difference in ranking of FPS across the three levels of EQ-5D-Y P/D ***(p<0.001)*** Good convergent validity	The EQ-5D-Y P/D dimension and Faces Pain Scale showed convergent validity at the AI only.
Correlations between **VAS** and **PedsQL total**	**EQ-5D-Y VAS** with **PedsQL total** (Spearman Rho) (p value)	NS, poor correlation (0 .156)	NS, poor correlation (0.241)	Significant, moderate correlation (-0.523) ***(0.002)***	Significant, moderate correlation (-0.564) ***(p<0.001)***	The highest correlation was between the PedsQL and WeeFIM (-0.48) and the lowestbetween the PedsQL and the VAS (-0.28).
Correlations between **VAS** and **WeeFIM total** scores	**EQ-5D-Y VAS** with **WeeFIM total** (Spearman Rho) (p value)	NS, poor correlation (0.188)	NS, poor correlation (-0.194)	NS, poor correlation (0.064)	Significant moderate correlations (0.525) ***(p<0.001)***	
Correlations between **PedsQL** and **WeeFIM total** scores	**PedsQL total** with **WeeFIM total** (Spearman Rho) (p value)	NS, fair correlation (0.353)	Significant, moderate correlation (0.577) ***(p=<0.001)***	Significant moderate correlation (-0.558) ***(p<0.001)***	Significant, moderate correlation (-0.529) ***(p<0.001)***	
Correlations between **EQ-5D-Y VAS** and **Index Score**	**EQ-5D-Y VAS** and **Index Score** (Spearman Rho) (p value)	NS, poor correlation (0.0548)	NS, fair correlation (0.316)	NS, poor correlation (0.078)	Significant, strong correlations (-0.767) ***(p<0.001)***	

Correlations between **EQ-5D-Y Index Score** and **PedsQL total**	**EQ-5D-Y Index Score** and **PedsQL total** (Spearman Rho) (p value)	Significant, poor correlation (0.199) ***(p=0.041)***	Significant, fair correlation (0.441) ***(p=0.009)***	Significant, fair correlation (0.443) (p=0.011)	Significant, moderate correlations (0.635) ***(p<0.001)***	
Correlations between **EQ-5D-Y Index Score** and **WeeFIM total**	**EQ-5D-Y Index Score** and **WeeFIM total** (Spearman Rho) (p value)	NS, poor correlation (0.047)	NS, poor correlation (-0.278)	Significant, fair correlation (-0.398) ***(p=0.024)***	Significant, moderate correlations (-0.659) ***(p<0.001)***	
Responsiveness to change, over time, as indicated by effect size (r=Z/Sq.Rt. N) (Z=Wilcoxon Signed rank) when using the:	**EQ-5D-Y dimensions** **Mobility** (p-value) effect size	(p=0.056) **0.31 Medium effect size**	(p=0.824) 0.05 Small effect size	***(p=0.034)*** **0.34 Large effect size**	***(p=0.015)*** **0.38 Medium effect size**	Using the dimensions the most responsiveness was evident at the AI, with a medium effect size in all dimensions, except UA and P/D.
	LAM	(p=0.301) 0.18 Small effect size	***(p=0.028)*** **0.45 Medium effect size**	(p=0.347) 0.21 Small effect size	***(p=0.044)*** **0.33 Medium effect size**	
	UA	(p=0.101) 0.24 Small effect size	(p=0.977) 0.01 Small effect size	(p=0.161) 0.50 Medium effect size	(p=0.108) 0.28 Small effect size	
	P/D	***(p=0.026)*** 0.26 Small effect size	(p=0.831) 0.04 Small effect size	(p=0.209) 0.24 Small effect size	(p=0.102) 0.27 Small effect size	
	WSU	(p=0.202) 0.17 Small effect size	(p=0.510) 0.12 Small effect size	(p=0.414) 0.15 Small effect size	**(p=0.020) 0.40 Medium effect size**	
	EQ-5D-Y Index Score	(p=0.85) 0.02 Small effect size	(p=0.29) 0.15 Small effect size	***(p=0.01)*** **0.38 Medium effect size**	**(p<0.001) 0.41 Medium effect size**	
	EQ-5D-Y VAS	(p=0.134) 0.15 Small effect size	(p=0.642) 0.08 Small effect size	(p=0.069) 0.36 Medium effect size	***(p=0.001)*** **0.43 Medium effect size**	The EQ-5D-Y VAS detected a medium difference in the Acute and CIs and were therefore the most responsive to change over time, in these children.
	PedsQL total	***(p<0.001)*** 0.27 Small effect size	***(p=.001)*** 0.43 Medium effect size	(p=0.247) 0.18 Small effect size	(p=0.001) 0.37 Medium effect size	The PedsQL only detected a medium size change in the SS children whose condition might have been expected to be more stable.

	WeeFIM total	Not tested	***(p<0.001)*** **0.52 Large effect size**	***(p=0.018)*** 0.43 Medium effect size	(***p<0.001***) 0.**52 Large effect size**	The WeeFIM detected a medium to large improvement in all groups tested, even though the children at the SS had fairly stable health conditions with minimal changes; the children at the CI were expected to change slightly and the AI was the only institution at which a large change was expected.
	FPS	***p=0.01*** 0.27 Small effect size	p=0.23 0.22 Small effect size	p=0.67 0.08 Small effect size	***p=0.01*** 0.37 Medium effect size	The FPS detected a significant small change in P/D at the MS school and a significant medium change in the AI group.
	Agreement between **Proxy reports** and **child self-report dimensions** (kappa)	Mobility, (0.000), unable to compute, due to no variance LAM (-0.05) slight agreement UA **(.36) fair agreement** P/D **(.22) fair agreement** (1 - 0.21) WSU (0.07) slight agreement	Mobility **(0.55) moderate agreement** LAM (0.20) slight agreement UA (0.08) slight agreement P/D (0.16) slight agreement WSU (0.000), unable to compute, due to no variance	Mobility, **(0.84) Good agreement** LAM **(0.51) Moderate agreement** UA Kappa **(0.42) moderate agreement** P/D, (0.08) slight agreement WSU (0.01) slight agreement	Mobility **(0.73) Substantial agreement** LAM (0.06) slight agreement Usual Activity **(0.41) moderate agreement** P/D **(0.46) Moderate agreement** WSU, P/D, **(0.35) fair agreement**	There was moderate to good agreement for the Mobility dimension at all institutions, but each institution only demonstrated agreement in two to three dimensions
	Agreement between Proxy reports and child self-report **VAS** (Spearman's Rho and p value)	Weak, but significant (r=0.297) (p=0.016)	Weak, NS	Weak, NS	Moderate, NS	Even though the proxy and self-report VAS scores demonstrated acceptable ICC overall 0.58, at an institutional level, there was only a significant correlation in the MS children.

NS = not significant

4.17 Summary of performance of EQ-5D-Y when used on children with different health states

In the sample of children ranging from TD at a MS school, chronically disabled, chronically ill, and acutely ill children, the EQ-5D-Y performed as follows:

4.17.1 Performance of the EQ-5D-Y dimensions

On a dimension level the various groups of children were able to identify their level of problem. The percentage and level of problems reported in each dimension was associated with the institution. The MS children had the least problems on all dimensions, except for the WSU dimension. As expected the acutely ill children reported the most problems in level 3 (a lot of problems) for all dimensions.

The stability of the dimension scores, over a period was poor at MS and SS where better stability was expected as their health status did not change over time. Over a seven-month period there was better stability of MS dimension scores, except for WSU dimension.

Discriminant validity was evident between only the AI and the MS school on all dimensions.

Convergent validity between all similar dimensions on EQ-5D-Y and PedsQL was evident at the AI only, while convergent validity between EQ-5D-Y and WeeFIM dimensions was evident at the SS, CI and AI. The EQ-5D-Y P/D dimension and Faces Pain Scale showed convergent validity at the AI only.

The EQ-5D-Y dimensions were the most responsiveness to change at the AI, with a medium effect size in three dimensions, except for UA and P/D.

4.17.2 Summary of EQ-5D-Y Index Score

Using the composite Index Score for dimensions, stability of these scores was evident at MS and SS groups, as expected.

The Index Score was able to discriminate between MS and AI and the other three groups, but not between SS and CI.

The EQ-5D-Y Index Score was the most responsive in the AI group, but also significantly responsive in the CI group.

4.17.3 Performance of the EQ-5D-Y VAS

The stability of the VAS scores was variable at the institutions which were expected to show better agreement between scores over time (MS and SS). Over the shorter three month period reliability in VAS scores was evident at the MS school as expected, but to a lesser extent over the seven month period. The VAS scores indicated some stability at the SS over the seven month period only and not over the shorter three month period. It is possible that these children based their overall HRQoL on different aspects of well-being at the different time intervals.

When comparing the VAS score against the ranking of different levels of the dimensions, across institutions, it was found to be significant on all dimensions at the AI only. The VAS was able to discriminate only between the AI and the other three groups, but not between the other three groups. Despite having problems in the various dimensions (alpha changes), the children with a chronic health condition or disability, scored similar levels on the VAS (beta changes) for overall HRQoL compared to the MS children who had minimal or no problems on the dimensions. This is possible evidence that the children with a chronic condition or disability and they do not equate functional problems with a lowered overall HRQoL. Instead, they report having a high overall HRQoL on the VAS. As a result VAS scores should be used with caution as an outcome measure in children with chronic health conditions or disability.

There were significant correlations between EQ-5D-Y VAS and PedsQL total at CI and AI and between EQ-5D-Y VAS and WeeFIM total at AI only. The EQ-5D-Y VAS and Index Score were significantly correlated at AI only.

The VAS demonstrated responsiveness in that the effect size was medium in the AI and CI groups, with significantly different scores in the AI, where the greatest improvement was expected.

4.17.4 Relationship between EQ-5D-Y proxy and self-report

There were mostly slight correlations (low kappa) between proxies and self-report, at the individual dimension levels, except for the Mobility dimension which was higher. The proxy reports of children with a health condition, generally reported more problems on the other functional dimensions (Self-care and UA) and fewer problems on the less obviously observed dimensions (P/D, WSU).

However the intra-class correlation between proxy and self-report VAS was acceptable, indicating better agreement of overall HRQoL.

The proxy and self-report should therefore not be used interchangeably, on an individual dimension level due to discrepancies between proxy and self-report scores.

4.17.5 Feasibility and usefulness of the EQ-5D-Y

The EQ-5D-Y took only five minutes to complete. The majority of the therapists (six out of nine) reported that they found the EQ-5D-Y easy to administer, that it provided additional information on the child's status that was not obtained in routine assessments and that they would continue to use it routinely. The therapists felt that the extra information gained from this measure could be used when planning a rehabilitation programme for the children.

5 CHAPTER 5: Discussion

5.1 Introduction

The main finding of the research was that the EQ-5D-Y performed better when used in children with an acute health condition than when used in other groups of children (that is, children with no health conditions and children with chronic health conditions or chronic disabilities). We therefore recommend that it should be used with caution in these groups. The EQ-5D-Y displayed test-retest reliability in all groups, but poor stability using the dimensions scores and VAS, at the institutions which were expected to show stability in scores over time; that is at the MS and SS. The Index Scores derived from a summary score of the dimensions on a QALY scale, however, did display intra-rater reliability and stability, at the MS and SS. The EQ-5D-Y could discriminate between acutely ill children and healthy typically developing children, and demonstrated convergent validity with the other outcome measures (PedsQL, WeeFIM and FPS), when used on acutely ill children. It displayed the most responsiveness in the acutely ill children, with a medium treatment effect size in three dimensions and in VAS. It generally performed less satisfactorily in the other groups of children. These specific findings are discussed in detail, below, as they pertain to each research objective.

The sample will first be discussed in order to determine the generalizability of the findings. The different groups of children's responses to the EQ-5D-Y and the performance of the measure compared to other outcome measures will form the basis of the discussion. Comparisons will be made with other similar samples found in the literature. The psychometric properties of the EQ-5D-Y and the other outcome measures, as they pertain to the different groups of children will be highlighted, and recommendations will be made as to which instrument might be preferable in different contexts. As there was evidence of a mismatch between the functional dimensions of the EQ-5D-Y and the global perception of HRQoL (as measured by the VAS), possible reasons for this will be elaborated on. The study limitations are highlighted and the methodology discussed. Finally, the conclusions and recommendations for future practice and research are presented.

5.2 Sample

The sample of 347 children was specifically recruited to represent South African children across a diverse range of health conditions, ranging from severe acute conditions (54), through a variety of disabling conditions (91) to transient, relatively minor and/or no ailments (201). In comparison previous studies have mostly described the use of the EQ-5D-Y as an outcome measure among TD school children without a health condition. Examples using the EQ-5D-Y with large samples of children, include studies in Spain with 620 children (90), 521 South African children (156), 3421 Canadian school children (157), 260 Swedish children (158), and a United Kingdom study with 160 children (122). Some studies have assessed HRQoL using the EQ-5D-Y on children with one chronic condition only. These studies assessed 96 children with Cystic Fibrosis (132); 310 children with eczema (159); 126 children with diabetes mellitus (78); 196 children with musculoskeletal deformities (160) and 450 children with mental health problems (161). A few studies have compared HRQoL between TD children and children with one health condition. These studies compared TD children and children with long standing chronic disabilities (15); children with and without Celiac Disease (162); and heathy children and children diagnosed with Acute Lymphoblastic Leukaemia (116). A study in 2013 assessed the feasibility and validity of assessing HRQoL using the EQ-5D-Y in 71 children and adolescents with a variety of chronic health conditions and 407 TD children from the general public (163). This study compared the performance of the EQ-5D-Y across a range of health states including chronic and acute health conditions.

One of the strengths of this study was that our sample size was large enough to allow for the comparison of the performance of the EQ-5D-Y across children with very different experiences of health conditions. The early validation study of the EQ-5D-Y (1)(2010) was done with large groups (between 258 and 756) of TD children from five different countries, who had few health conditions and consequently might have had a small variance in their responses. This could have inflated the reliability and validity as there would have been large numbers reporting no problems on different dimensions. In contrast, the studies using the EQ-5D-Y in children with one specific health conditions generally had smaller sample sizes (between 96 and 310) and even smaller samples when comparing TD with one health condition (between 25 and 103). In addition, only a few of these studies (122)(90)(116)(122)(163) explicitly examine the psychometric properties of reliability and validity of the EQ-5D-Y. Even though they comment on the need to examine responsiveness of the measure in longitudinal studies, no paper has as yet reported this.

As has been reported in a similar paediatric population (164), several of the MS children in this study did not return the informed consent forms signed by parents and that might have resulted in a bias towards children from generally more organized families, possibly with a more stable socio-economic status. This was not explored but should be kept in mind. The attrition rate was acceptable in the MS, CI and SS children. It would have been desirable to follow up the AI children after discharge, but as the institution was a central hospital serving a large geographical catchment area, this would have been logistically very difficult.

The gender distribution was balanced, although there were more males with health conditions, a common finding particularly in children with disabilities (15)(116)(165). Studies have shown that in older children, females generally report worse HQRoL than males (166)(114)(157) (166). However, the difference in responses between the genders was not explored in this study as it was felt that the children were still of an age where variances were not anticipated.

The age range for which the EQ-5D-Y is specifically recommended was represented in the study and the children from each institution were similar in terms of age. As could be expected, the children in the MS school were enrolled at the appropriate grade level for age whereas children in the CI were enrolled in lower grades for their ages. Although schooling is provided in the CI, the long-term nature of the health conditions might have had a negative impact on the academic achievement of these children. This might need to be taken into account if studies in the lower age group were to be undertaken.

5.3 Responses to the EQ-5D-Y

5.3.1 Typically developing children at a MS school

The majority of published studies that include TD children report the most problems in the WSU dimension, followed by P/D , with the least problems in LAM, followed by Mobility and UA (1)(158)(90)(156). This pattern was not evident in our study. The children attending a MS school reported the most problems in the P/D dimension, followed by UA, with the least problems in LAM, Mobility, and followed by WSU. It is difficult to interpret these findings in the absence of additional information regarding the reasons for their self-reported ratings. However, the low income area in which these children live, has limited recreational facilities and the area is often affected by violence/crime. Parents tend to restrict children from playing outside (UA) when gang violence escalates. The children seem to view this restriction as a problem with UA, but do not seem to perceive the dangers as affecting their psychological QoL (WSU).

Despite the relatively small number of MS children reporting anxiety or depression (16.2%), compared to approximately 40% in the children with a health condition, there is still cause for concern among what is supposed to be TD children. This may be an indication that these children are very aware of the conditions in which they live and that the situation prevents them from performing UA, which causes anxiety.

It is unclear why so many MS children reported experiencing some problems with Mobility (15.3%) and P/D (24.8%). A possible explanation for this finding is in the interpretation of the phrase "walking about" in the EQ-5D-Y. Even though it is a *health-related* QoL measure, many of the children interpreted this to mean "the freedom to walk around" without environmental hindrances, and not whether they had physical limitations due to a health condition. This would apply to this group of children who might have difficulty walking about their community because of the crime.

A study on anxiety in South African children by Muris e al (2008) (167)highlighted several contributing factors which still exist in post-apartheid South Africa. A difference in living conditions persists between the different population groups, with the majority of black and coloured populations living in poorer areas. All the children in our study came from these poorer communities and were exposed to crime, violence, gangs, weapons, drugs and rape, which as Muris discussed were possible causes of increased anxiety in children, which could have affected some aspects of their QoL.

A study examining correlations between pain, function and HRQoL found poor correlations between a specific pain outcome measure and HRQoL measures (EQ-5D and SF-36), with increased responsiveness to a change in pain on the more specific pain measure (168). These authors felt that the different measures were measuring different constructs of pain. This might be the case in our study, as the MS children did not report high levels of pain when reporting on the FPS, which specifically assesses the level of pain. The P/D being reported on in the EQ-5D-Y may be related to relatively minor and transient pain in these children.

There was a considerable ceiling effect in dimension scores in the MS children, with over half reporting no problems on any dimension, as was found in other studies (116) (122) (132) (90). This finding emphasizes that the EQ-5D-Y should be used with caution in children who have no health condition. The original reliability and validity study of the EQ-5D-Y (1) acknowledged the high ceiling effects in all dimensions for TD children and that the measure may be limited in detecting moderate impairments or discriminating between groups of healthy children.

It is interesting to note that when comparing the performance of the adult and youth versions of the EQ-5D-Y on the same participants, the youth version (EQ-5D-Y) lowered the ceiling effect, by the use of slightly different wording (169) (156) and generally performed better in the general population.

The VAS scores, an indication of overall HRQoL, were not the highest (indicating better HRQoL), for the MS children. These children scored their overall HRQoL lower than the SS and CI children, but not significantly so. This was a similar finding to Jelsma and Ramma (15) who found that able bodied children scored slightly lower on the VAS compare to children with disabilities at a SS.

The Index Score, however, representing a summary dimension score on a QALY scale, indicated minimal problems on the various dimensions for MS children, which was significantly lower than the Index Score of the other groups.

The recent valuation study by Craig et al (2015) (66) raises the question as to who should value the children's health states, the parents or the child. It has been found that adults do not assign the same

value to a particular health state in adults, as they do to the same health state in children (170). An earlier Canadian study, conducted in 2014, was the first study to use the EQ-5D-Y to value health states in children, using approximately 6 800 children's own ratings of health (171). Positive evidence showed that 10 to 11-year-old children can assess their own health status for valuation studies. These studies indicate an interest in valuating children's health for research and economic purposes.

5.3.2 Children with health conditions

The SS children reported the most problems in the Mobility and LAM dimensions; the CI children in Mobility, and AI children in the UA and Mobility dimensions. The children with a chronic condition or disability (CI and SS) reported the least problems in the P/D dimension, which was to be expected as the children were still relatively young and may not yet be experiencing pain related to poor posture, severe contractures or condition related pain. The responses of the children with either a chronic or acute condition seemed to be consistent with their health conditions and therefore, the EQ-5D-Y dimensions performed adequately in these children.

Similarly, a Swedish study (158) exploring self-reported dimension ratings across groups of children known to differ in health status, also reported an increased percentage of problems on all dimensions, in children with a disabling handicap. In particular, these children reported more problems on the Mobility and P/D dimensions compared to the healthy children. The increased reporting on the P/D dimension in these children was in contrast to our study.

Despite their physical limitations, over half of the CI children reported having no problems with LAM, as they were generally all encouraged by caregivers and therapist at the institution to be as independent with self-care as possible and had devised methods of coping in this dimension. Many of the CI and SS children also reported no difficulty with doing their UA. It seems that, as long as they were able to socialise with their peers, they scored "no problems" in this dimension and did not perceive their physical limitations as an obstacle. This could indicate reconceptualization of their concept of UA and even though it may not be the same as the MS children's concept for this dimension, the CI and SS children did not perceive themselves as having problems with UA. This dimension is affected in both MS group and chronic health condition group, by their environment. As mentioned earlier the MS children are limited in performing their UA by the unsafe environment in which they live, while the CI and SS groups live in a sheltered institutional environment which is able to facilitate socialising and doing UA. The factors contributing to the scores in this dimension are worth exploring in future studies.

A Dutch HRQoL study (172) also found that children with muscular dystrophy responded differently to healthy children when rating UA, as a result of different experiences and conceptions.

In contrast to the high ceiling effect observed in the MS group, only 12% of the children with a health conditions reported no problems on any dimension but, as they were all specifically included in the study due to their having a health condition, this still seems quite a high number to report no impact on HRQoL.

Ceiling effects have also been observed in other studies using the EQ-5D-Y on TD children (1)(116)(157), children with cystic fibrosis (132) or asthma (67), and children with musculoskeletal problems (160), with a high percentage of "no problems" on dimensions being reported. This finding may be a reflection of the generic nature of the EQ-5D-Y, which may not be responsive to an adequate range of health states. Some of the authors mentioned above have suggested increasing the number of possible response choices on dimensions from three to five, as has been done with adult EQ-5D, to reduce the ceiling effect.

The ceiling effect with regard to the VAS was most obvious in MS, SS and CI and contributed to the skewness of the distribution of data in our study. A study by Canaway and Frew (2013)(122), found that the large percentage of respondents reporting “no problems” on all EQ-5D-Y dimensions or a median VAS of 100 as was the case in our study, presented a ceiling effect, resulting in a clustering of values at one end of the distribution curve.

The median VAS was 100 in all but the AI, which would indicate that either the children do not understand the VAS or that functional limitations as result of a health condition do not influence HRQoL as much as one would imagine. As no correlation was found between age and VAS, and no pattern was evident with regard to the ages of the children scoring 100, the latter explanation is more likely. This is discussed below in section 5.8.

The Swedish study (158) found that children with a health condition reported a lower VAS compared to healthy children, which is in contrast to this study.

The AI group scored significantly lower Index Scores, than the other groups, indicating that these children experienced the most problems in the various dimensions, as was expected.

When comparing the EQ-5D-Y dimension scores to the other outcome measure, the EQ-5D-Y dimensions and WeeFIM indicated that the MS children had the least problems, followed by the SS and CI and that the AI children had the most problems. However, this was not the case in the PedsQl dimension scores, except for the Activity dimension. Interestingly however, when comparing PedsQL and WeeFIM total scores, the same pattern emerges with the MS children experiencing the least problems, followed by CI and SS children and lastly AI children with the most problems, indicating that the different measures performed similarly overall.

5.4 Evaluation of the psychometric properties of the EQ-5D-Y.

The COSMIN checklist (25) which is a consensus based checklist developed to ensure methodological rigour in the reporting of the psychometric properties of outcome measures, was used in the study (Appendix 1). General requirements for discussion when evaluating psychometric properties of an outcome measure include:

Missing items:
A weakness of the study was that there were no missing responses on the EQ-5D-Y or PedsQL to report on because, unfortunately, the researcher checked for these while the children were completing the self-reported outcome measures. This could have potentially affected the results concerning the usefulness of the EQ-5D-Y compared to the PedsQL, as it was not possible to compare the number of missing responses between the two outcome measures. However, the researcher did anecdotally report that children required more reminding to rate all items when using the longer PedsQL than for the EQ-5D-Y. Other studies, using either outcome measure, have also reported very few missing responses and this has no effect on the results (145)(1)(114)(113). An Italian study (106) and an English study (156) found no missing responses on the EQ-5D-Y, rating it easy to understand and use in children between eight and twelve years.

Sample size:
The sample size was sufficiently powered to depict a change in HRQoL in those groups hypothesised to change over time (CI and AI). There was some attrition over time from the AI group as they were discharged before they could be reassessed, but the overall number of children with a health condition, remained sufficient.

Number of measurements taken:
A strength of the study was the number of measurements taken. Most participants repeated the measure three times. The majority of the acutely ill children repeated the measure at least twice, which was considered sufficient to determine responsiveness in the studies mentioned below. Only five of the 13 generic HRQoL measures reviewed had repeated the measures over a period of time and were able to comment on responsiveness. These included the CHQ (118), HUI (62), KINDL (99), PedsQL (114) and QOML (118). No repeated measure studies assessing the responsiveness of the EQ-5D-Y on a clinical sample was found, which highlights the relevance of this study.

Time interval:
A possible weakness was the short interval between test-retest periods. Being only one day apart some of the children may have remembered their previous scores. However a longer period may have resulted in a change in the acute health condition of the children in hospital. Measurements repeated at three monthly intervals allowed for changes in the children with a chronic health condition, due to better management of their condition. However over the seven month period, developmental changes and altered priorities may have occurred in all the groups of children, changing their concept of HRQoL.

Stability of respondents in interim periods:
The stability of the children's' health in interim periods could be determined from the contextual information provided by the therapists and from medical files. The performance of the instruments was tested under the assumption that the MS and SS children's health status would remain stable and that AI and CI children would show an improvement in HRQoL over time. It was possible to identify which children in each group did not fit the hypothesis based on worsening health status and they were removed from the analysis.

Test conditions:
Test conditions at all facilities were described and were exactly the same at each assessment period, for all outcome measures.

The following psychometric properties of the EQ-5D-Y were evaluated and will be discussed as they pertain to the research study objectives:

Internal consistency:
As the EQ-5D-Y and the other HRQoL measures used were multi-dimensional, the internal consistency of the various dimensions was expected to be low and was not assessed in this study. The ICC of test-retest VAS scores was determined and relationships between the dimensions and overall HRQoL (VAS) were examined and will be discussed.

Reliability:
The various forms of reliability were examined by assessing the measurement error. The measuring instrument error was assessed through test-retest of the children's responses 24 hours apart. The variability or stability of the characteristic was assessed by repeated measures over a three and seven month period (intra-rater reliability). The error attributed to the individual taking the measurement was assessed by correlating the responses of proxies and children's self-report (inter-rater reliability). It is recommended that further research investigate the inter-rater reliability between different proxies for the same child, as this aspect was not investigated in this study.

Content validity:
The EQ-5D-Y has been developed from the adult EQ-5D into a paediatric HRQoL measure, and it is recognised that it does not contain child specific dimensions. However, all dimensions measure aspects of HRQoL, but this was not specifically assessed.

Structural validity:
The relationship between dimension scores and overall HRQoL was examined

Hypothesis testing:
The ways in which the EQ-5D-Y was expected to perform with the different groups of children was hypothesised, before data collection.

Cross-cultural testing:
The original English version of the EQ-5D-Y and the version translated into Afrikaans were used, but both versions were not tested on the same participant. The isiXhosa version was not used as the isiXhosa speaking children could not read the language, but could read the English version. The PedsQL had not previously been translated into Afrikaans or isiXhosa and this was performed, following the forwards and backward translation process, before data collection. However the isiXhosa version was also not used, for the reason given above.

Criterion validity:
The ability of the EQ-5D-y to discriminate between children with different health statuses was examined. Convergent validity between similar dimensions on the PedsQL, WeeFIM and FPS were investigated, for the different groups of children.

Responsiveness:
The responsiveness of the measure to depict a change in HRQoL over time, was examined.

Interpretability:
The mean age (Std. Dev. and range) of the sample population was analysed, as well as the gender distribution and educational level of the children. The disease characteristics for each group was given. The selection of participants and the study settings were described. The distribution of scores in the different groups of children was analysed.

5.5 Reliability of the EQ-5D-Y

The first specific objectives of the study related to the reliability of the instrument. Reliability is a reflection of the relative amount of true or fixed value and random error. Sources of error include the measuring instrument, the variability of the characteristics and the individual taking the measure (108). Measuring instrument error was assessed through test-retest of the children's responses 24 hours apart. The variability or stability of the characteristic was assessed by repeated measures over a three and seven month period. The error attributed to the individual taking the measurement was assessed by correlating the responses of proxies and children.

5.5.1 Measuring instrument error

Generally the EQ-5D-Y dimensions demonstrated acceptable test-retest reliability in all the groups of children. As the dimension ratings are ordinal in nature, the Kappa statistic was used. This measures the consistency; that is whether or not the highest and lowest scoring children remain the highest and

lowest, respectively, on the second assessment. In other words, it measures consistency of the dimensions rather than absolute agreement across the groups. The EQ-5D-Y was found to be consistent in all dimensions except for the UA dimension. A possible reason for poor agreement in this dimension, in the present reliability study, is that several examples of usual activities are included in the questionnaire to explain the construct, for example "going to school, hobbies, sports, playing, doing things with family or friends" and the child might have been relating to a different, specific activity each time.

The encouraging results are similar to an Italian study (116) which assessed test-retest reliability of the EQ-5D-Y on children from the general population and children with Acute Lymphoblastic Leukaemia. The authors reported fair agreement across the dimensions, including UA.

Apart from the nature of the instrument, there may be other reasons that could contribute to the test-retest result. The test-retest method was applied on consecutive days and the short time interval between measurements may have resulted in measurement bias. According to Devon et al (2007)(103), the recommended time interval between original test and retest should be long enough to ensure that participants don't remember their initial scoring, but not too long for changes to have taken place. These authors recommend an interval of between two weeks to one month, but there is little consensus as to what time interval is suitable for children (83). Seven to ten days was the time interval used in the EQ-5D-Y Feasibility, Reliability and Validity multi-national study (1), two weeks apart in the development of the KIDSCREEN outcome measure (106) and DISABKIDS (86) and morning/afternoon in a study comparing the performance of the CHU and EQ-5D-Y (122). Test-retest on consecutive days was deemed appropriate in this study as the health status was likely to change in the AI children.

The absolute agreement of test-retest VAS, calculated using Interclass Correlation Coefficient (ICC) was found to be 0.77, indicating good agreement. It would appear that the VAS gives reliable results, similar to the Italian study by Scalone et al which also reported a high level of agreement (ICC of 0.82)(116). This was a heartening result, as there are concerns as to whether children as young as eight have the numeracy skills to respond appropriately to this task. However, as mentioned above, the marked ceiling effect could have influenced this.

It would appear that the EQ-5D-Y is a reliable measurement instrument in children with different health states.

5.5.2 Variability of the characteristic (HRQoL)

The stability of the self-reported HRQoL was measured over three and seven months in the MS and SS children as their health status was not anticipated to change much over this time. Transient variability, as is the nature of HRQoL over time (5)(6)(7), was considered and therefore it was expected that there would be moderate agreement between scores over time. HRQoL was however less stable than expected and only slight agreement was reported in each dimension over three months at the MS school. The SS children demonstrated moderate agreement between mobility and LAM scores only, over three months. However, over the longer seven month period there was better stability between scores in all dimensions, except for the WSU dimension. There are several possible explanations for the weak agreement between scores. There have been a true alteration in HRQoL due to fluctuating levels of perceived independence in the various dimensions over time, which could be attributed to fatigue on a particular day. This explanation was reported on in a United Kingdom study (122) which found limited reliability in EQ-5D-Y and CHU test-retest results in younger TD children. The test was performed in the morning and then in the afternoon of the same day, when the children were fatigued. Even though the measures were administered in repeated assessments at roughly the same

time of day in our study, it is possible that the children, especially the SS children with limited mobility, would be easily fatigued if they had recently performed a physically activity which they found challenging. This could affect their responses on the EQ-5D-Y dimensions.

Also, different levels of maturity might account for a decrease in stability. As children's abilities, outlook and priorities may change as they develop, this could account for some inconsistencies in reporting over a period of time (90). In the light of the increased stability at the SS over the longer seven month period however, however, it is unlikely that this played a role in their responses.

This lack of stability might also be an indication that children are not able to differentiate between specific HRQoL and general QoL which may fluctuate more over time. This was not case in the United Kingdom study (122) which reported that children understood the construct being measured and children as young as five years can reliably self-report if given age appropriate measures (81). The poor stability observed in MS and SS children in the current study may not be due to a misunderstanding of the constructs, but rather to confusion as to whether the problem encountered was due to a health condition or to other factors, such as being hungry or feeling unsafe on that day, emotional distress due to personal conflicts with others or being excluded from a particular social peer group. The argument that children are unable to distinguish between general QoL and specific HRQoL is weakened, however, by the lack of impact of major life events on HRQoL. In this study it was found that life events such as a death in the family, divorce, birth of a new sibling, moving house or a surgical intervention, did not have an effect on the VAS or the WSU dimension.

The stability of the VAS, compared to the fluctuations in the dimension scores, may indicate that, for children, a functional limitation does not necessarily translate into a reduction in HRQoL, so that although the function of a child may alter with time, the VAS remains relatively stable or even improves over time. This is discussed below in Section 5.7.

When the summary EQ-5D-Y dimension score (Index Score), derived from a QALY scale developed by Craig et al (2015) (66), is used the Index Score shows stability at the MS and SS, as expected. This is in line with the stable VAS at these institutions. This could indicate that the fluctuations in individual dimensions are levelled out when a summary score is used. Although the values developed by Craig et al were used in this study to allow for the calculation of a composite score, these have not yet been adopted by the EuroQoL Foundation.

5.5.3 Agreement between individuals taking the measure

Correlations between children and their proxies report were used to explore the amount of error that might be due to the respondents. Similar to the self-report, there was a significant difference in the mean ranking of the proxy Index Scores between institutions, but whereas the self-reports Index Scores rankings were significantly different between all institutions except for SS and CI, there were no differences in ranking between the MS and CI, and the SS and AI in the proxy reports. It would seem that the children are better able to identify their problems on the dimensions than the proxies.

For all groups, there was limited agreement between self-report and proxy at a dimension level, except for Mobility, being an easily observable dimension, for which there was moderate to good agreement. The highest levels of agreement between child and proxy were seen in the AI, with only LAM dimension having slight disagreement.

As the validity and reliability study of the Spanish EQ-5D-Y proxy version (90) found high agreement between proxy and self-report on the Mobility, LAM and UA dimensions and lower agreement for P/D and WSU dimensions, it was expected that the proxies might report similar or maybe more problems in the first three dimensions, which are observable dimensions and less in the other two which are more subjective. This is often true for the psychological component in children's health (16). Children with a health condition clearly are more anxious than their proxies realise. There were very large discrepancies in the percentages of children versus proxies reporting problems in the WSU dimension (for example 40% of the children at the SS reported problems, whereas the proxy reported that no child had any problems). Of similar concern is that, in the CI, there was a large difference between the high percentage of children reporting problems and the relatively low number of children identified as having problems in all domains by the proxies. In contrast, the proxies of the AI children reported more problems in all dimensions apart from WSU. This highlights the need for clinicians, who were the proxies in children with health conditions, to use self-reported HRQol measures to become more aware of how a health condition affects the child's experience of HRQoL.

For all groups combined there was moderate agreement between self-report and proxy VAS scores, but this was not the case at institutional level, where the two respondent's VAS scores were only significantly correlated at the MS school. The majority of respondents from the MS reported full health and there were a larger number of respondents at this institution, which may have inflated the correlation. This discrepancy between raters was similar to a number of other studies (93) (92) (15) (77). The correlation between proxy and child report was the highest at the AI and this might have been significant if the sample size had not been so small.

These discrepancies in raters' description and perception of HRQoL may be due to a limited understanding by the proxy on how the child functions in different contexts, such as at school, home and when socialising with peers. Proxies and children may also have different perceptions of HRQoL as seen in a study by Kaartina et al (93). When they compared parent proxy and self-report HRQoL from 379 Malaysian adolescents, these authors found that the proxies reported lower HRQoL than the children. While the proxy may not be aware of the extent of the emotional impact the health condition has on their child, they do have a broader perspective and comparison base on which to base their report. As it is the adult proxy who makes decisions on the child's health care and utilisation of services, their perspective needs to be considered. It is therefore generally recommended that both proxy and self-report measures be taken into account when assessing a child's HRQoL whether for planning a management programme in which the child's perspective is most important or for determining the cost effectiveness of a programme and resource allocation, where the proxy report may be of more benefit (93)(26).

In this study it would appear that a proxy report might be useful in an acute setting, especially if the child is too ill to self-report, but less reliable in settings where children have chronic illness or disability as the proxies seem to under-estimate the emotional distress of the condition on these children. It is clear, however, that the proxy and self-report should not be used interchangeably.

5.6 Validity of the EQ-5D-Y

The next study objective related to the construct validity of the EQ-5D-Y, which was examined by assessing both discriminant and convergent validity.

5.6.1 Discriminant validity based on dimensions

The discriminant validity of the EQ-5D-Y dimensions was examined by comparing the HRQoL profiles of the different groups. As hypothesised, the MS children reported significantly fewer problems in each

dimension than the children in the AI. In addition, they reported fewer problems in the functioning domains of Mobility and LAM than the SS children, many of whom had functional limitations due to their health conditions. It was interesting that, apart from the AI children, there was no difference between the rankings of UA, P/D or Discomfort or WSU between the other groups. This would indicate that, despite being in special institutions, the children appear to correctly reference their UA against themselves and their class mates rather than against children in general. In addition, the existence of a chronic health condition or disability did not significantly impact on their emotional state, which may be an indication that RS took place. This however, was not analysed in sufficient detail to ascertain that RS indeed took place. As discussed previously, it is unclear why so many MS children reported P/D and although the mean ranking of the SS and CI was higher, there was no significant difference between these groups.

These results are similar to an Italian study(116) which compared TD children to acutely ill children, and which found that TD children reported significantly fewer problems in all dimensions of the EQ-5D-Y compared to the acutely ill children. Therefore discriminant validity between these two groups was evident. The Swedish study (158) using children in groups known to differ in health status, also found discriminant validity between children with severe illness and/or handicap and children with chronic rhinitis and/or asthma. A Spanish and Catalonian study (173) also found the EQ-5D-Y able to satisfactorily discriminate between known groups of TD children with and without a chronic or mental health condition.

It would appear that the EQ-5D-Y dimensions exhibited discriminant, known group validity.

5.6.1 Discriminant validity of the VAS

Similar to the dimensions, the VAS of the AI children was significantly lower than all other groups but there was little difference in VAS between the other groups. Jelsma et al. (15) reported that children with disabilities in Cape Town actually reported higher overall HRQoL than their able-bodied peers so the lack of difference between the MS and SS children was not surprising. However, the lack of difference between the children in the CI and the other children was unexpected, as it was anticipated that children with health conditions would have a poorer overall perception of their health state. The Swedish study (158) found that children with two or more clinical characteristics reported significantly lower on the VAS. In contrast to a Canadian study (2010)(157) which found disparities between dimension scores and VAS, depending on the children's socio-demographics, the children in this study were all from the same socio-economic background and this was unlikely to have been a factor.

The VAS was able to clearly discriminate between the AI and MS children.

While the EQ-5D-Y VAS could discriminate between the acutely ill children and the other three groups, the PedsQL total score was only able to discriminate between the MS children and the AI children. The WeeFIM total score was able to discriminate between the children at the CI and SS and AI. However, it was unable to depict a difference in functional independence between the AI and SS. It would seem that the WeeFIM is not sensitive enough to depict a functional difference between children using a mobility device and children confined to a hospital bed.

5.6.2 Convergent validity

The study objective of assessing convergent validity was investigated by determining whether there were correlations between the scores on similar dimensions of previously validated p-HRQoL outcome measures across the three levels of the EQ-5D-Y dimensions.

Convergent validity was found between the different levels of all similar dimensions of the EQ-5D-Y and the PedsQL, for the AI group only. However, convergence between the functional (observable) items on EQ-5D-Y and the WeeFim was evident across all groups tested (SS, CI and AI). The dimensions of P/D and WSU did not show the expected relationship with the PedsQL items in the SS or CI children, but only in the AI children. As the PedsQL references the construct to "the last month" and the EQ-5D-Y to that day, it might not be surprising that the dimensions do not concur. It was previously noted that there were more MS children reporting P/D than would be expected and the lack of concurrence between the PedsQL "I hurt" item and the P/D dimension might be due to the relatively minor and transient nature of the pain or discomfort experienced by these children.

Similar to this study, Ravens-Sieberer et al (2010) (1) reported convergent validity between EQ-5D-Y WSU and PedsQL Feelings dimension in TD children. It is not clear why this relationship was not evident in children with a chronic health condition.

The Italian study (116) investigating concurrent validity between EQ-5D-Y and PedsQL found good levels of convergence for Self-care in TD and acutely ill children. However, it is not clear which dimensions or items were compared, as there is no specific self-care dimension in the PedsQL measure, only "I have problems with showering or bathing myself", which is only one item of self-care.

As with the dimensions, the VAS performed best at the AI and was correlated with both the WeeFim and the PedsQL totals in this group. It was noted that the WeeFIM Social cognition subscale focuses on language and cognitive skills, while the PedsQL Socialising dimension focuses on the child's emotional and social abilities, which are different constructs and therefore associations between the two would be expected to be poor. The EQ-5D-Y does not assess either of these constructs, so they were not included in the analysis.

The total PedsQL and WeeFim scores did show convergent validity across all groups, which is consistent with some of the literature. A Turkish study (174) of 32 children with cerebral palsy found significant relationships between all the WeeFIM subscales and PedsQL dimensions. A Canadian study (175) of 224 children with physical disabilities ranging from cerebral palsy to spina bifida and developmental delay found fair to moderate correlations between WeeFIM Self-care and Mobility subscales and PedsQL Health and Activities dimension. However, this was in contrast to another Canadian study (176) of 124 children with physical disabilities and a USA study (177) of 562 children with ambulatory cerebral palsy. This finding highlights the usefulness of using more than one outcome measure to accurately determine a child's HRQoL from their own perspective and an objective measure of their actual functional independence.

No other studies were found comparing the EQ-5D-Y with WeeFIM. The WeeFIM demonstrated high ceiling effects in the various dimensions in the CI children and should be used with caution in these children.

The convergent validity of the EQ-5D-Y with the other outcome measures, therefore, appears to be acceptable in the AI children, but variable in other groups of children.

5.7 Responsiveness to change in HRQoL over time

The responsiveness to change, when a change in HRQoL does occur, is an extension of validity and was examined in the three outcome measures, EQ-5D-Y, PedsQL and WeeFIM. Responsiveness was quantified as the effect size (178) in order to give an indication of the minimally clinically important observed change in these children (179)(126). In order to interpret the meaning of the change in scores, the types of change which occurred was examined by linking it to underlying clinical change

(178)(125) according to the hypothesis. A mixed, rather confusing picture emerged with regard to the dimensions with MS and CI children reporting a medium effect for Mobility, children at the SS reporting a medium effect for LAM and the AI children reporting medium effects for Mobility, LAM and WSU. Although it was not surprising that there was no change in UA in the AI children who were still in hospital at the time of the second assessment, the lack of responsiveness of the P/D dimension was unexpected, especially since the FPS did demonstrate a medium effect. However, when the summary Index Scores were used, a clearer picture emerged with a medium effect size evident in the two groups, the CC and AI groups, who were expected to show improvement. There is an obvious need to do further studies using this composite score derived from summarising the EQ-5D-Y scores on a QALY scale, as is emphasised by Craig et al (66).

A similar pattern of responsiveness was detected in the VAS scores in that the CC and AI were the only two groups to demonstrate a medium change. The AI children improved as their acute condition and pain was managed. The CI children improved after being admitted to a facility for better management of their chronic health condition.

The WeeFIM detected a large improvement in SS and AI groups and medium in CI. Even though the children at the SS had fairly stable health conditions with minimal physical changes occurring within the study period, most of their changes occurred in the cognition dimension. A USA study (180) did find the WeeFIM responsive in detecting a change over a one year period, in 205 children identified with chronic disabilities such as spina bifida and cerebral palsy (similar conditions to the SS children) and developmental delays, using effect size. The difference between this study and the present one is the time period and, even though the SS children were not expected to change within the three month period, they might change over a longer period as their chronic disability deteriorates.

The WeeFIM performed the best in the CI children, who were expected to change slightly, and the AI where a large change was expected.

The PedsQl did not respond as expected in any of the groups, only detecting a change in the children at the SS. A possible reason for the lack of responsiveness in the AI could be that the PedsQL questions how much of a problem the child had in the various dimensions during the last month and it may not be suitable for repeated assessments on acutely ill children, who are assessed more often than once a month. It did not seem to detect a change in the CI children and may not be sensitive enough to changes occurring in this group. This was in contrast to a PedsQL study (181) which found the generic measure to be responsive to change in 231 children with chronic rheumatoid arthritis, assessed 60 days apart. Another two year study, using the PedsQL to assess HRQoL in children who regularly participated in sport compared to those who did not, found an increase in overall HRQoL in the children regularly participating. As this was expected, it indicated that the PedsQL was able to show change when it was expected, but interestingly no change was observed on the physical dimension or emotional dimension (even though this was expected to change) (182).

The MS group were not expected to show change, so it would seem that the PedsQL only performed as expected in this group.

The FPS performed as expected in detecting a small change in pain levels in the MS, SS and CI. A medium change over time was seen in the AI group as their pain medication took effect.

5.8 Response shift

It might be expected that scores on the different dimensions and overall HRQoL (VAS score) would be associated (39). However, in this study the VAS scores were only significantly correlated with all

dimension scores in the children at the AI. Based on the descriptive analyses, the reported changes on the five dimensions of the EQ-5D-Y do not seem to be strongly related to the VAS scores in the CI and SS children. It is not possible to conclude whether this is an example of Response Shift or not, as the children's responses were not tested retrospectively with a "then-now" test to determine whether they would rate their health at baseline differently.

It has been recognized in other studies that individuals with a chronic disability adapt to their situation by either changing their internal standard and values or their conception of HRQoL, a phenomenon known as RS (38)(99)(183)(96)(184)(16). The children with a chronic health condition or disability were able to correctly identify their problems on a dimensional level, but scored similar levels on the VAS for overall HRQoL as the MS children with minimal or no problems on the dimensions.

A previous South African study (15) found that children with long standing, chronic conditions reported a higher general HRQoL on the VAS than those of TD children, whereas their dimension scores indicated greater problems. This was similar to other studies that found that children with a chronic disability did not equate limited functional independence with a lowered HRQoL (26)(36)(15). It would seem, in children with a chronic disabling condition who are otherwise healthy, the VAS and the dimension scores might be measuring different constructs and that functional limitations might not always equate to poor HRQoL on the VAS.

This disparity between dimension scores and VAS could have an impact on the psychometric properties of a HRQoL outcome measure, which may influence the interpretation of clinical research findings (99) and, therefore, the EQ-5D-Y should be used with caution in children with a chronic condition or permanent disability.

The VAS could be seen by these children as a measure of their satisfaction level with their participation within their community, despite their physical limitations. Albrecht and Devlieger (1999) (96) described the QoL of people with a disability as a balance between body, mind and spirit within their community context. They suggested the factors that negatively influenced the QoL of people with a disability are a lack of social and community ties, poor access to resources and a disabling environment. None of these negative influences were present in the sample of children with a chronic disability or health condition, which could account for their high satisfaction levels, as seen by high VAS scores. The children were all at schools which catered for their disability, were well resourced, with an adaptive environment and situated within the community from which the children came.

A similar finding was evident in a study of HRQoL, in 40 children from Dutch neuromuscular referral centres, living with muscular dystrophy (172) and 30 children with physical disabilities attending a special school in the North West Province, South Africa (165). The severity of their condition and physical limitations did not negatively affect their HRQoL, while being able to socialise with their peers positively influenced their HRQoL.

It is also unlikely that the children with a long-standing chronic health condition, some of whom were born with the health condition, would have the same concept of HRQoL as children with no health condition or those with a sudden acute health condition, as discussed by Swartz in a paper on the application of RS theory (99). It is possible that MS children, who have had no experience of a health condition resulting in a disability or needing hospitalisation, would have a very different concept of HRQoL compared to children with a chronic or acute health condition.

This could affect the psychometric properties of HRQoL measures (99)(39). For high discriminant validity, it is assumed that the participants all have the same conception of what is being measured

and similar experiences to refer to (103), which is not the case in the children in this study. RS may also influence the sensitivity of HRQoL measures to depict differences between different health conditions and interpretation of clinical research findings (39), which could ultimately introduce bias (185).

Despite observable, objective limitations (alpha change) in the children with chronic disabilities, the phenomena of possible RS could result in underestimation of the effect of the health condition on their HRQoL. This has been discussed by Swartz (99) and could lead to the incorrect impression, if the HRQoL measure is used to calculate quality-adjusted-life-years (QALY). As the QALY is the basis of economic evaluations used for resource allocation (52), RS could result in fewer resources being allocated to these children. It is important for users of HRQoL outcome measures to be aware of possible RS when interpreting results for planning changes in management of the condition or for economic evaluations, as it seems to have occurred in the children with chronic conditions in this study.

5.9 Feasibility and Usefulness of EQ-5D-Y outcome measure

Feasibility was assessed by the length of time taken to complete the measure and not by the number of missing responses as other studies have done (1)(90)(116)(158), as the researcher was present to ensure that the children did not leave out any responses. The EQ-5D-Y, being a short questionnaire, took only five minutes to complete compared to 15 minutes to complete the PedsQL. Children became easily distracted when completing the longer PedsQL and had to be reminded to complete the task at hand more often, compared to when completing the EQ-5D-Y. This would indicate a preference for using the EQ-5D-Y in these children. The EQ-5D-Y would seem to consist of enough items to assess HRQoL, but does not take long to complete, a desirable attribute of a paediatric HRQoL measure (144). As the EQ-5D-Y displayed an equivalent or greater responsiveness to change, there is strong argument to use the shorter instrument as a routine outcome measure.

Although there have been several reports on the usefulness of the EQ-5D-Y as an outcome measure (13)(1)(90)(15), this is the first study, to our knowledge, which examined the usefulness within a routine clinical context.

The sample of physiotherapists who reported on the measure was quite small (nine) but these were the only clinicians involved with the children in the study. In general, these clinicians reported that the EQ-5D-Y was acceptable, useful and easy to apply. The majority stated that they would continue to use it in future as it provided them with insight into the children's subjective perception of their own health condition.

Two therapists at the AI found some eight to nine-year-old children had difficulties understanding the "UA" dimension, but this was not the case elsewhere. It was found that the children generally understood the questions, rendering the measure feasible to use in this population.

All the therapists noticed a relationship between responses and clinical signs, mostly in the Mobility and LAM dimensions. As was discussed earlier, the large discrepancy between self-report and proxy, particularly in the SS and CI, emphasises the need for therapists to become more aware of the subjective experiences of the children that they treat. This was recognised and all therapists found the measure provided them with additional information in the P/D and WSU dimensions and the extent to which these aspects affected the children's HRQoL. They found this less obviously observable information most useful when planning the management of the child, as they gained better insight

into the areas most affecting the child's HRQoL and could, therefore, pay more attention to these aspects.

The EQ-5D-Y appears to be a useful measure to include in assessing children, for completeness of information and for aiding the planning of a management programme for the child. However, in contrast to the FPS, the dimension of P/D did not demonstrate responsiveness in the AI children and it is suggested that the EQ-5D-Y be supplemented by the FPS within the AI group.

6 CHAPTER 6: CONCLUSION

Research into HRQoL, resource allocation and decisions regarding the management of a health condition may be based on the scores obtained using a HRQoL measure. For these purposes, it is important that the measure is reliable and valid when used on a particular population sample (25)(31). HRQoL measures should be stable in the absence of observable change, but responsive to change if changes do occur.

We found that the EQ-5D-Y performed well when used on acutely-ill children. The VAS scores in this group of children were significantly correlated with all dimension scores in these children. This measure was found to be reliable and could clearly discriminate between children with an acute illness and children in the MS schools with no health condition. In addition, good convergent validity between the EQ-5D-Y and the other outcome measures, the PedsQL, WeeFIM and FPS was demonstrated in the AI children. It seems that children with an acute health condition are able to respond most appropriately to the EQ-5D-Y as they are able to recognise the impact of their health condition on their QoL, comparing their acute condition to their normal condition. They have a comparison of HRQoL on which to base their responses. Likewise, responsiveness to change in EQ-5D-Y dimension scores, except for P/D and VAS, was most noticeable in the acutely-ill children. It would seem that a specific pain measure, FPS, is better able to depict changes in pain in these children and should be used in conjunction with the EQ-5D-Y in a hospital setting.

The EQ-5D-Y dimensions did reflect the SS and CI problems appropriately, but this was not the case with the VAS. Despite reporting problems in the various dimensions, these children scored a high overall HRQoL on the VAS. It would seem that these children do not equate lowered functional ability on dimensions with lowered overall HRQoL on VAS.

The measure did not perform well in the MS children with unexpected results obtained for the P/D dimension. In addition, some children reported problems with Mobility, although this was not reflected on the functional independence outcome measure, the WeeFIM. This may reflect an interpretation and contextual issue. It would seem that a healthy child might not relate problems with "walking about" to a health state, but rather to environmental barriers.

There was a disparity between dimension scores and VAS in the MS children. It is possible that children who have not experienced a health condition may not be able to differentiate between *health* related quality of life and *general* quality of life. High ceiling effects, as were evident in the MS children, also affected the psychometric properties of the measure.

However, when comparing the Index Score (a composite summary score of dimensions, based on a QALY scale (66)) across institutions, the MS children's score indicated almost no problems on any dimension, followed by the CI, SS and the AI with the most problems, as expected. There was a relationship between the responsiveness of the VAS and Index Score, with both being the most responsive to change in the AI and CI children, who were expected to show a change over time. This highlights the need for more studies to be done using the composite score summarising the health condition for comparison against other measures which produce a single summary score. Proxy and self-reports should not be used interchangeably. The proxy report is appropriate for use if the acutely ill child is unable to self-report, but is not recommended for use in the SS and CI children.

6.1.1 Recommendations for practice:

When choosing a HRQoL outcome measures for practice, the population on which it will be used should be considered.

The EQ-5D-Y can be used with confidence in acutely ill children as it fulfils all psychometric requirements in this group. It was quick and easy to use and provided the therapists with additional information of which they were not previously aware. The EQ-5D-Y dimensions performed adequately in the SS and CI groups, but not the VAS.

The PedsQL could discriminate between the MS and AI groups, but was not responsive in depicting a change in HRQoL in the AI group. Therefore PedsQL is recommended for use in assessing HRQoL in MS children, but not in acutely ill children.

The WeeFIM, which is an objective measure of function rather than HRQoL, performed well in identifying different levels of independence between the SS, CI and AI groups. It was most responsive to change in the AI group, but is also recommended for use in children with a chronic disability. Ceiling effects may be apparent of this measure is used on children with a chronic health condition and without a physical disability. In this case, the suitability of this measure for assessing function in these children may be limited.

6.1.2 Recommendations for research:

When choosing a HRQoL measure for research purposes, the effect of RS on the psychometric properties of the measure, when used on children with a chronic health condition, should be considered. Response shift as such, was not investigated in this study but it was found that these children appear not to equate lowered functional ability with lowered overall HRQoL. It is therefore recommended that further qualitative research is necessary to understand the constructs underpinning children with a chronic health condition's VAS reported health.

The need to develop paediatric-based weights for use on EQ-5D-Y health states is currently being considered. Further research should be conducted in valuation studies assessing preferred EQ-5D-Y health states in children, in order to develop a value set for use in children. As the EQ-5D-Y has been found to be reliable, valid and responsive in acutely-ill children and less so in chronically-ill children, there is a need to develop a EuroQoL composite score summarising the health condition for comparison against other measures which produce a single summary score. There has been progress in this direction with the 2015 valuation study of Craig et al (66). However, further studies using the summary score, based on the QALY scale, are needed before the results of Craig et al study can be adopted by the EuroQoL Group.

6.1.3 Recommendations for use of the EQ-5D-Y outcome measure in economic valuations

Children with a chronic health condition tend to under-report on the effect of limited functional ability on their HRQoL. This may lead to fewer resources being allocated to their health care. In this case, the use of proxy reports is advisable. The adult proxy is ultimately responsible for the child's health care needs and has a broader outlook on which to base responses.

The EQ-5D-Y was therefore found to be a useful instrument to measure HRQoL and its use is recommended in children with acute illness and, to a lesser extent, in those with chronic health conditions. Routine use may result in more holistic care being delivered to these children as their perspective is taken into account in the planning of their own management.

REFERENCES

1. Ravens-Sieberer U, Wille N, Badia X, Bonsel G, Burström K, Cavrini G, et al. Feasibility, reliability, and validity of the EQ-5D-Y: results from a multinational study. Quality of Life Research [Internet]. 2010 Aug [cited 2014 Mar 19];19(6):887–97. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2892614&tool=pmcentrez&rendertype=abstract

2. Costanza R, Fisher B, Ali S, Beer C, Bond L, Boumans R, et al. Quality of life: An approach integrating opportunities, human needs, and subjective well-being. Ecological Economics [Internet]. 2007 Mar [cited 2014 Mar 23];61(2-3):267–76. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0921800906000966

3. Milton CL. The Ethics of Defining Quality of Life. Nursing Science Quarterly [Internet]. 2013 Apr [cited 2014 Dec 13];26(2):121–3. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23575485

4. Felce D, Perry J. Quality of life: Its Definition and Measurement. Research in Developmental Disabilities [Internet]. 1995 Jan;16(1):51–74. Available from: http://linkinghub.elsevier.com/retrieve/pii/0891422294000288

5. Wu M, Brazier JE, Kearns B, Relton C, Smith C, Cooper CL. Examining the impact of 11 long-standing health conditions on health-related quality of life using the EQ-5D in a general population sample. The European Journal of Health Economics [Internet]. 2015 Jan 10 [cited 2014 Dec 13];16(2):141–51. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24408476

6. Spieth LE, Harris C aroline V. Assessment of Health-Related Quality of Life in Children and Adolescents: An Integrative Review. Journal of Pediatric Psychology [Internet]. 1996 Apr;21(2):175–93. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8920152

7. Ackerman IN, Busija L, Tacey M a, Bohensky M a, Ademi Z, Brand C a, et al. Performance of the Assessment of Quality of Life Measure in People With Hip and Knee Joint Disease and Implications for Research and Clinical Use. Arthritis Care & Research [Internet]. 2014 Mar [cited 2014 Jul 10];66(3):481–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23983000

8. Fallowfield L. What is quality of life ? [Internet]. Health Economics. 2009. p. 1–8. Available from: http://www.whatisseries.co.uk/whatis/pdfs/What_is_QOL.pdf

9. Cella David, Yount Susan, Roth rock Nan, Gershon Richard CK, Bryce. R, Deborah A, Fries James F, Bruce Bonnie RM. The Patient-Reported Outcomes Measurement Information System (PROMIS). Medical Care. 2007;45(1):1–16.

10. Sprangers M a G, Schwartz CE. Reflections on changeability versus stability of health-related quality of life: distinguishing between its environmental and genetic components. Health and Quality of Life Outcomes [Internet]. 2008 Jan [cited 2014 Jul 19];6(89). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2607267&tool=pmcentrez&rende

rtype=abstract

11. Varni JW, Burwinkle TM, Lane MM. Health-related quality of life measurement in pediatric clinical practice: an appraisal and precept for future research and application. Health and Quality of Life Outcomes [Internet]. 2005 Jan [cited 2013 Mar 2];3(34). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1156928&tool=pmcentrez&rendertype=abstract

12. Solans M, Pane S, Estrada M-D, Serra-Sutton V, Berra S, Herdman M, et al. Health-Related Quality of Life Measurement in Children and Adolescents: A Systematic Review of Generic and Disease-Specific Instruments. Value in Health [Internet]. International Society for Pharmacoeconomics and Outcomes Research (ISPOR); 2008 [cited 2013 May 23];11(4):742–64. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18179668

13. Wille N, Badia X, Bonsel G, Burström K, Cavrini G, Devlin N, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. Quality of Life Research [Internet]. 2010 Aug [cited 2014 Mar 19];19(6):875–86. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2892611&tool=pmcentrez&rendertype=abstract

14. Jelsma J, Pronk M, Ferguson G, Jelsma-Smit D. The effect of the Nintendo Wii Fit on balance control and gross motor function of children with spastic hemiplegic cerebral palsy. Developmental Neurorehabilitation [Internet]. 2013 Jan [cited 2013 May 5];16(1):27–37. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23030836

15. Jelsma J, Ramma L. How do children at special schools and their parents perceive their HRQoL compared to children at open schools? Health and Quality of life Outcomes [Internet]. 2010 Jan;8(72). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2919454&tool=pmcentrez&rendertype=abstract

16. Sattoe JNT, van Staa A, Moll H a. The proxy problem anatomized: Child-parent disagreement in Health related Quality of Life reports of chronically ill adolescents. Health and Quality of lLfe Outcomes [Internet]. BioMed Central Ltd; 2012 Jan [cited 2014 Jul 20];10(10). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3299605&tool=pmcentrez&rendertype=abstract

17. Barofsky I. Can quality or quality-of-life be defined? Quality of Life Research [Internet]. 2012 May [cited 2014 Dec 13];21(4):625–31. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21725867

18. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. Journal of Clinical Epidemiology. 2003;56:395–407.

19. Smith K evin W, Avis NE, Assmann SF. Distinguishing between quality of life and health status in quality of life research: a meta-analysis. Quality of Life Research. 1999;8:447–59.

20. Rajmil L, Herdman M, Fernandez de Sanmamed M-J, Detmar S, Bruil J, Ravens-Sieberer U, et al. Generic Health-Related Quality of Life Instruments in Children and Adolescents: A Qualitative Analysis of Content. Journal of Adolescent Health [Internet]. 2004 Jan [cited 2014 May 3];34(1):37–45. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1054139X03002490

21. Skevington SM, Lotfy M, O'Connell K a. The World Health Organization's WHOQOL-BREF quality of life assessment: Psychometric properties and results of the international field trial. A Report from the WHOQOL Group. Quality of Life Research [Internet]. 2004 Mar;13(2):299–310. Available from: http://link.springer.com/10.1023/B:QURE.0000018486.91360.00

22. Corr Susan PA. The Exploration of Quality of Life through Q Methodology. Occupational Science for Occupational Therapy. 2014;233–45.

23. Renwick R. Quality of Life for Young People with Neurological and Developmental Conditions : Issues and Challenges. Life Quality Outcomes in Children and Young People with Neurological and Developmental Conditions. 2013.

24. Jenney MEM, Campbell S. WHOQOL Measuring Quality of Life. Archives of Disease in Childhood [Internet]. 1997 Oct 1;77(4):347–50. Available from: http://www.who.int/mental_health/media/68.pdf

25. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. 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. 2010;19(4):539–49.

26. Bjornson KF, Belza B, Kartin D, Logsdon RG, McLaughlin J. Self-Reported Health Status and Quality of Life in Youth With Cerebral Palsy and Typically Developing Youth. Archives of Physical Medicine and Rehabilitation [Internet]. 2008 Jan [cited 2015 Jan 8];89(1):121–7. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3653574&tool=pmcentrez&rendertype=abstract

27. CONSTITUTION OF THE WORLD HEALTH ORGANIZATION 1 [Internet]. WHO. Constitution of the World Health Organization. 2006. 2006. p. 1–18. Available from: http://www.who.int/governance/eb/who_constitution_en.pdf

28. Huber M, Knottnerus JA, Green L, van der Horst H, Jadad AR, Kromhout D, Leonard B, Lorig K, Loureiro MI, van der Meer JWM, Schnable P, Smith R, van Weel, C SH. How should we define health ? BMJ. 2011;343;d4163:1–3.

29. Sawin KJ, Bellin MH. Quality of Life in Individuals With Spina Bifida: A Research Update. Developmental Disabilities Research Reviews [Internet]. 2010 Jan [cited 2014 Sep 10];16(1):47–59. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20419771

30. Gabriel Z, Bowling A. Quality of life from the perspectives of older people. Ageing and Society

[Internet]. 2004 Sep [cited 2014 Mar 25];24(5):675–91. Available from: http://www.journals.cambridge.org/abstract_S0144686X03001582

31. Basch E, Abernethy AP, Mullins CD, Reeve BB, Smith M Lou, Coons SJ, et al. Recommendations for Incorporating Patient-Reported Outcomes Into Clinical Comparative Effectiveness Research in Adult Oncology. Journal of Clinical Oncology [Internet]. 2012 Dec 1 [cited 2014 Jul 10];30(34):4249–55. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23071244

32. Eiser C, Morse R. A review of measures of quality of life for children with chronic illness. Archives of Disease in Childhood [Internet]. 2001 Mar;84(3):205–11. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1718699&tool=pmcentrez&rendertype=abstract

33. Palmermo TM, Long AC, Lewandowski AS, Drotar D, Quittner AL WL. Commentary on Evidence-based Assessment of Health-related Quality of Life and Functional Impairment in Pediatric Psychology. Journal of Pediatric Psychology [Internet]. 2008;33(9):997–8. Available from: http://www.jpepsy.oxfordjournals.org/cgi/doi/10.1093/jpepsy/jsn072

34. Hetherington R, Dennis M, Barnes M, Drake J, Gentili F. Functional outcome in young adults with spina bifida and hydrocephalus. Child's Nervous System. 2006;22:117–24.

35. Leger RR. Severity of Illness, Functional Status, and HRQOL in Youth with Spina Bifida. Rehabilitation Nursing. 2005;30(5):180–7.

36. Müller-Godeffroy E, Michael T, Poster M, Seidel U, Schwarke D, Thyen U. Self-reported health-related quality of life in children and adolescents with myelomeningocele. Developmental Medicine and Child Neurology [Internet]. 2008 Jun [cited 2015 Jan 9];50(6):456–61. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18371093

37. Pirpiris M, Gates PE, McCarthy JJ, D'Astous J, Tylkowksi C, Sanders JO, et al. Function and Well-Being in Ambulatory Children With Cerebral Palsy. Journal of Pediatric Othopedics [Internet]. 2006 [cited 2013 Jul 10];26(1):119–24. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16439915

38. Sprangers M a, Schwartz CE. Integrating response shift into health-related quality of life research: a theoretical model. Social Science & Medicine [Internet]. 1999 Jun;48(11):1507–15. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10400253

39. Schwartz CE, Andresen EM, Nosek M a, Krahn GL. Response Shift Theory: Important Implications for Measuring Quality of Life in People With Disability. Archives of Physical Medicine and Rehabilitation [Internet]. 2007 Apr [cited 2014 May 24];88(4):529–36. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17398257

40. Hainsworth KR, Davies WH, Khan KA, Weisman SJ. Development and Preliminary Validation of the Child Activity Limitations Questionnaire: Flexible and Efficient Assessment of Pain-Related Functional Disability. The Journal of Pain [Internet]. 2007 Sep [cited 2013 Jul 10];8(9):746–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17631055

41. Petersen S, Hägglöf BL, Bergström EI. Impaired Health-Related Quality of Life in Children With Recurrent Pain. Pediatrics. 2009;124(4):e759–67.

42. Gold Jeffrey I, Mahrer Nicole E, Yee Joyce PTM. Pain, fatigue and health-related quality of life in children and adolescents with chronic pain. Clinical Journal of Pain. 2009;25(5):407–12.

43. Greenhalgh J, Meadows K. The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: a literature review. Journal of Evaluation in Clinical Practice [Internet]. 1999 Nov;5(4):401–16. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10579704

44. Efficace F, Feuerstein M, Fayers P, Cafaro V, Eastham J, Pusic A, et al. Patient-reported Outcomes in Randomised Controlled Trials of Prostate Cancer: Methodological Quality and Impact on Clinical Decision Making. European urology [Internet]. European Association of Urology; 2013 Oct 30 [cited 2014 Jul 12];1–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24210091

45. Wang R-H, Wu L-C, Hsu H-Y. A path model of health-related quality of life in type 2 diabetic patients: a cross-sectional study in Taiwan. Journal of Advanced Nursing [Internet]. 2011 Dec [cited 2014 Jul 10];67(12):2658–67. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21615464

46. Schwartz CE, King M, Sprangers MAG. Let's engage the patient within us: a personal perspective on patient engagement. Quality of Life Research [Internet]. 2014 May [cited 2014 Jul 10];23(4):1045–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24277233

47. Chawla N, Arora NK. Why do some patients prefer to leave decisions up to the doctor: lack of self-efficacy or a matter of trust? Journal of Cancer Survivorship : research and practice [Internet]. 2013 Dec [cited 2014 Jul 12];7(4):592–601. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23892559

48. Sloan J, Symonds T, Vargas-Chanes D, Fridley B. Practical Guidelines for Assessing the Clinical Significance of Health-Related Quality of Life Changes within Clinical Trials. Drug Information Journal [Internet]. 2003 Jan 1 [cited 2014 Jul 12];37(1):23–31. Available from: http://dij.sagepub.com/lookup/doi/10.1177/009286150303700105

49. Le QA, Doctor JN, Zoellner LA, Feeny NC. Minimal clinically important differences for the EQ-5D and QWB-SA in Post-traumatic Stress Disorder (PTSD): results from a Doubly Randomized Preference Trial (DRPT). Health and Quality of life Outcomes [Internet]. 2013 Jan;11(59). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3635945&tool=pmcentrez&rendertype=abstract

50. Räsänen P, Roine E, Sintonen H, Semberg-Konttinen V, Ryynänen O-P, Roine R. Use of quality-adjusted life years for the estimation of effectiveness of health care: A systematic literature review. International Journal of Technology Assessment in Health Care [Internet]. 2006

Jan;22(2):235–41. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16571199

51. Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. British Medical Bulletin [Internet]. 2010 Jan [cited 2014 Jul 13];96:5–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21037243

52. Stevens K, Ratcliffe J. 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 [Internet]. Elsevier; 2012 Dec [cited 2014 Oct 1];15(8):1092–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23244812

53. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. Journal of Evaluation in Clinical Practice [Internet]. 2006 Oct;12(5):559–68. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16987118

54. Izard J, Hartzler A, Avery DI, Shih C, Dalkin BL, Gore JL. User-centered design of quality of life reports for clinical care of patients with prostate cancer. Surgery [Internet]. Mosby, Inc.; 2014 May [cited 2014 Jul 12];155(5):789–96. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24787105

55. Hofhuis JGM, van Stel HF, Schrijvers AJP, Rommes JH, Bakker J, Spronk PE. Conceptual issues specifically related to health-related quality of life in critically ill patients. Critical Care [Internet]. 2009 Jan [cited 2014 Jul 19];13(118). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2688122&tool=pmcentrez&rendertype=abstract

56. Landolt M a, Vollrath M, Niggli FK, Gnehm HE, Sennhauser FH. Health-related quality of life in children with newly diagnosed cancer: a one year follow-up study. Health and Quality of Life Outcomes. 2006;4(63).

57. Voepel-Lewis T, Zanotti J, Dammeyer J a., Merkel S. Reliability and Validity of the Face, Legs, activity, Cry, Consolability Behavioral Tool in Assessing Acute Pain in Critically Ill Patients. American Journal of Critical Care Nurses. 2010;19(1):55–61.

58. Aitken ME, Tilford JM, Barrett KW, Parker JG, Simpson P, Landgraf J, et al. Health Status of Children After Admission for Injury. Pediatrics. 2002;110(2):337–42.

59. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life: A conceptual model of patient outcomes. Journal of the American Medical Association [Internet]. 1995;273(1):59–65. Available from: http://jama.jamanetwork.com/article.aspx?articleid=385444

60. Stevens KJ. Working With Children to Develop Dimensions for a Preference-Based, Generic, Pediatric, Health-Related Quality-of-Life Measure. Qualitative Health Research [Internet]. 2010 Mar [cited 2014 Jul 19];20(3):340–51. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20054040

61. Horsman J, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI): concepts, measurement properties and applications. Health and Quality of Life Outcomes [Internet]. 2003 Jan;1(54). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=293474&tool=pmcentrez&rendertype=abstract

62. Shaw JW, Johnson JA, Coons SJ. US Valuation of the EQ-5D Health States: Development and Testing of the D1 Valuation Model. Medical Care. 2005;43(3):203–20.

63. Sintonen H, Richardson J. The 15-D Measure of Health Related Quality of Life: Reliability, Validity and Sensitivity of its Health State Descriptive System. National Centre for Health Program Evaluation, Working Paper 41, Melbourne. Finland; 1994. Report No.: 42.

64. Moock J, Kohlmann T. Comparing preference-based quality-of-life measures: results from rehabilitation patients with musculoskeletal, cardiovascular, or psychosomatic disorders. Quality of Life Research [Internet]. 2008 Apr [cited 2014 Jul 10];17(3):485–95. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18288590

65. Chiou CF, Weaver MR, Bell M a., Lee T a., Krieger JW. Development of the multi-attribute Pediatric Asthma Health Outcome Measure (PAHOM). International Journal for Quality in Health Care. 2005;17(1):23–30.

66. Craig Benjamin M; Greiner Wolfgang; Brown Derek S; Reeve Bryce B. Valuation of Child Health-Related Quality of Life in the United States. Health Economics. 2015;

67. Noyes J, Edwards RT. EQ-5D for the Assessment of Health-Related Quality of Life and Resource Allocation in Children: A Systematic Methodological Review. Value in Health [Internet]. Elsevier Inc.; 2011 Dec [cited 2014 Jul 9];14(8):1117–29. Available from: http://dx.doi.org/10.1016/j.jval.2011.07.011

68. Ebrahim S, Parshuram C. Comparison of utility scores from the Visual Analog Scale and Health Utilities Index 3 in children following pediatric intensive care unit admission. Journal of Child Health Care [Internet]. 2013 Aug 12 [cited 2014 Jul 10]; Available from: http://www.ncbi.nlm.nih.gov/pubmed/23939724

69. Shenkman Elizabeth A; Revicki Dennis A; Wen Pey-Shan;; Huang I-Chan. Quality of Life Measurements for Children with Life-Threatening Conditions: Limitations and a New Framework. Child Indicator Research. 2011;4(1):145–60.

70. Guyatt GH, Veldhuyzen Van Zanten SJ, Feeny DH, Patrick DL. Measuring quality of life in clinical trials: a taxonomy and review. Canadian Medical Association Journal. 1989;140(12):1441–8.

71. Yarlas A, Yen L, Hodgkins P. The relationship among multiple patient-reported outcomes measures for patients with ulcerative colitis receiving treatment with MMX® formulated delayed-release mesalamine. Quality of Life Research [Internet]. 2014;24(3):671–83. Available from: http://link.springer.com/10.1007/s11136-014-0797-2

72. Longworth L, Yang Y, Young T, Mulhern B, Hernández Alava M, Mukuria C, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. Health Technology Assessment [Internet]. 2014 Feb [cited 2014 Jul 11];18(9):1–224. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24524660

73. Pocoski J, Benjamin K, Michaels LA, Flood E, Sasane R. An overview of current trends and gaps in patient-reported outcome measures used in haemophilia. European Journal of Haematology. 2014;93(75):1–8.

74. Svavarsdóttir Erla Kolbrún; Orlygsdottir Brynja. Comparison of Health-Related Quality of Life Among 10- to 12-Year-Old Children With Chronic Ilnesses and Healthy Children: The Parents' Perspective. The Journal of School Nursing. 2006;

75. Law M, Hanna S, Anaby D, Kertoy M, King G, Xu L. Health-related quality of life of children with physical disabilities: a longitudinal study. BMC Pediatrics [Internet]. BMC Pediatrics; 2014 Jan [cited 2014 Jul 20];14(26). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3911963&tool=pmcentrez&rendertype=abstract

76. Varni James W; Seid Michael; Kurtin Paul S. Pediatric Health-Related Quality of Life Measurement Technology : A Guide for Health Care Decision Makers. Journal of Clinical Outcomes Management. 1999;6(4):33–40.

77. Matza LS, Swensen AR, Flood EM, Secnik K, Leidy NK. Assessment of Health-Related Quality of Life in Children: A Review of Conceptual, Methodological, and Regulatory Issues. Value in Health. 2004;7(1):79–92.

78. Velasco Miranda MJ;Domingues Martin E; Diez Arroyo FJ; Perez Mendez P; de Buitrago Amigo Gonzales J. Health related quality of life in type 1 diabetes mellitus . Anales de Pediatría [Internet]. 2012 [cited 2013 May 1];77(5):329–33. Available from: http://www.sciencedirect.com/science/article/pii/S1695403312001646

79. Hinds, Pamela S, Burfhen, Elizabeth A, Haase, Joan E, Phillips CR. Advances in Defining, Concepualizing, and Measuring Quality of Life in Pediatric Patients With Cancer. Oncology Nursing Forum. 2006;33(1):23–30.

80. Young B, Rice H, Dixon-Woods M, Colver AF, Parkinson KN. A qualitative study of the health-related quality of life of disabled children. Developmental Medicine and Child Neurology [Internet]. 2007 Sep;49(9):660–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17718821

81. Varni JW, Limbers CA, Burwinkle TM. 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 [Internet]. 2007 Jan [cited 2014 Jul 17];5(1). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1769360&tool=pmcentrez&rendertype=abstract

82. Fekkes M, Theunissen NC, Brugman E, Veen S, Verrips EG, Koopman HM, et al. Development and psychometric evaluation of the TAPQOL: a health-related quality of life instrument for 1-5-year-old children. Quality of Life Research [Internet]. 2000 Jan;9(8):961–72. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11284215

83. Riley AW. Evidence That School-Age Children Can Self-Report on Their Health. Ambulatory Pediatrics. 2004;4:371–6.

84. Pell T, Jarvis T. Developing attitude to science scales for use with children of ages from five to eleven years. International Journal of Science Education [Internet]. 2001 Aug [cited 2014 Jul 17];23(8):847–62. Available from: http://www.tandfonline.com/doi/abs/10.1080/09500690010016111

85. Petersson C, Simeonsson RJ, Enskar K, Huus K. Comparing children's self-report instruments for health-related quality of life using the International Classification of Functioning, Disability and Health for Children and Youth (ICF-CY). Health and Quality of Life Outcomes [Internet]. 2013 Jan;11(1). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3648353&tool=pmcentrez&rendertype=abstract

86. Simeoni M-CC, Schmidt S, Muehlan H, Debensason D, Bullinger M, Petersen C, et al. Field testing of a European quality of life instrument for children and adolescents with chronic conditions: the 37-item DISABKIDS Chronic Generic Module. Quality of Life Research [Internet]. 2007 Jun [cited 2013 Aug 17];16(5):881–93. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17404899

87. Varni JW, Limbers C a, Neighbors K, Schulz K, Lieu JEC, Heffer RW, et al. The PedsQL^{TM} Infant Scales: feasibility, internal consistency reliability, and validity in healthy and ill infants. Quality of Life Research [Internet]. 2011 Feb [cited 2014 Jul 15];20(1):45–55. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20730626

88. Varni JW, Limbers C a, Burwinkle TM. Parent proxy-report of their children's health-related quality of life: an analysis of 13,878 parents' reliability and validity across age subgroups using the PedsQL 4.0 Generic Core Scales. Health and Quality of Life Outcomes [Internet]. 2007 Jan [cited 2013 Jun 16];5(2). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1769359&tool=pmcentrez&rendertype=abstract

89. Scalone Luciana, Carlo T, Broccoli S, Cristiani, Cortesi P, Borghetti F, Selleri P CG. A Investigation of suitablility of EQ-5D-Y proxy version to assess health of children aged from 4 years. 27th Scientific Plenary Meeting of the EuroQol Group. 2011.

90. R. GNP-SMAG-DMOP. Validity and reliability of the Spanish EQ-5D-Y Proxy version. Anales de Pediatría. 2014;81(4):212–9.

91. Morrow A M; Hayen A; Quine S; Scheinberg A; Craig J C. A comparison of doctors', parents' and

children's reports of health states and health-related quality of life in children with chronic conditions. Child: Care, Health and Development [Internet]. 2012 Mar [cited 2014 Jul 20];38(2):186–95. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21651605

92. Wu R, Zhang J, Sun J, Zhou M, Wu JS, Li N, et al. Validation of the Chinese version of the Canadian Haemophilia Outcomes-Kids' Life Assessment Tool (the CHO-KLAT). Haemophilia [Internet]. 2014;20:794–9. Available from: http://doi.wiley.com/10.1111/hae.12489

93. Kaartina S, Chin YS, Fara Wahida R, Woon FC, Hiew CC, Zalilah MS, et al. Adolescent self-report and parent proxy-report of health-related quality of life: an analysis of validity and reliability of PedsQL^{TM} 4.0 among a sample of Malaysian adolescents and their parents. Health and Quality of Life Outcomes [Internet]. 2015;13(44):1–9. Available from: http://www.hqlo.com/content/13/1/44

94. Rajmil L, López AR, López-Aguilà S, Alonso J. Parent-child agreement on health-related quality of life (HRQOL): a longitudinal study. Health and Quality of Life Outcomes [Internet]. 2013 Jan [cited 2014 Jul 20];11. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3706362&tool=pmcentrez&rendertype=abstract

95. Gerharz EW, Eiser C, Woodhouse CRJ. Current approaches to assessing the quality of life in children and adolescents. British Journal of Urology. 2003;

96. Albrecht GL, Devlieger PJ. The disability paradox: high quality of life against all odds. Social Science & Medicine [Internet]. 1999 Apr;48(8):977–88. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10390038

97. von Faber M, van der Wiel AB, van Exel E, Gussekloo J, Lagaay AM, van Dongen E, Knook DL, van der Geest S WR. Successful Aging in the Oldest Old. Who Can Be Characterized as Successfully Aged? Archives Internal Medicine. 2001;161(22):2694–700.

98. Schwartz CE, Ahmed S, Sawatzky R, Sajobi T, Mayo N, Finkelstein J, et al. Guidelines for secondary analysis in search of response shift. Quality of Life Research [Internet]. 2013 Dec [cited 2014 Apr 2];22(10):2663–73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23572398

99. Schwartz CE. Applications of Response hSift Theory and Methods to Participation Measurement: A Brief History of a Young Field. Archives of Physical Medicine and Rehabilitation [Internet]. Elsevier Inc.; 2010 Sep [cited 2014 May 24];91(1):S38–43. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20801278

100. Norman G. Hi! How are you? Response shift, implicit theories and differing epistemologies. Quality of Life Research. 2003;12(3):239–49.

101. Ravens-Sieberer U, Bullinger M. Assessing health-related quality of life in chronically ill children with the German KINDL: first psychometric and content analytical results Assessing health-related quality of life in chronically ill children with the German KINDL: first psychometric an.

Quality of Life Research [Internet]. 1998;7:399–407. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=pbh&AN=11501139&site=ehost-live

102. Castelino M, Abbott J, McElhone K, Teh L-S. Comparison of the psychometric properties of health-related quality of life measures used in adults with systemic lupus erythematosus: a review of the literature. Rheumatology [Internet]. 2013 Apr [cited 2014 Jul 10];52(4):684–96. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23264550

103. DeVon HA, Block ME, Moyle-Wright P, Ernst DM, Hayden SJ, Lazzara DJ, et al. A Psychometric Toolbox for Testing Validity and Reliability. Journal of Nursing Scholarship [Internet]. 2007 Jan;39(2):155–64. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17535316

104. Reinfjell T, Diseth TH, Veenstra M, Vikan A. Measuring health-related quality of life in young adolescents: Reliability and validity in the Norwegian version of the Pediatric Quality of Life Inventory 4.0 (PedsQL) generic core scales. Health and Quality of Life Outcomes [Internet]. 2006 Jan [cited 2014 Dec 14];4(61). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1584218&tool=pmcentrez&rendertype=abstract

105. Niemeijer Anuschka S; Reindeers-Messelimk Heleen A; Disseldorp Laurien M; Nieuwenhuis Marianne K. Feasibility , Reliability , and Agreement of the WeeFIM Instrument in Dutch Children With Burns. Physical Therapy. 2012;92(7):958–66.

106. Ravens-Sieberer U, Herdman M, Devine J, Otto C, Bullinger M, Rose M, et al. The European KIDSCREEN approach to measure quality of life and well-being in children: development, current application, and future advances. Quality of Life Research [Internet]. 2014;23(3):791–803. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3953538&tool=pmcentrez&rendertype=abstract

107. Hullmann SE, Ryan JL, Ramsey RR, Chaney JM, Mullins LL. Measures of General Pediatric Quality of Life: Child Health Questionnaire (CHQ), DISABKIDS Chronic Generic Measure (DCGM), KINDL-R, Pediatric Quality of Life Inventory (PedsQL) 4.0 Generic Core Scales, and Quality of My Life Questionnaire (QoML). Arthritis Care & Research [Internet]. 2011 Nov [cited 2014 Jul 20];63(S11):S420–30. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22588762

108. Hopkins WG. Measures of reliability in sports medicine and science. Sports medicine (Auckland, NZ). 2000;30(1):1–15.

109. Johnson DR, Creech JC. Ordinal Measures in Multiple Indicator Models: A Simulation Study of Categorization Error. American Sociological Review. 1983;48(3):398–407.

110. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics. 1977;33(1):159–74.

111. Roach KE. Measurement of Health Outcomes: Reliability, Validity and Responsiveness. Journal of Prosthetics and Orthotics. 2006;18(1):8–12.

112. Pickard A. Simon; Knight Sara. J. Proxy Evaluation of Health-related Quality of Life: A Conceptual Framework for understanding Multiple Proxy Perspectives. Medical Care. 2005;43(5):493–9.

113. Saigal S, Rosenbaum P, Stoskopf B, Hoult L, Furlong W, Feeny D, et al. Development, reliability and validity of a new measure of overall health for pre-school children. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2005;14(1):243–57.

114. Kalyva E, Malakonaki E, Eiser C, Mamoulakis D. Health-related quality of life (HRQoL) of children with type 1 diabetes mellitus (T1DM): self and parental perceptions. Pediatric Diabetes [Internet]. 2011 Feb [cited 2013 Jul 27];12(1):34–40. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20546163

115. Hsieh Y, Wu C, Lin K, Chang Y, Chen C, Liu J. Responsiveness and Validity of Three Outcome Measures of Motor Function After Stroke Rehabilitation. Stroke Journal of Cerebral Circulation [Internet]. 2009 Apr [cited 2014 Jul 14];40(4):1386–91. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19228851

116. Scalone L, Tomasetto C, Matteucci MC, Selleri P, Pacelli B, Cavrini G. Assessing quality of life in children and adolescents : development and validation of the Italian version of the EQ-5D-Y. Italian Journal of Public Health. 2011;8(4):331–41.

117. Bharmal M, Thomas J. Comparing the EQ-5D and the SF-6D Descriptive Systems to Assess Their Ceiling Effects in the US General Population. Value in Health. 2006;9(4):262–71.

118. Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi J a, et al. Psychometric Evaluation and Calibration of Health-Related Quality of Life Item Banks: Plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). Medical Care. 2007;45(5):22–31.

119. Wyrwich KW, Tierney WM, Wolinsky FD. Further Evidence Supporting an SEM-Based Criterion for Identifying Meaningful Intra-Individual Changes in Health-Related Quality of Life. Journal of Clinical Epidemiology [Internet]. 1999 Sep;52(9):861–73. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0895435699000712

120. Ramos-Goñi JM, Rivero-Arias O, Errea M, Stolk E a, Herdman M, Cabasés JM. Dealing With the Health State “Dead” When Using Discrete Choice Experiments to Obtain Values For EQ-5D-5L Heath States. The European Journal of Health Economics [Internet]. 2013 Jul [cited 2014 Oct 6];14(1):S33–42. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3728441&tool=pmcentrez&rendertype=abstract

121. Varni JW, Seid M, Rode C a. The PedsQL: Measurement Model for the Pediatric Quality of life Inventory. Medical Care [Internet]. 1999 Feb;37(2):126–39. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10024117

122. Canaway AG, Frew EJ. 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. 2013 Feb;22(1):173–83.

123. Varni JW, Burwinkle TM, Seid M, Skarr D. The PedsQL 4.0 as a Pediatric Population Health Measure: Feasibility, Reliability, and Validity. Ambulatory Pediatrics [Internet]. 2003;3(6):329–41. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14616041

124. Haraldstad K, Christophersen K-A, Eide H, Nativg GK, Helseth S. Health Related Quality of Life in Children and Adolescents: Reliability and Validity of the Norwegian Version of KIDSCREEN-52 Questionnaire, A Cross Sectional Study. International Journal of Nursing Studies [Internet]. 2010 May [cited 2013 Jun 18];48(5):573–81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21067750

125. Eurich DT, Johnson J a, Reid KJ, Spertus J a. Assessing Responsiveness of Generic and Specific Health Related Quality of Life Measures in Heart Failure. Health and Quality of Life Outcomes [Internet]. 2006 Jan [cited 2014 Jul 13];4(89). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1675990&tool=pmcentrez&rendertype=abstract

126. Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PMM. On Assessing Responsiveness of Health-Related Quality of Life Instruments : Guidelines for Instrument Evaluation. Quality of Life Research. 2003;12:349–62.

127. Quittner AL, Modi AC, Wainwright C, Otto K, Kirihara J, Montgomery a B. Determination of the Minimal Clinically Important Difference Scores for the Cystic Fibrosis Questionnaire-Revised Respiratory Symptom Scale in Two Populations of Patients with Cystic Fibrosis and Chronic Pseudomonas aeruginosa Airway Infection. Chest Journal [Internet]. 2009 Jun [cited 2013 Jun 19];135(6):1610–8. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2821291&tool=pmcentrez&rendertype=abstract

128. Cohen J. Statistical Power Analysis for the Behavioral Sciences [Internet]. Statistical Power Analysis for the Behavioral Sciences. 1988. 274-287 p. Available from: http://books.google.com/books?id=Tl0N2IRAO9oC&pgis=1

129. Ottenbacher KJ, Msall ME, Lyon N, Duffy LC, Ziviani J, Granger C V, et al. The WeeFIM Instrument: Its Utility in Detecting Change in Children with Developmental Disabilities. Archives of Physical Medicine and Rehabilitation [Internet]. 2000 Oct [cited 2013 Aug 9];81(10):1317–26. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11030496

130. Pickard Simon A; Johnson Jeffrey A; Feeny David H. Responsiveness of Generic Health-Related Quality of Life Measures in Stroke. Quality of Life Research [Internet]. 2005 Feb;14(1):207–19. Available from: http://link.springer.com/10.1007/s11136-004-3928-3

131. Engel L, Bryan S, Evers SM a a, Whitehurst DGT. Exploring Psychometric Properties of the SF-6D , a Preference- Based Health-Related Quality of Life Measure , in the context of Spinal Cord Injury. Quaility of Life Research. 2014;23:2383–93.

132. Eidt-Koch D, Mittendorf T, Greiner W. Cross-Sectional Validity of the EQ-5D-Y as a Generic Health Outcome Instrument in Children and Adolescents with Cystic Fibrosis in Germany. BioMed Central Pediatrics [Internet]. 2009 Jan [cited 2013 Apr 30];9(55). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2753333&tool=pmcentrez&rendertype=abstract

133. Starfield B, Bergner M, Ensminger M, Riley A, Ryan S, Green B, et al. Adolescent Health Status Measurement: Development of the Child Health and Illness Profile. Pediatrics. 1993;91(2):430–5.

134. Landgraf JM, Maunsell E, Speechley KN, Bullinger M, Campbell S, Albetz L, et al. Canadian-French, German and UK versions of the Child Health Questionnaire: Methodology and Preliminary Item Scaling Results. Quality of Life Research. 1998;7(5):433–45.

135. Collier J, MacKinlay D, Phillips D. Norm Values for the Generic Children's Quality of Life Measure (GCQ) From a Large School-Based Sample. Quality of Life Research. 2000;9(6):617–23.

136. Le Coq Elise M., Colland Vivian T. , Boeke Joan A., Bezemer, Dick P. VEJTM. Reproducibility, Construct Validity, and Responsiveness of the "How Are You?" (HAY), a Self-Report Quality of Life Questionnaire for Children with Asthma. Journal of Asthma. 2000;37(1):43–58.

137. Ravens-Sieberer U; Bullinger M. Assessing Health-Related Quality of Life in Chronically Ill Children. Quality of Life Research. 1998;7:399–407.

138. Bjorner, Jakob B, Rose Matthias, Gandek Barbara, Stone Arthur A, Junghaenel Doerte U WJE. Methods of Administration of PROMIS. Journal Clinical Epidemiology. 2014;67(1):108–13.

139. Gong GWK, Young NL, Dempster H, Porepa M, Feldman BM. The Quality of My Life questionnaire: the minimal clinically important difference for pediatric rheumatology patients. The Journal of rheumatology. 2007;34(3):581–7.

140. Verrips, E G H, Vogels A G C, Verloove-Vanhorick S P, Fekkes M, Koopman H M, Kamphuis R P, Thuenissen N C M WJM. Health-Related Quality of Life Measure for Children - The TACQOL. Journal of Applied Therapeutics. 1997;

141. Carvalho MFC, Bloch K V, Oliveira SKF. Quality of Life of Children and Adolescents with Rheumatic Fever. Jornal de Pediatria (Brazil). 2009;85(5):438–42.

142. Rentz AM, Matza LS, Secnik, Swensen A R DA. Psychometric Validation of the Child Health Questionnaire (CHQ) in a Sample of Children and Adolescents With Attention-Deficit/Hyperactivity Disorder. Quaility of Life Research [Internet]. 2005;14(3):719–34. Available from: http://www.ncbi.nlm.nih.gov/pubmed/

143. Maddigan SL, Feeny DH, Majumdar SR, Farris KB, Johnson J a. Health Utilities Index Mark 3 Demonstrated Construct Validity in a Population-Based Sample With Type 2 Diabetes. Journal

of Clinical Epidemiology. 2006;59(5):472–7.

144. Taylor RM, Wray J, Gibson F. Measuring Quality of Life in Children and Young people After Transplantation: Methodological considerations. Pediatric Transplantation. 2010;14(10):445–58.

145. Robitail S, Ravens-Sieberer U, Simeoni MC, Rajmil L, Bruil J, Power M, et al. Testing the Structural and Cross-Cultural Validity of the KIDSCREEN-27 Quality of Life Questionnaire. Quality of Life Research. 2007;16:1335–45.

146. Bullinger M, Brütt AL, Erhart M, Ravens-Sieberer U. Psychometric Properties of the KINDL-R Questionnaire: Results of the BELLA Study. European Child and Adolescent Psychiatry. 2008;17:125–32.

147. Forrest CB, Bevans KB, Pratiwadi R, Moon J, Teneralli RE, Minton JM, et al. Development of the PROMIS Pediatric Global Health (PGH-7) Measure. Quality of Life Research. 2014;23:1221–31.

148. Kamphuis ROBP, Fekkes M, Wit JANM, Verloove-vanhorick SP. Measuring Health-Related Quality of Life in a Child Population. European Journal of Public Health. 1999;9(3):188–93.

149. Complied by Strategic Development Information and GIS Department C of CT 2011 and 2011 C data supplied by SSA. City of Cape Town – 2011 Census – Cape Town (Based on information available at the time of compilation as released by Statistics South Africa) Cape Town Overview – 2011 Census Demographic Profile – 2011 Census. 2012.

150. Varni JW, Burwinkle TM, Seid M. The PedsQL 4.0 as a School Population Health Measure: Feasibility, Reliability, and Validity. Quality of Life Research [Internet]. 2006 Mar [cited 2013 May 23];15(2):203–15. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16468077

151. Boyle SE, Jones GL WS. Quality of Life, Physical Activity, Weight Status and Diet in Adolescent School Children. Quality of Life Research. 2010;19(7):943–54.

152. Uniform Data System for Medical Rehabilitation. The WeeFIM II Clinical Guide. 2006. p. 1–182.

153. Hicks CL, von Baeyer CL, Spafford P a, van Korlaar I, Goodenough B. The Faces Pain Scale-Revised: Toward a Common Metric in Pediatric Pain Measurement. Pain [Internet]. 2001 Aug;93(2):173–83. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11427329

154. Declaration of Helsinki World - Medical Association Declaration of Helsinki. Bulletin of the World Health Organization. 2001. p. 373–4.

155. Rankin G; Stokes M. Clinical Rehabilitation. Clinical Rehabilitation. 1998. p. 187–99.

156. Jelsma J. A Comparison of the Performance of the EQ-5D and the EQ-5D-Y Health-Related Quality of Life Instruments in South African Children. International Journal of Rehabilitation

Research [Internet]. 2010 Jun [cited 2013 Apr 30];33(2):172–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20351650

157. Wu XY, Ohinmaa A, Veugelers PJ. Sociodemographic and Neighbourhood Determinants of Health-Related Quality of Life Among Grade-Five Students in Canada. Quality of Life Research [Internet]. 2010 Sep [cited 2013 May 1];19(7):969–76. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20446044

158. Burström K, Svartengren M, Egmar AC. Testing a Swedish Child-Friendly Pilot Version of the EQ-5D Instrument - Initial Results. European Journal of Public Health. 2011;21(2):178–83.

159. Thomas KS, Sach TH. A Multicentre Randomized Controlled Trial of Ion-Exchange Water Softeners for the Treatment of Eczema in Children: Protocol for the Softened Water Eczema Trial (SWET) (ISRCTN: 71423189). British Journal of Dermatology. 2008;159(3):561–6.

160. Vitale MG, Levy DE, Johnson MG, Gelijns a C, Moskowitz a J, Roye BP, et al. Assessment of Quality of Life in Adolescent Patients With Orthopaedic Problems: Are Adult Measures Appropriate? Journal of Pediatric Othopedicsrthopedics. 2001;21(5):622–8.

161. Colver AF, Merrick H, Deverill M, Le Couteur A, Parr J, Pearce MS, et al. Study Protocol: Longitudinal Study of the Transition of Young People with Complex Health Needs from Child to Adult Health Services. BioMed Central Public Health [Internet]. BMC Public Health; 2013 Jan [cited 2014 Jul 20];13(1):675–86. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3724698&tool=pmcentrez&rendertype=abstract

162. Nordyke K, Norström F, Lindholm L, Stenlund H, Rosén A, Ivarsson A. Health-Related Quality of Life in Adolescents with Screening-Detected Celiac Disease, Before and One Year After Diagnosis and Initiation of Gluten-Free Diet, A Prospective Nested Case-Referent Study. BioMed Central Public Health [Internet]. 2013 Jan [cited 2013 Apr 10];13(142). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3585471&tool=pmcentrez&rendertype=abstract

163. Burström K, Bartonek A, Broström EW, Sun S, Egmar a. C. EQ-5D-Y as a Health-Related Quality of Life Measure in Children and Adolescents With Functional Disability in Sweden: Testing Feasibility and Validity. Acta Paediatrica, International Journal of Paediatrics. 2014;103(4):426–35.

164. Jelsma J, Burgess T, Henley L. Does the Requirement of Getting Active Consent from Parents in School-Based Research Result in a Biased Sample? An Empirical Study. Journal of Empirical Research on Human Research Ethics [Internet]. 2012;7(5):56–62. Available from: http://www.jstor.org/stable/info/10.1525/jer.2012.7.5.56

165. Ras M, Human A T. The Health Related Quality of Life of Children Living with a Physical Disability Attending Physiotherapy at Meerhof School in the North West Province. Journal of Community and Health Sciences. 2014;9(2):48–61.

166. Wille N, Bullinger M, Holl R, Hoffmeister U, Mann R, Goldapp C, et al. Health-Related Quality of Life in Overweight and Obese Youths: Results of a Multicenter Study. Health and Quality of Life Outcomes [Internet]. 2010 Jan;8(36). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2868813&tool=pmcentrez&rendertype=abstract

167. Muris P, du Plessis M, Loxton H. Origins of common fears in South African children. Journal of Anxiety Disorders [Internet]. 2008 Dec [cited 2015 Jan 8];22(8):1510–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18417318

168. DeVine J, Norvell DC, Ecker E, Fourney DR, Vaccaro A, Wang J, et al. Evaluating the Correlation and Responsiveness of Patient-Reported Pain With Function and Quality-of-Life Outcomes After Spine Surgery. Spine. 2011;36(21):S69–74.

169. Craig BM, Pickard a. S, Lubetkin EI. Health Problems are More Common, But Less Severe When Measured Using Newer EQ-5D Versions. Journal of Clinical Epidemiology [Internet]. Elsevier Inc; 2014;67(1):93–9. Available from: http://dx.doi.org/10.1016/j.jclinepi.2013.07.011

170. Kind P, Klose K, Gusi N, Olivares PR, Greiner W. Can Adult Weights be Used to Value Child Health States? Testing the Influence of Perspective in Valuing EQ-5D-Y. Quality of Life Research [Internet]. Springer International Publishing; 2015; Available from: http://link.springer.com/10.1007/s11136-015-0971-1

171. Wu XY, Ohinmaa A, Johnson J a., Veugelers PJ. Assessment of Children's Own Health Status Using Visual Analogue Scale and Descriptive System of the EQ-5D-Y: Linkage Between Two Systems. Quality of Life Research. 2014;23(2):393–402.

172. Grootenhuis M a, de Boone J, van der Kooi AJ. Living with Muscular Dystrophy: Health Related Quality of Life Consequences for Children and Adults. Health and Quality of Life Outcomes. 2007;5(31).

173. Robles N, Rajmil L, Rodriguez-Arjona D, Azuara M, Codina F, Raat H, et al. Development of the Web-Based Spanish and Catalan Versions of the Euroqol 5D-Y (EQ-5D-Y) and Comparison of Results With the Paper Version. Health and Quality of Life Outcomes [Internet]. ???; 2015;13(72). Available from: http://www.hqlo.com/content/13/1/72

174. Boyaci Ahmet, Akal Ali, Tutoglu Ahmet, Kandemir Hasan, Koca Irfan, Boyraz Ismail, Celen Esra OU. Relationship among Ocular Diseases, Developmental Levels, and Clinical Characteristics of Children with Diplegic Cerebral Palsy. Journal of Physical Therapy Science. 2014;26:1679–84.

175. Grilli L, Feldman DE, Majnemer A, Couture M, Azoulay L, Swaine B. Associations Between a Functional Independence Measure (WeeFIM) and the Pediatric Quality of Life Inventory (PedsQL4.0) in Young Children With Physical Disabilities. Quality of Life Research. 2006;15(6):1023–31.

176. Feldman DE, Swaine B, Gosselin J, Meshefedjian G, Grilli L. Is Waiting for Rehabilitation Services Associated With Changes in Function and Quality of Life in Children with Physical Disabilities?

Physical & Occupational Therapy in Pediatrics. 2008;28(4):291–304.

177. Sullivan E, Barnes D, Linton JL, Calmes J, Damiano D, Oeffinger D, et al. Relationships Among Functional Outcome Measures Used For Assessing Children With Ambulatory CP. Developmental Medicine and Child Neurology. 2007;49(5):338–44.

178. Beaton DE, Bombardier C, Katz JN, Wright JG. A Taxonomy for Responsiveness. Journal of Clinical Epidemiology [Internet]. 2001 Dec;54(12):1204–17. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11750189

179. Adair B, Said CM, Rodda J, Morris ME. Psychometric Properties of Functional Mobility Tools in Hereditary Spastic Paraplegia and Other Childhood Neurological Conditions. Developmental Medicine and Child Neurology [Internet]. 2012 Jul [cited 2013 Jul 27];54(7):596–605. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22524640

180. Ottenbacher KJ, Msall ME, Lyon N, Duffy LC, Ziviani J, Granger C V, et al. The WeeFIM Instrument: Its Utility in Detecting Change in Children with Developmental Disabilities. Archives of Physical Medicine and Rehabilitation [Internet]. 2000 Oct [cited 2013 Aug 9];81(10):1317–26. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11030496

181. Varni JW, Seid M, Knight TS, Burwinkle T, Brown J, Szer IS. The PedsQL in Pediatric Rheumatology: Reliability, Validity, and Responsiveness of the Pediatric Quality of Life Inventory Generic Core Scales and Rheumatology Module. Arthritis and Rheumatism. 2002;46(3):714–25.

182. Vella S a, Cliff DP, Magee C a, Okely AD. Sports Participation and Parent-Reported Health-Related Quality of Life in Children: Longitudinal Associations. The Journal of Pediatrics [Internet]. Elsevier Ltd; 2014 Jun [cited 2014 Jul 18];164(6):1469–74. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24657117

183. Howard JS, Mattacola CG, Howell DM, Lattermann C. Response Shift Theory: An Application for Health-Related Quality of Life in Rehabilitation Research and Practice. Journal of Allied Health [Internet]. 2011 Jan;40(1):31–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21399850

184. Sullivan M. The New Subjective Medicine : Taking the Patient's Point of View on Health Care and Health Medicine. Social Science & Medicine. 2003;56:1595–604.

185. Oort FJ, Visser MRM, Sprangers M a G. Formal Definitions of measurement Bias and Explanation Bias Clarify Measurement and Conceptual Perspectives on Response Shift. Journal of Clinical Epidemiology [Internet]. Elsevier Inc; 2009 Nov [cited 2014 May 24];62(11):1126–37. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19540722

APPENDICES

Appendix 1: COSMIN checklist

Properties that have been assessed in this study are marked with x			
A. Internal consistency			x
B. Reliability			x
C. Measurement error			x
D. Content validity (including face validity)			o
Construct validity			
E. Structural validity			x
F. Hypotheses testing			x
G. Cross-cultural validity			x
H. Criterion validity			x
I. Responsiveness			x
J. Interpretability			x
A. Internal consistency			
Design requirements	Yes	No	NA
1. Was the percentage of missing items given?			x
2. Was there a description of how missing items were handled?	x		
3. Was the sample size included in the analysis adequate?	x		
Statistical methods			
4. Was the Cronbach's alpha calculated?	x		
Reliability (including test-retest, intra- and inter-rater reliability)			
Design requirements			
1. Was the percentage of missing items given?			x
2. Was there a description of how missing items were handled?	x		
3. Was the sample size included in the analysis adequate?	x		
4. Were at least two measures available?	x		
5. Were the administrations independent?	x		
6. Was the time interval stated?	x		
7. Were the patients stable in the interim period for the construct being measured?	x		
8. Was the time interval appropriate?	x		
9. Were the test conditions similar for both measurements?	x		
Statistical methods			
10. Was the intraclass correlation coefficient (ICC) calculated, for continuous scores?	x		
11. Was the kappa calculated for dichotomous scores?	x		
B. Measurement error			
Design requirements and checks were the same as for reliability.			
Statistical methods			
1. Were limits of agreement assessed	x		
D. Content validity (including face validity)			
Design requirements			
1. Was there an assessment of whether all items refer to relevant aspects of the construct being assessed?			x
E. Structural validity			
Design requirements			
1. Does the scale consists of effective indicators?			x

G. Cross-cultural validity			
Design requirements			
1. Were both the original language in which the HR- PRO instrument was developed and the language into which it was translated described?	x		
2. Were the expertise of the persons translating the measure described?	x		
3. Did the translators work independently from one another?	x		
4. Were the items translate backwards and forwards?	x		
5. Was there adequate description of how the differences between the original and translated were resolved?	x		
6. Was the translation reviewed by a committee (i.e. original developer?)	x		
7. Was the HR-PRO instrument pre-tested (cognitive interviews to check for interpretation, cultural relevance and ease of comprehension?	x		
Statistical methods			
8. Was confirmatory factor analysis performed?		x	
9. Was differential item function between language groups assessed?		x	
H. Criterion validity			
Design requirements			
1. Can the criterion used be considered as a reasonable "gold standard"	x		
G. Responsiveness			
Design requirements			
1. Was a longitudinal design of at least two measurements used?	x		
2. Was the time interval stated?	x		
3. If anything occurred in the interim period was it adequately described?	x		
4. Was a portion of patients changed (improved or deteriorated)?	x		
5. Were hypotheses about changes in score formulated a priori?	x		
Statistical methods			
1. Were design and statistical methods adequate for the hypotheses to be tested?	x		
Generalisability			
Was the sample for which the HR-PRO was evaluated adequately described in terms of:	x		
1. Mean or median age with Std Dev and range?	x		
2. Distribution of sex?	x		
3. Important disease characteristics	x		
4. Settings at which the study was conducted?	x		
5. Language in which the instrument was evaluated?	x		
7. Was the method used for selection of participants described?	x		
8. Was the percentage of missing responses acceptable?	x		

Appendix 2 Specific health conditions

Diagnosis	Count
Specific health conditions of MS participants	
Asthma	3
Asthma and Eczema	1
Attention Deficit Hyperactive Disorder	1
Developmental Coordination Disorder	2
Eyesight problems	1
High blood pressure	1
Sinusitis	1
Stomach ulcer	1
TOTAL	11
Specific health conditions of participants at the SS	
Cerebral Palsy, cerebella ataxia	1
Cerebral Palsy, flaccid quadriplegic	1
Cerebral Palsy, right hemiplegic	5
Cerebral Palsy, spastic diplegic	4
Cerebral Palsy, spastic quad	1
Congenital Muscle Disease, centronuclear myopathy	1
Duchennes Muscular Dystrophy	2
Lymphatic venous malformation, Klippel-Trenaunay Syndrome	1
Osteogenesis Imperfecta	1
Paraplegia, acute ascending flaccid paralysis	1
Spastic paraparesis, congenital	1
Spina Bifida, lumbar	7
Spina Bifida, sacral	2
Spinal Cord Injury C6-C7, incomplete	1
Spinal Cord Injury L2-L3, complete	1
Spinal Cord Injury T4-T6, complete	1
Spinal Muscular Atrophy- type 2	4
TOTAL	35
Specific health conditions of participants at the chronic care facility	
Acute Lymphoblastic Leukaemia	1
Cerebral Palsy, spastic diplegia	2
Cerebral Palsy, spastic hemiplegia	1
Chronic lung disease, oxygen dependent	1
Diabetes mellitus, type 1	8
Epilepsy	1
Failure to Thrive	3
Fanconi's Syndrome	1
Guillain–Barré Syndrome	1
Hepatitis	1
Human Immunodeficiency Virus *(*HIV)	4

Human Immunodeficiency Virus *(*HIV) and Pulmonary Tuberculosis (PTB)	4
Prada Willi Syndrome	1
Transverse Myelitis	1
Traumatic Brain Injury, ataxia	1
Tuberculous meningitis	1
TOTAL	32
Specific health conditions of participants at the acute care facility	
Acute glomerulonephritis	1
Acute Lymphoblastic Leukaemia	2
Acute pancreatitis post fall	1
Appendicitis, acute	6
Appendix, perforated	1
Arthritis, rheumatoid, right knee	1
Arthritis, septic, ankle	1
Arthritis, septic, knee	1
Asthma	1
Burkitts lymphoma	1
Burn, flame	3
Burn, hot water	2
Cardiac arrest	1
Cardiac, Mitral Valve Disease	1
Cardiac, Tetralogy of Fallot	2
Cardiac, Transposition of the Great Arteries, Ventricular Septal Defect, Patent Ductus Arteriosis	1
Caustic soda injestion	1
Complex seizure disorder	1
Craniopharngioma	2
Cyst, cervical spinal cord	1
Dog bite, thigh	1
Fractured neck of femur	1
Fractured tibia with compartment syndrome	1
Hepatitis	1
Human Immunodeficiency Virus (HIV)	2
Hypospadia	1
Idiopathic abdominal P/D	1
Motor Vehicle Accident, compound fracture left ankle with degloving of foot	1
Motor Vehicle Accident, pelvic fracture, degloving of perineum	1
Neurocystercicicosis	1
Neuropathic bladder	1
P/Dful left knee, post fall	1
Parapneumonic effusion	1
Pedestrian Vehicular Accident,ankle degloving	1
Scoliosis, idiopathic	1
Spina Bifida, sacral	1
Tuberculosis of the hip	1
Tumour, paraspinal T4-T8	1
Tumour, post fossa	1

Tumour, supratentorial	1
TOTAL	52

CONTEXTUAL INFORMATION REQUIRED FOR HRQoL RESEARCH

Please complete section 7 BEFORE administering EQ-5D-Y to child.

1. **Date of 1st administration** of EQ-5D-Y/survey 1: year/month/day.....
2. **Institution** where survey was conducted, please circle appropriate one:

Mainstream School - MS children St. Joseph's Home – CI

Astra School – SS RXH – AI

3. **Name of person** conducting the survey: .....
4. **Date of discharge**: year/month/day.....

5. DEMOGRAPHICS

Name of patient: .....

Folder number (if available): .....

Date of birth: year/month/date.....

Age: .....

Gender, please circle: male female

Present level of education: .....:

Date of admission to institution: year/month/day.....

Full Diagnosis: Primary.....

Secondary.....

5. HEALTH CHARACTERISTICS

Reason for admission:

.....

.....

.....

Medication:

.....

.....

.....

.....

Health status of child at 1st administration of EQ-5D-Y:

- Is child **acutely ill** or **chronically ill**? Please circle one
- Does child use an **assistive device** and/or **is confined to bed**, please circle?

  Yes No

  If **yes**, please circle one or more options?

  Confined to bed

  Motorised wheelchair Self-propelled wheelchair

  Walking frame elbow crutches

  Axillary crutches Orthoses

  Catheterised: indwelling or self-catheterised

  Other, please specify:

6. **EQ-5D-Y**: NOTE: This section is only to be completed by therapist if the patient is known to you.

Please complete BEFORE administering the EQ-5D-Y to the patient.
If patient is not known to you, continue with section 5.2

- As the health professional how would you rate the child's health status today with regard to the following, please tick one:

Mobility? *(walking about)*

Has no problems walking about ❑

Has some problems walking about ❑

Has a lot of problems walking about ❑

LAM?

Has no problems washing or dressing self ❑

Has some problems washing or dressing self ❑

Has a lot of problems washing or dressing myself ❑

Doing UA? *(for example, going to school, hobbies, sports, playing, doing things with family or friends)*

Has no problems doing UA ❑

Has some problems doing UA ❑

Has a lot of problems doing UA ❑

Having P/D?

Has no P/D or discomfort ❑

Has some P/D or discomfort ❑

Has a lot of P/D or discomfort ❑

Feeling WSU?

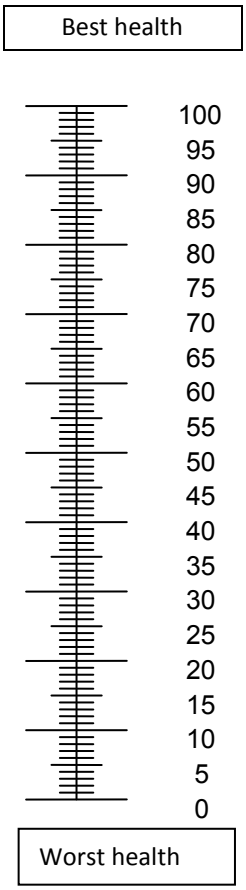
Is not WSU ❑

Is a bit WSU ❑

Is very WSU ❑

How would you rate the child's health today, if 100 is the best health imaginable and 0 the worst health imaginable.

Please mark with an X on the line to show how good or bad the child's health is TODAY



7. During the last two weeks has the child had to **cut down** on any of the things he/she usually does (for example at school or leisure) because of illness or injury?

Yes ❑ No ❑

8. In the last two weeks has the child needed to **see a doctor** for any reason? (Excluding this admission, if hospitalised)

Yes ❑ No ❑

9. **Briefly describe general management of the health condition:**

MEDICAL MANAGEMENT ……

……

……

……

SURGICAL MANAGEMENT……

……

……

……

PALLIATIVE MANAGEMENT ……

……

……

……

10. **Does the child receive physiotherapy, please circle?**

Yes No

11. **Has there been an occurrence of any life event related to the child or family in the preceding 6 months,** such as a death in family, divorce, moving house, arrival of new baby, bullying at school etc.?

Yes No

If yes, which of the above? ……

……

12. Number of visits from family members since admission

……

……

13. Does child go home for weekends, please circle?

Yes No

14. How long did it take the child to fill in the EQ-5D-Y form, in minutes?

……

……

Health status at SECOND administration of EQ-5D-Y

Date: .....

Changes in medication?:

.....

.....

.....

15. Is child **acutely ill** or **chronically ill**? Please circle one

16. Any **change** in assistive device since 1st survey, please circle?

Yes No

17. (i) EQ-5D-Y: Only complete if child is known to you, before administering the form to child, otherwise continue with (ii). As the health professional how would you rate the child's health status today with regard to the following, please tick one:

Mobility? *(walking about)*

Has no problems walking about ❑

Has some problems walking about ❑

Has a lot of problems walking about ❑

LAM?

Has no problems washing or dressing self ❑

Has some problems washing or dressing self ❑

Has a lot of problems washing or dressing myself ❑

Doing UA? *(for example, going to school, hobbies, sports, playing, doing things with family or friends)*

Has no problems doing UA ❑

Has some problems doing UA ❑

Has a lot of problems doing UA ❑

Having P/D?

Has no P/D or discomfort ❑

Has some P/D or discomfort ❑

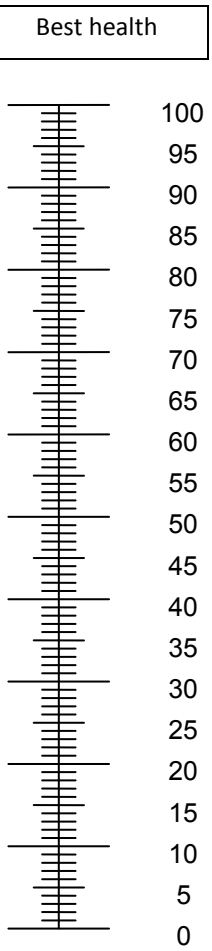
Has a lot of P/D or discomfort ❑

Feeling WSU?

Is not WSU ❑

Is a bit WSU ❑

Is very WSU ❑



How would you rate the child's health today, if 100 is the best health imaginable and 0 the worst health imaginable. Mark with an X.

18. **(ii)** Since the 1st EQ-5D-Y has the child had to **cut down** on extra things he/she usually does
(for example at school or leisure) because of illness or injury?

Yes No

- Since the 1st EQ-5D-Y has the child needed to **see a doctor** for any reason? (Excluding this admission, if hospitalised)

Yes No

Briefly describe any changes in **general management of the health condition:**

MEDICAL MANAGEMENT CHANGES:

.....

.....

.....

.....SURGICAL MANAGEMENT CHANGES:

.....

.....

.....

.....

Does the child receive physiotherapy, please circle?

Yes No

Has there been an occurrence of any life event related to the child or family since 1st EQ-5D-Y, such as a death in family, divorce, moving house, arrival of new baby, bullying at school etc.?

Yes No

If yes, which of the above? .....

.....

Number of visits from family members since 1st EQ-5D-Y

.....

.....

Does child go home for weekends, please circle?

Yes No

How long did it take the child to fill in the 2nd EQ-5D-Y form, in minutes?

.....

.....

Questionnaire on feasibility and utility of including the EQ-5D-Y form in routine assessments

Thank you for taking the time to provide your input, by completing this questionnaire.

Name of facility at which you are a physiotherapist .....

Question 1: How easy was it to implement the EQ-5D-Y measure in an assessment of a child?

- Please circle one option or make a cross next to choice:

Very easy Moderately easy

Difficult Did not use on a child

- If moderately easy or difficult, was this due to time: (please tick appropriate box/es)

Time constraints

Lack of resources

Language barrier

Lack of understanding on the child's part

IF YOU ADMINISTERED THE EQ-5D-Y TO A CHILD, PLEASE CONTINUE WITH ALL QUESTIONS

IF YOU ONLY COMPLETED A PROXY MEASURE i.e. FILLED IN THE FORM IN THE CONTEXTUAL INFO AS YOU WOULD EXPECT THE CHILD TO, THEN PROCEED TO QUESTION 4

Question 2: If you administered form to a child did the child generally encounter problems with completing the form?

- Please indicate response by circling: YES or NO

- If YES, which age group had difficulties? Please tick appropriate box

8 year olds

9 year olds

10 year olds

11 year olds

12 year olds

- If YES, in which domain/s did they have difficulties and/or with VAS?

Mobility

LAM

Doing UA

P/D

WSU

VAS

Question 3: Did you notice a relationship between child's responses on the EQ-5D-Y form and clinical signs and symptoms?

- Please indicate response by circling YES or NO
- If YES, in which dimension/dimensions in which you observed a close link to clinical signs and symptoms

Mobility

LAM

UA

P/D

WSU

VAS

TO BE COMPLETED BY ALL

Question 4: Whether you administered a form to a child or filled in a proxy, do you think using the outcome measure routinely during patient assessment might assist your planning of the management of the child?

- Please indicate response by circling YES or NO

Question 5: Do you feel that any information from a specific domain and/or VAS is the most useful?

Please indicate response by circling YES or NO

- If YES, which domain/domains/VAS were most useful

Mobility

LAM

UA

P/D

WSU

VAS

Question 7: Will the EQ-5D-Y provide you with additional information on the child's health status?

- Please indicate response by circling YES or NO

- If YES, indicate which domain/s/VAS provided new information

Mobility

LAM

UA

P/D

WSU

VAS

Question 8: Would you continue to include the EQ-5D-Y as part of your patient assessment in future?

- Please indicate response by circling YES or NO

Thank you for participating in the study and for completing the questionnaire.

You will be informed of the results

Report on the linguistic validation process in translating PedsQL4.0, Generic-Core-Child and Generic-Core-Parent Proxy, from English into Afrikaans and isiXhosa

The PedsQL questionnaire was developed in United States English and consists of instructions, items being investigated and response choices. When translating it into another language, a linguistic validation process is necessary to ensure that the original concept is maintained, so that the meaning and content of the translated version is no different to the original version. The translated version however must be easily understood by the population it is being translated for. When translating text, the translator has to be familiar with both languages in order to first read the text and mentally paraphrase it into comprehensible chunks. The next step is to work out how the words and phrases relate to one another and then determine the accurate meaning from the context. Finally the translator has to figure out how to capture that contextual meaning in the other language.

Before translating the PedsQL into Afrikaans and isiXhosa, two of the most common languages spoken in the Western Cape, apart from English, a Translation Agreement was entered into between the researcher, who was the user of the instrument and MAPI Research Trust, acting on behalf of Dr. James W. Varni, the copyright owner in the PedsQL (**Error! Reference source not found.**. Authorization was granted to the researcher to translate the English PedsQL into these two languages, for use in South Africa subject to the following conditions, all of which were acknowledged and adhered to:

"1. Dr. James W. Varni owns all copyright in the PedsQL and in all PedsQL versions including but not limited to existing and future translations of the PedsQL.

2. The researcher acknowledges Dr. James W. Varni's copyright in the PedsQL and shall not contest such copyright or perform any act or omission adverse to such exclusive right. Further the researcher acknowledges that Dr. James W. Varni holds the unfettered right to use, reproduce and exploit the aforesaid translation(s), throughout the world, for its full term without any cost or conditions.

3. The researcher agreed to the new translation undergoing a full linguistic validation process according to guidelines and recommendations that have been established in collaboration with Dr. James W. Varni as to the process to be followed in order to obtain a conceptually equivalent translation."

The following recommended methodological steps were followed:

- Forward translation
- Backward translation
- Patient testing
- Proofreading and finalization
- Report

The researcher agreed to translate both the child self-report and parent proxy-report PedsQL 4.0 Generic Core Scales for 8-12 year olds. The reason for this was that the PedsQL 4.0 Generic Core Scales use parallel child self-report and parent proxy-report forms, therefore any translations needed to include both child and parent proxy forms for the same age-group.

Procedure:

The researcher sourced local professional language translators from a database of freelance language service providers, used by the Department of Cultural Affairs & Sport, Western Cape Government.

Method:

Phase 1: Forward translation process

For this process, two translators' bilingual in Afrikaans and English and two isiXhosa and English bilingual translators, were each independently given the original US English version of the PedsQL 4.0 Generic Core Child and PedsQL4.0 Generic Core Parent-Proxy, to translate into their respective native languages. They were required to independently produce a forward translation of the original items, instructions and response choices, of the above mentioned forms. The translators all listed their field of expertise in Health, Education or Medical and had more than 10 years' experience in translating. After a discussion between both translators and the researcher, a single reconciled version of each language, namely an Afrikaans reconciled version and an isiXhosa reconciled version of the PedsQL Child-Report and PedsQL Parent-Proxy-Report was produced.

Forward translation process of Afrikaans version:

The researcher being familiar with the Afrikaans language could compare the two Afrikaans versions and found them to be similar in language use. Afrikaans does not have the continuous tense form so some accommodations had to be made for this grammatical issue. Some discrepancies were however noticed. Translator 1 had clear, specific instructions and response choice words were easier to understand and closer to the original meaning. However the wording of task items was complicated with pronouns often being left out. Translator 2 had simpler wording of task items, closer to the original meaning, but cumbersome wording of response choices.

The following issues were discussed between the researcher and the two translators:

On page one of both forms, it was agreed upon to use the shorter "kinderverslag" rather than "verslag van kinders", as they mean the same thing but "kinderverslag" is closer in meaning to the English "child report".

The word "aanwysings" has two meanings either "directions" or "instructions", so it was considered suitable.

It was decided by all that "Lewenskwaliteit" "Quality of Life" should be kept instead of "Lewensgehalte", as they mean the same thing, but "Lewenskwaliteit" is slightly closer to the original meaning.

In the instructions block, it was decided to use the same word "byna", meaning "almost" instead of introducing a new word "amper" (also meaning "almost") to be consistent.

On the items page, the word "moeilik" was kept as a translation of "hard" as the direct Afrikaans translation of "hard" would imply ' firm' or 'not yielding to pressure', which has a different contextual meaning.

The word "rondom die huis" was inserted to keep to the orginal "around the house"

Translator 2 used the word "sukkel met…" to describe "trouble with…", whereas translator 1 kept to the word for "difficult" being "moeilik". It was decided by all to use the word "sukkel met…" as a translation for "trouble with…" as it was closer to the original contextual meaning of the word.

Following the discussions a combined reconciled version, Afrikaans Version 1 was agreed on, taking the best wording from each translation, while aiming for a version as close to the original one as possible.

Forward translation process of isiXhosa version:

The linguistic process of translating the forms into isiXhosa was challenging as the researcher did not have an adequate understanding of the language and had to rely solely on the abilities of the translators. The Xhosa language, or isiXhosa, is spoken predominantly in the Eastern Cape province of South Africa and for numerous historical, social, and political reasons, it is not often used in natural

language processing (NLP) research (Xhosa-English Machine Translation: Working with a Low-Resource Language. A report by Kristine K. Johnson kkjohnson@wesleyan.edu Summer 2011).ref As a result one of the major issues in translating the language is that it is that there are very few available language resources such as dictionaries or morphological analysers to determine the internal structure of the words. There is no well-researched, carefully compiled collection of written parallel Xhosa and English text. One has to rely on the expertise of the translators.

In English, there are several ways of expressing the present tense, but not so in isiXhosa. For example: I have a problem with…. / I am having a problem with…. / I do have a problem with …. would all be translated using the same isiXhosa words.

There is also no grammatical gender in isiXhosa. For example, u- represents both him and her, or you. There are no different pronouns for different sexes. Also, ndi- which means I and u- which means you are not separate words which can stand on their own, as in English. They are always attached to the verb and they are an equivalent of the English pronoun.

Many isiXhosa phrases have more than one meaning: the literal or obvious one and a figurative or symbolic one. Often the arrangement of words in isiXhosa cannot be translated literally or directly into English and remain meaningful as the example “difficulty in walking one block” – the literal translation resulted in “difficulty in walking on top of a block”. There is no equivalent for the distance measurement of one block in isiXhosa; therefore the wording had to be changed to “difficulty in walking a short distance between houses”.

Additionally, isiXhosa is a morphologically rich language using a variety of affixes to modify or form a new word. Therefore when attempting to translate a word into isiXhosa or back from isiXhosa into English, different forms of the word may be used and still mean the same thing. This became evident after the researcher received what looked like two quite different forward translations of the same form. After discussion with the translators it became clear that both versions of the forward translation were contextually equivalent.

A reconciled version 1 in isiXhosa was agreed upon after these discussions, using the most contextually accurate phrases and phrases that would be used more extensively by isiXhosa speaking people.

Phase 2: Backward translation process

The backward translation process involved translating the first reconciled version of the Afrikaans and isiXhosa translations back into English. A third translator, bilingual in both Afrikaans/isiXhosa and English was approached and asked to translate the forms back into English, without have access to the original English version.

The backward translation of the Afrikaans version was then compared with the original English version to determine whether there were any miss-interpretations or misunderstandings in the forward translated version. In the backward translation some words were different from the original English word, but have the equivalent contextual meaning. These words are:

Original English version	Afrikaans backward translated version
“hard to…”	“difficult to…”
“chores”	“tasks”
“ache”	“P/D”
“low energy”	“little energy”
“blue”	“depressed”
“trouble”	“struggle”
“kids”	“children”
“miss school”	absent from school”

After discussions with the backwards translator, version 2 of the questionnaires was decided on

The backward translation of the isiXhosa version was compared with the original English version for comparison of items, instructions and response choices. In the backward translation some phrases were different from the original English version, but were considered to be the best contextual fit.

Original English version	isiXhosa backward translated version
"hard to…"	"difficult to…"
"difficulty in walking one block"	difficulty in walking a short distance between houses"
"ache"	"P/D"
"low energy"	"I lack energy"
"sad or blue"	"depressed"
"trouble"	"difficulty"
"kids"	"children"
"It is hard to keep up when I play with other kids"	"It is difficult to follow when I am playing with other children"
"It is hard to pay attention in the classroom"	"I can't be attentive in the classroom"
"I have trouble keeping up with my school work"	"I have problems with doing my schoolwork"

After discussions with the backward translator, version 2 of the isiXhosa questionnaires was decided on.

Report:

At this stage before the translated versions were tested on participants a report with comments was sent to Dr. James Varni, at Mapi Research www.mapigroup.com | www.mapi-trust.org for author's feedback.

On review the following comments were made:

1. "Depressed" in both the Afrikaans and isiXhosa languages was too strong. "I feel sad" for self-report or "Feels sad" for parent proxy-report would be better. "Blue" should not be translated since it is culturally specific to the US.

2. "No energy" or "lack of energy" in the isiXhosa version was too strong. "I feel tired" for self-report or "Feels tired" for parent proxy-report should be used if "low energy" was not possible.

The translator made the changes and cognitive interviewing took place on five Afrikaans and five isiXhosa speaking children between the age of eight and twelve years.

Cognitive interviews:

All five Afrikaans speaking participants ranging from eight to twelve years easily understood the form and could complete it.

However it was discovered that even though the isiXhosa children spoke their home language fluently, most of them could not read isiXhosa as they attended English or Afrikaans speaking schools and had learnt to read in one of these languages. It was possible to find five children who attended an isiXhosa school could read isiXhosa and the form was tested on them. They also easily understood the form and could complete it without difficulty. The forms were proof read and finalised. A final report was submitted to Dr Varni.

Conclusion:

The translated PedsQL forms were considered appropriate for the research participants.

NAME OF CHILD :

DATE OF ADMINISTRATION OF FORM :

Health Questionnaire

English version for South Africa

EQ-5D-Y

Describing your health TODAY

Under each heading, please tick the ONE box that best describes your health TODAY

Mobility *(walking about)*

I have no problems walking about ❑
I have some problems walking about ❑
I have a lot of problems walking about ❑

Looking After Myself

I have no problems washing or dressing myself ❑
I have some problems washing or dressing myself ❑
I have a lot of problems washing or dressing myself ❑

Doing Usual Activities *(for example, going to school, hobbies, sports, playing, doing things with family or friends)*

I have no problems doing my usual activities ❑
I have some problems doing my usual activities ❑
I have a lot of problems doing my usual activities ❑

Having Pain or Discomfort

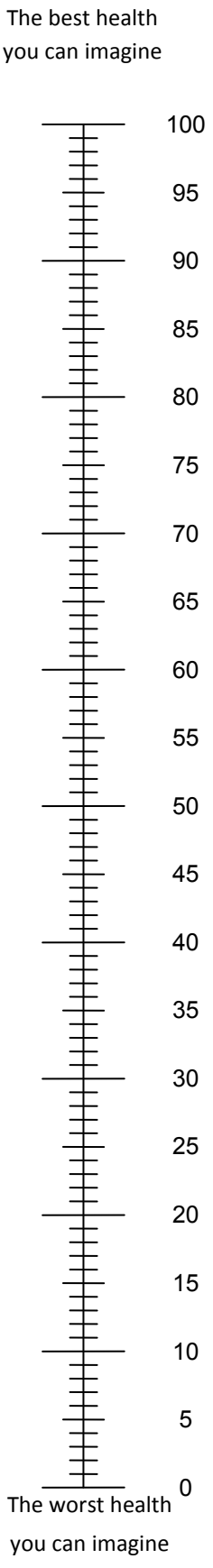
I have pain or discomfort ❑
I have some pain or discomfort ❑
I have a lot of pain or discomfort ❑

Feeling Worried, Sad or Unhappy

I am not worried, sad or unhappy ❑
I am a bit worried, sad or unhappy ❑
I am very worried, sad or unhappy ❑

How good is your health TODAY

- We would like to know how good or bad your health is TODAY.
- This line is numbered from 0 to 100.
- 100 means the <u>best</u> health you can imagine.
  0 means the <u>worst</u> health you can imagine.
- Please mark with an X on the line to show how good or bad your health is TODAY.



EQ-5D Registration Form

Dear Des,

Thank you for registering your study/trial/project or other at the EuroQol website. You sent us the following information:

Job title: **Clinical Educator**

First name: **Des**

Surname: **Scott**

Organization: **University of Cape Town**

Postal address: Faculty of Health Sciences, Department of Health and Rehabilitation Sciences, Physiotherapy Division, F45 Old Main Building, Groote Schuur Hospital, Observatory

Postal/Zip code: **7935**

City: **Cape Town**

Country: **ZA**

Telephone: **+27839498333**

E-mail: des.scott@uct.ac.za

Work environment: **Academia**

Title / Description of your study, trial, project or other: The association between Health Related Quality of Life, health condition and function in South African children.

Objective: **The overall aim of the study is investigate the HRQoL in children with a health condition and to determine associations between different health conditions, with regard to function and P/D. The ability of EQ-5D-Y measure to detect change in these children (between 8-12 years) within different contexts over a 7 month period will be assessed. The feasibility of the data collection process, using the EQ-5D-Y will be monitored, as well as examining how useful the collected data are to the therapists.**

Design: Other

Other design: A longitudinal, descriptive, analytical observational cohort study

Clinical area: **Physiotherapy**

Other clinical area: **Paediatric HRQoL**

Source of funding: **The EuroQol Group Foundation**

Number of patients / respondents: **240**

Starting date (year only): **2013**

Finishing date (year only): **2015**

Which version of the EQ-5D would you like to use? **EQ-5D-Y Paper (Youth version: 7-12 years)**

Countries: **South Africa**

Languages: **English (South Africa), Xhosa (South Africa), Sesotho (South Africa), Zulu (South Africa), Afrikaans (South Africa)**

Which other generic health measures will you use? **PedsQol, Faces Pain Scale, WeeFIM**

Which other disease / condition specific health measures will you use? **None**

Are you prepared to have this information published in any EuroQol reports/surveys regarding usage of EQ-5D? **Yes**

Are you prepared to have your details made available to colleagues who are involved in research in a similar area? **Yes**

Terms of use: **I agree with the Terms of use √**

NAME OF PARENT :

NAME OF CHILD :

DATE WHEN COMPLETING FORM :

Health Questionnaire

English version for South Africa

EQ-5D-Y

Describing <u>your child's</u> health TODAY

Under each heading, please tick the ONE box that best describes **your child's** health TODAY

Mobility *(walking about)*

I have <u>no</u> problems walking about ❑
I have <u>some</u> problems walking about ❑
I have <u>a lot</u> of problems walking about ❑

Looking After Myself

I have <u>no</u> problems washing or dressing myself ❑
I have <u>some</u> problems washing or dressing myself ❑
I have <u>a lot</u> of problems washing or dressing myself ❑

Doing Usual Activities *(for example, going to school, hobbies, sports, playing, doing things with family or friends)*

I have <u>no</u> problems doing my usual activities ❑
I have <u>some</u> problems doing my usual activities ❑
I have <u>a lot</u> of problems doing my usual activities ❑

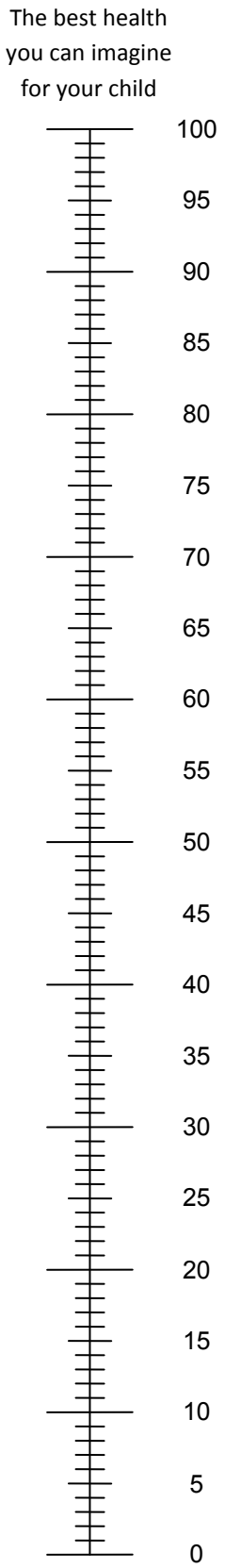
Having Pain or Discomfort

I have <u>no</u> pain or discomfort ❑
I have <u>some</u> pain or discomfort ❑
I have <u>a lot</u> of pain or discomfort ❑

Feeling Worried, Sad or Unhappy

I am <u>not</u> worried, sad or unhappy ❑
I am <u>a bit</u> worried, sad or unhappy ❑
I am <u>very</u> worried, sad or unhappy ❑

How good is your child's health TODAY



We would like to know how good or bad **your child's** health is TODAY.

This line is numbered from 0 to 100.

100 means the **best** health you can imagine for your child.

0 means the **worst** health you can imagine for your child.

Please mark with an X on the line to show how good or bad your child's health is TODAY.

ID# _____

Date:_____

PedsQL™

Pediatric Quality of Life Inventory

Version 4.0

CHILD REPORT (ages **8-12**)

DIRECTIONS

On the following page is a list of things that might be a problem for you. Please tell us **how much of a problem** each one has been for you during the **past ONE month** by circling:

0 if it is **never** a problem
1 if it is **almost never** a problem
2 if it is **sometimes** a problem
3 if it is **often** a problem
4 if it is **almost always** a problem

There are no right or wrong answers.
If you do not understand a question, please ask for help.

*In the past **ONE month**, how much of a **problem** has this been for you …*

ABOUT MY HEALTH AND ACTIVITIES *(problems with…)*	Never	Almost Never	Some-times	Often	Almost Always
1. It is hard for me to walk more than one block	0	1	2	3	4
2. It is hard for me to run	0	1	2	3	4
3. It is hard for me to do sports activity or exercise	0	1	2	3	4
4. It is hard for me to lift something heavy	0	1	2	3	4
5. It is hard for me to take a bath or shower by myself	0	1	2	3	4
6. It is hard for me to do chores around the house	0	1	2	3	4
7. I hurt or ache	0	1	2	3	4
8. I have low energy	0	1	2	3	4

ABOUT MY FEELINGS *(problems with…)*	Never	Almost Never	Some-times	Often	Almost Always
1. I feel afraid or scared	0	1	2	3	4
2. I feel sad or blue	0	1	2	3	4
3. I feel angry	0	1	2	3	4
4. I have trouble sleeping	0	1	2	3	4
5. I worry about what will happen to me	0	1	2	3	4

HOW I GET ALONG WITH OTHERS *(problems with…)*	Never	Almost Never	Some-times	Often	Almost Always
1. I have trouble getting along with other kids	0	1	2	3	4
2. Other kids do not want to be my friend	0	1	2	3	4
3. Other kids tease me	0	1	2	3	4
4. I cannot do things that other kids my age can do	0	1	2	3	4
5. It is hard to keep up when I play with other kids	0	1	2	3	4

ABOUT SCHOOL *(problems with…)*	Never	Almost Never	Some-times	Often	Almost Always
1. It is hard to pay attention in class	0	1	2	3	4
2. I forget things	0	1	2	3	4
3. I have trouble keeping up with my schoolwork	0	1	2	3	4
4. I miss school because of not feeling well	0	1	2	3	4
5. I miss school to go to the doctor or hospital	0	1	2	3	4

MAPI RESEARCH TRUST
Education Information Dissemination

PedsQL™

USER-AGREEMENT

Use of the PedsQL™ 4.0 Generic Core Scales, Modules and Translations

Date : 31/07/2013

day month year

PART 1. LICENSEE'S DETAILS

LICENSEE Name: ***Please have the information type written*** Des Scott.....

LICENSEE Title: Mrs.....

Company : University of Cape Town .....

Address : Faculty of Health Sciences .....

PBag X3 .....

Observatory, 7935 .....

Country : South Africa .....

Phone : +27214066401 / 6428..... Fax :+27214066323 .....

Email : des.scott@uct.ac.za..... .....

VAT number (if applicable): ....

2. CONTEXT OF PEDSQL USE

1. **Individual clinical practice** X☐ *(please go directly to section 4)*

- Expected duration of use: Indefinite ☐ or Number of years _____

2. <u>Mode of administration</u> **X**☐ Paper

☐ Electronic version

If electronic administration, please precise the type of medium:

- PDA ☐ - Web-based ☐
- CDr / DVD ☐ - OtherX ☐ (please precise):..IPAD.....

3. <u>Research study</u> X☐

Title: The association between Health Related Quality of Life, health condition and function in South African children.

Disease or disorder:_Multiple disorders, from acute illness e.g. pneumonia to chronic conditions such as spina bifida_____

- **Type of research**:
  - ☐ clinical trial - Phase II ☐ / Phase III ☐
  - ☐ epidemiologic/observational
  - ☐ X other: longitudinal, descriptive, analytical observational cohort study .....
- **PedsQL used as primary end point**: yes X☐ no ☐
- **Number of expected patients (total):** 240
- **Number of administrations of the questionnaire per patient**: 3
- **Length of the follow-up (if any) for each patient**: 7 months
- **Planned study date**: *start* 01 2014 (month/year) *end* 07 2014 (month/year)

3. PROJECT FINANCING

- Not funded academic research ☐
  Not funded academic research: if your project is not explicitly funded, but funding comes from overall departmental funds or from the University or individual funds then fees are waived.
- Funded academic research **X**☐
  Funded academic research: academic projects receiving funding from commerce, government, EU or registered charity should anticipate paying the corresponding fees

Note: Funded academic research sponsored by industry fits "commercial study" category

- Large non-commercial organization Research and Evaluation ☐
  (per-study license)

  Large non-commercial organization Research and Evaluation; e.g. states, nations, hospitals, healthcare systems (includes an important number of patients and/or centres)

- Large non-commercial organization Unlimited Research and Evaluation and clinical use ☐
  (annual license, unlimited use)

  Large non-commercial organization Research and Evaluation; e.g. states, nations, hospitals, healthcare systems (includes an important number of patients and/or centres)

Please specify number of centres----------

- Commercial study ☐
  Commercial studies (industry, CRO, any for-profit companies)

  Please specify number of centres----------

Granting / Sponsoring from (if any) *(name of the governmental/foundation/company or other funding/sponsoring source): EuroQol Foundation .....*

4. REQUESTED PEDSQL™ SCALES *(please tick the appropriate box(es))*

PedsQL™ Generic Core Scales *Please specify: Standard X☐ Acute ☐ Both ☐*										
Adult (over 26)		**Young Adult (18-25)**		**Adolescent (13-18)**		**Child (8-12)**		**Young Child (5-7)**		**Toddler (2-4)**
Self-report	Parent proxy-report	Self-report	Parent proxy-report	Child self-report	Parent proxy-report	Child self-report	Parent proxy-report	Child self-report	Parent proxy-report	Parent proxy-report
☐	☐	☐	☐	☐	☐	X☐	☐	☐	☐	☐

5. TRANSLATIONS

Please indicate in which language(s) and for which country(ies) the above requested PedsQL™ scale(s) is/are needed:

Language:	***For use in the following country***	***Language:***	***For use in the following country***	***Language:***	***For use in the following country***

e.g. English	*USA*				
e.g. Spanish	*USA*				
ENGLISH	SOUTH AFRICA				
AFRIKAANS	SOUTH AFRICA				
isiXHOSA	SOUTH AFRICA				

The PedsQL™ translation(s) may not be available in the country required. Please check availability of translations with Mapi Research TRUST or consult the PedsQL™ website at www.pedsql.org section "Translations".

If not available in the language(s) required, a Linguistic Validation must be undergone.

USER AGREEMENT

This agreement is between Mapi Research Trust and Des Scott ..... ("user").

Mapi Research Trust shall deliver the original **PedsQL™** and/or the translations requested by "User" subject to the following conditions:

- The translations requested are available, and
- The present contract is duly completed and signed by "User"

The use of the PedsQL™ in the above mentioned context is subject to the following conditions:

1.This user agreement is for the use of the PedsQL™, i.e., the PedsQL™ Pediatric Quality of Life Inventory™ report forms, registered copyrights in the PedsQL™ (e.g., U.S. copyright registration No. TXu 856-101) and related treaty, convention and common law rights pertaining thereto, with all rights reserved to Dr. James W. Varni, licensor and author of the PedsQL™ .

IN WITNESS WHEREOF, the parties hereto have caused this agreement to be executed by their duly authorised representatives as of the date first above written.

AGREED

User's Signature:  Title: Mrs _____ Company/Organisation: University of Cape Town _____ Date: 31/07/2013_____	**Company/Organisation Stamp (if applicable):**

FIM™ score sheet LTCS FIM score sheet Feb 08 FIM is a non-registered trademark of Uniform Data System for Medical Rehabilitation, a division of UB Foundation Activities, Inc. AROC (the Australasian Rehabilitation Outcomes Centre) holds the territorial licence for the FIM in Australia.

FIM™ score sheet

Name: Date of birth:
Date of assessment:
Hospital/unit:
Method of administration: Direct observation Interview with:

* Injury: TBI Multiple amputations Burns *(Note FIM is not required for SCI or blindness)* **Area**	**Score**	**Explain reasons for giving this score**	***Due to injury?**
SELF CARE			
1.Eating			Yes No
2.Grooming			Yes No
3.Bathing			Yes No
4.Dressing– Upper			Yes No
5.Dressing– Lower			Yes No
6.Toileting			Yes No
7.Bladder			Yes No
8.Bowel			Yes No
Self care total			
MOBILITY			
9.Transfers: Bed/ Chair/Wheelchair		Mode: W– Walk C- Wheelchair B- Both	Yes No
10.Transfers: Toilet			Yes No
11.Transfers: Bath/Shower			Yes No
12.Walk/ Wheelchair			Yes No
13.Stairs			Yes No
Mobility total			

COMMUNICATION			
14.Comprehension		Mode: A – Auditory V - Visual C - Both	Yes No
15.Expression		Mode: V – Vocal N - Non-vocal B - Both	Yes No
SOCIAL COGNITION			
16.Social interaction			Yes No
17.Problem solving			Yes No
18.Memory			Yes No
Cognition total			
FIM total			

UDSMR: Uniform Data System for Medical Rehabilitation
270 Northpointe Parkway, Suite 300
Amherst, New York 14228
Attention: Legal Services Department
Telephone: 716-817-7800
Fax: 716-568-0037

Licensee: Mrs. Des Scott
University of Cape Town
23A Rose Street
Newlands
7700
Cape Town
South Africa
Telephone: +27 (0) 21 406 6401/6428
Fax: +27 (0) 21 406 6323

IN WITNESS WHEREOF, Licensee and UDSMR have duly executed this Agreement on the date first indicated above.

UB FOUNDATION ACTIVITIES, INC.

By: _____
Edward P. Schneider, Executive Director

Date: 9/19/2013

MRS. DES SCOTT
UNIVERSITY OF CAPE TOWN

Signature: _____

Print Name: D.J. SCOTT

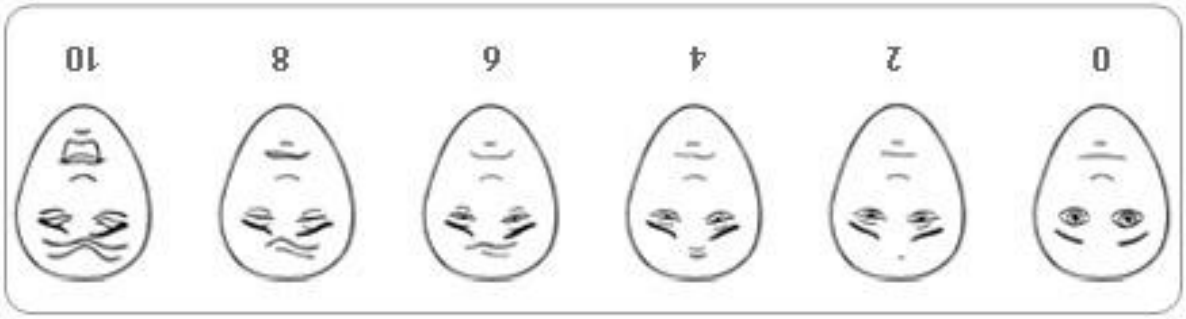
Date: 18/09/2013

Faces Pain Scale – Revised (FPS-R)
In the following instructions, say "hurt" or "pain," whichever seems right for a particular child.
"These faces show how much something can hurt. This face *[point to left-most face]* **shows no pain**.
The faces show more and more pain *[point to each from left to right]* **up to this one** *[point to right-most face]* **- it shows very much pain**.
Point to the face that shows how much you hurt [right now]**."**

*Score the chosen face **0**, **2**, **4**, **6**, **8**, or **10**, counting left to right, so '0' = 'no P/D' and '10' = 'very much P/D.' Do not use words like 'happy' and 'sad'. This scale is intended to measure how children feel inside, not how their face looks.*

Sources. Hicks CL, von Baeyer CL, Spafford P, van Korlaar I, Goodenough B. The Faces Pain Scale – Revised: Toward a common metric in pediatric P/D measurement. *P/D* 2001;93:173-183. Bieri D, Reeve R, Champion GD, Addicoat L, Ziegler J. The Faces Pain Scale for the self-assessment of the severity of P/D experienced by children: Development, initial validation and preliminary investigation for ratio scale properties. *P/D* 1990;41:139-150.
(fold along dotted line)

……



Aim of Study:

The overall aim of the pilot study was to investigate the reliability of the primary outcome measure, EQ-5D-Y, on a sample of children between 8-12 years, in the same contexts as the main study.

Specific objectives:

The specific objectives of the pilot study were to:

- Establish test-retest reliability of the EQ-5D-Y in children with a health condition and MS children by re-administering the measure the day after the initial assessment.
- Determine the feasibility of administering the EQ-5-D-Y by assessing the time taken by the children to complete the form.
- Determine the usefulness of obtaining contextual information from the medical files by assessing some of the demographic and medical information of participants.

Participants:
A sample of convenience of 40 children fulfilling the inclusion criteria of the main study and with parents readily available to sign consent, were selected. Two children, one from the SS and one from the CI were absent on the retest day, leaving a total of 38 participants.

Research settings:

All four institutions at which the main study would take place i.e. the mainstream school for MS (MS) children, the SS, the CI and the acute care hospital were included in the pilot study.

Methodology

Research design:

A descriptive, analytical study design was used.

Permission and recruitment:

Permission to perform the study was obtained from the Faculty of Health Sciences Human Research Ethics Committee (HREC REF: 354/2013) (Appendix 15). Following this, approval from all the institutions at which the study took place was obtained. Children eligible for recruitment into the pilot study were then identified by the clinical physiotherapist at the various facilities.

A parent of each child was asked to sign an informed consent form and the children were asked to sign an assent form, after the purpose of the study had been explained to them. Clinical physiotherapists interested in assisting with data collection signed a consent form and obtained parental consent from the children at the institutions', parents. This was performed by the researcher at the mainstream school.

Participant's inclusion criteria:

Approximately ten children from each institution with stable vital signs were recruited and only children who were willing and returned a signed consent form from a parent were included. Children absent on the day of retest were excluded.

Instrumentation:

The primary outcome measure of the main study, EQ-5D-Y, was used.
Time taken to complete the questionnaire was recorded
Demographic information obtained from medical files was recorded on a contextual information sheet, drawn up the researcher.

Procedure

The clinical physiotherapists were instructed on the use of the EQ-5D-Y and administered it to the identified children on day one. They recorded how long each child took to complete the form. Additional demographic information was also recorded on a contextual information sheet.

The measure was repeated the following day at all facilities.

Ethical considerations:

The same ethical considerations for informed consent, beneficence/ non-malfeasance, confidentiality and privacy which were taken for the main study were also taken in the pilot study.

Data analysis:

Descriptive statistics were used to describe the demographics of the participants and how long they took to complete the form. Frequency tables indicated how many participants were drawn from each institution and how many were diagnosed with a specific health condition.
To determine the test-retest reliability of the EQ-5D-Y in the pilot study:

- Test-retest reliability for the individual items of the EQ-5D-Y was tested by calculating Cohen's kappa coefficient to establish consistency between test and retest scores for each item.
- The correlation coefficient (ICC) for the total EQ-5D-Y score between test and retest was used to determine the strength of concordance between scores.
- The same was done for the first and second VAS score.

Results

Demographic and medical conditions of the participants:

There were 38 children who participated in the pilot study, 21 of whom were female and 38 were male. The institutions from which they were drawn are listed in Pilot study table 2 and their educational grade in Pilot study table 3

Pilot study table: 1: Institutions from which the participants came (n=38)

	Count	Percent
MS	9	23.7
SS	5	13.2
CI	9	23.7
AI	15	39.5
Total	38	100

There were more children drawn from the AI and from Grade 5 than from the other grades.

Pilot study table: 2: Level of education (grade) of participants (n=38)

Grade	Count	Percent
2	2	5.3
3	6	15.8
4	5	13.2
5	16	42.1
6	6	15.8
7	3	7.9
Total	38	100

The range of primary diagnoses of the participants are recorded in Pilot study table 4

Pilot study table: 3: Primary diagnosis of participants (n=38)

	Count	Percent
Hodgkins lymphoma	1	2.6
Diabetes Mellitus	3	7.9
Transverse Myelitis	1	2.6
Acute lymphoblastic leukemia	1	2.6
HIV+ve	2	5.3
Autoimmune hepatitis	1	2.6
Spinal muscular atrophy	1	2.6
Lymphatic venous malformation R foot	1	2.6
Spinal cord injury, MVA	1	2.6
Duchennes Muscular Dystrophy	2	5.3
MS	9	23.7
Burn	1	2.6
Caustic ingestion	1	2.6
Pneumonia meningitis	1	2.6
Congenital heart defect	1	2.6
Caustic ingestion at 18mnths	1	2.6
Burkitts lymphoma	1	2.6

Appendectomy	1	2.6
Rheumatic fever	1	2.6
traumatic eye injury	1	2.6
PVA	2	5.3
Sclerosing cholangitis	1	2.6
bowel obstruction	1	2.6
Menstrual P/D, vomiting	1	2.6
Spina bifida	1	2.6
Missing	0	0.0

First test and second re-test scores of EQ-5D-Y

The scores for each item of the EQ-5D-Y in the first (test) and second (retest) assessments are shown in Pilot study table 5.

Pilot study table: 4: First and second EQ-5D-Y dimension scores (n=38)

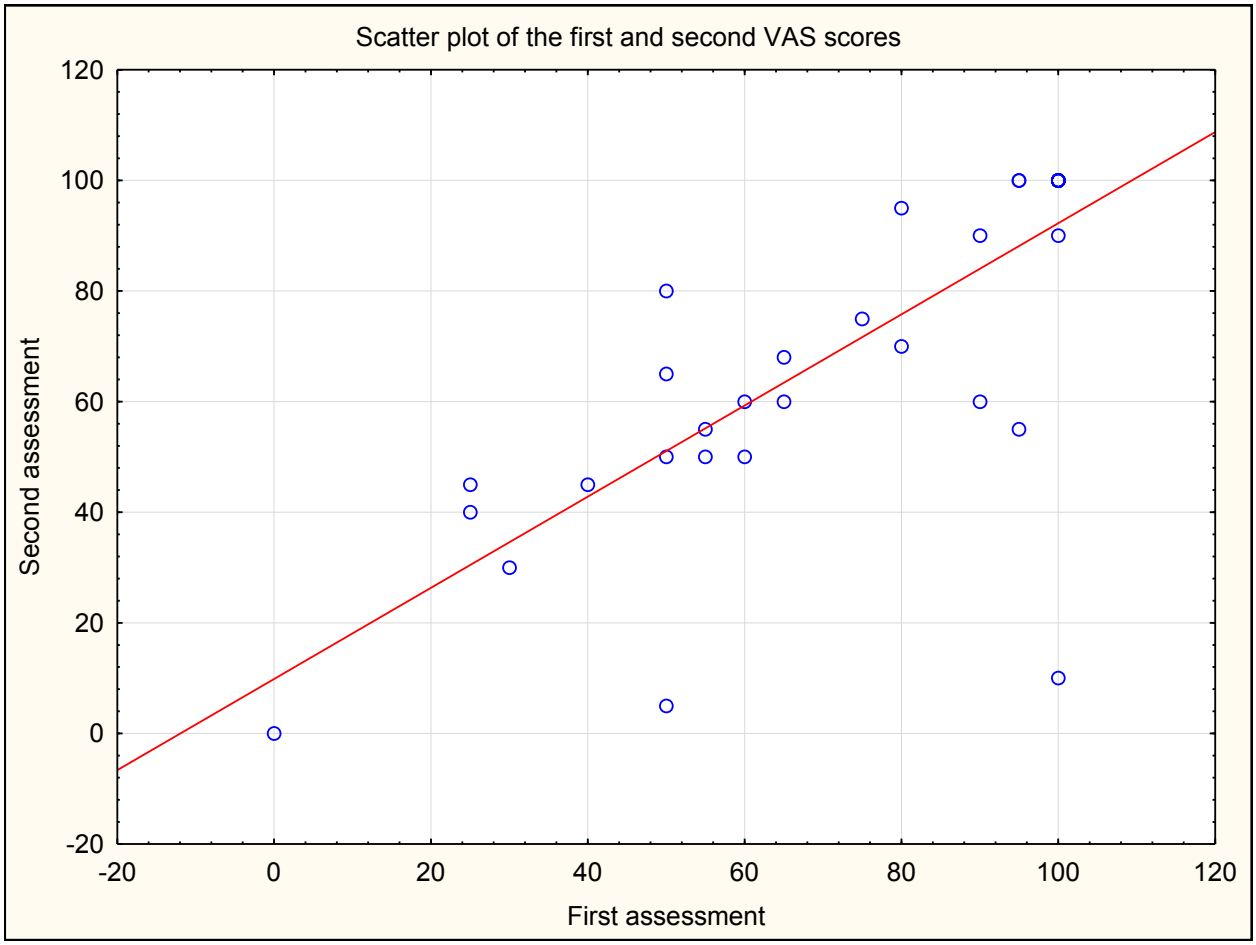
Mobility	**Second assessment**	**No problems**	**Some problems**	**Lots of problems**	**Row total**	**Kappa**	**App p**
First assessment							
No problems		20	4	1	25	0.546	p<.001
Some Problems		1	4	0	5		
Lots of problems		1	2	4	7		
Totals		22	10	5	37		
LAM	**Second assessment**						
First assessment							
No problems		25	0	0	25	0.653	p<.001
Some Problems		3	6	0	9		
Lots of problems		1	2	1	4		
Totals		29	8	1	38		

UA	**Second assessment**						
First assessment							
No problems		19	3	3	25	0.199	p<.127
Some Problems		6	4	1	11		
Lots of problems		0	2	0	2		
Totals		25	9	4	38		
P/D	**Second assessment**						
First assessment							
No problems		15	3	18		0.365	p<.0.08
Some Problems		6	9	15			
Lots of problems		2	2	4			
Totals		23	14	37			
Worried	**Second assessment**						
First assessment						0.551	p<.001
No problems		18	2	0	20		
Some Problems		4	10	0	14		
Lots of problems		1	1	1	3		
Totals		23	13	1	37		

The intraclass correlation (ICC) for absolute agreement for single measures for summed dimension scores (total score/misery index) was 0.662 (95% Confidence intervals (CIs) =0.437-0.809) which indicates good agreement between the two scores

The mean scores on the VAS were 75.8(SD=27.9) and 72.3 (SD=30) for the first and second assessments respectively.

The intraclass correlation (ICC) for absolute agreement for single measures for VAS scores was 0.765 (95% Confidence intervals (CIs) =0.594-0.870) which indicates strong agreement between the two VAS scores. As can be seen in figure 1, apart from two outliers, the first and second scores of the children were similar.



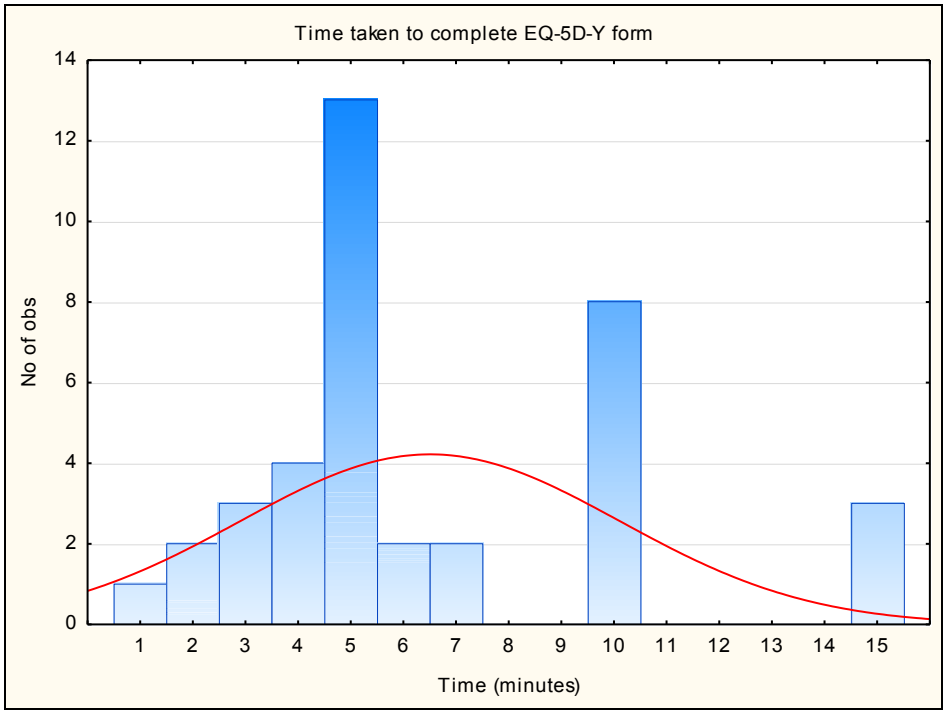
Pilot study figure: 1: Test-retest EQ-5D-Y VAS score

Time taken to complete EQ-5D-Y form

Pilot study table: 5: Time, in minutes, taken to complete EQ-5D-Y form (n=38)

Mean	Std.Dev	Minimum	Maximum	n
6.47	3.59	1	15	38

The mean time taken to complete the measure was 6.5 minutes as can be seen in Pilot study table 6. Thirteen children took 5 minutes to complete the form as can be seen in Pilot study figure 3 with eight children taking 10 minutes and three children taking 15 minutes to complete the form.



Pilot study figure: 2: Time taken to complete EQ-5D-Y, in minutes

Discussion

From the results of the pilot study it would seem that the study is feasible in that a sample of convenience easily elicited a sufficient number of participants within the eight to twelve year age group, from all four institutions, from whom it was possible to obtain demographic and medical information.

Informed signed consent was obtained from thirty eight children, ranging from grade two to grade seven educational levels. The two children in grade two should have been excluded from the study as they were not quite eight years old. The upper level of the inclusion criteria needs to be amended to include twelve year old children who are in grade seven, as there were three children in this category. Sixteen of the children were at a grade five educational level. There were seventeen more male than female children. More children, a total of fifteen, from the AI were recruited into the study than from the other facilities and the fewest, only five children from the SS. This institution was the most difficult to obtain parental consent from as most of the children board at the school and contacting parents is difficult. There were twenty four different health conditions in the participants, varying from acute to chronic and also MS children.

Percentage of agreement between the test and retest scores for each item of the EQ-5D-Y was determined using Cohen's kappa coefficient. According to Landis and Koch's guidelines (110), kappa of < 0.2 indicates poor agreement, 0.21-0.40 fair agreement, 0.41- 0.60 moderate agreement, 0.61- 0.80 substantial agreement.

The kappa for mobility 0.546, self-care 0.53 and worried 0.551, were all within the moderately agreeable range with a p value of < 0.001. The P/D item with a kappa of 0.365 was fairly agreeable with a p value of <0.08. However UA just fell into the poor agreement category, with a kappa of 0.199 and p value of <0.127

For the first and second test VAS scores, Intraclass Correlation Coefficient(ICC) was found to be 0.765 (95% Confidence intervals (CIs) =0.594-0.870) which indicated good agreement between the two assessment scores (155)

Conclusion

In summary the EQ-5D-Y is a reliable measure to use as the primary outcome measure in assessing HRQoL in the participants for the study. It is feasible as it only takes about six minutes to complete. It is possible to obtain useful contextual information on each participant from the medical file, which could be analysed further with information from the outcome measure. The inclusion criteria need to be amended to include children in grade seven, who fulfil the other criteria. More time needs to be allowed for obtaining consent from parents with children at the SS.

Appendix 15: Ethical Approval for the study

UNIVERSITY OF CAPE TOWN

Faculty of Health Sciences

Human Research Ethics Committee

Room E52-24 Old Main Building
Groote Schuur Hospital
Observatory 7925
Telephone [021] 406 6338 • **Facsimile** [021] 406 6411
Email: linsey.samuels@uct.ac.za
Website: www.health.uct.ac.za/research/humanethics/forms

21 October 2013

HREC REF: 625/2013

Prof J Jelsma
Physiotherapy
Health and Rehabilitation Sciences
F46, OMB

Dear Prof Jelsma

PROJECT TITLE: THE ASSOCIATION BETWEEN HEALTH RELATED QUALITY OF LIFE, HEALTH CONDITION AND FUNCTION IN SOUTH AFRICAN CHILDREN

Thank you for submitting your new study to the Faculty of Health Sciences Human Research Ethics Committee, received 21st October 2013.

It is a pleasure to inform you that the HREC has **formally approved** the above-mentioned study.

Approval is granted for one year until the 30th October 2014

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period.
(Forms can be found on our website: www.health.uct.ac.za/research/humanethics/forms)

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please quote the HREC. REF in all your correspondence.

Yours sincerely

PROFESSOR M BLOCKMAN
CHAIRPERSON, FHS HUMAN ETHICS
Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938
This serves to confirm that the University of Cape Town Human Research Ethics Committee complies to the Ethics Standards for Clinical Research with a new drug in patients, based on the Medical Research Council (MRC-SA), Food and Drug Administration (FDA-USA), International Convention on Harmonisation Good Clinical Practice (ICH GCP) and Declaration of Helsinki guidelines.
The Human Research Ethics Committee granting this approval is in compliance with the ICH Harmonised Tripartite Guidelines E6: Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95) and FDA Code Federal Regulation Part 50, 56 and 312.

UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Department of Health and Rehabilitation Sciences
Divisions of Communication Sciences and Disorders; Nursing and Midwifery; Occupational Therapy; Physiotherapy; Disability Studies
F45 Old Main Building, Groote Schuur Hospital
Observatory, Cape Town, W Cape, 7925
Tel: +27 (0) 21 406 6401/ 6428/ 6628/ 6534
Fax: +27 (0) 21 406 6323

Authorisation from institution

Investigator: Des Scott (B.Sc. Physiotherapy)
University of Cape Town
Department of Health and Rehabilitation Sciences
Division of physiotherapy

Supervisor: Professor Jennifer Jelsma
University of Cape Town
Department of Health and Rehabilitation Sciences
Division of physiotherapy

Title of study: The association between Health Related Quality of Life, health condition and function in South African children.

I am a physiotherapist from the University of Cape Town, conducting a research project investigating quality of life in children with a health condition, for a master's degree. I am requesting your permission to conduct part of this study at your institution.

The aim of the study is to investigate whether and how children's health conditions affect their daily life, physically, psychologically and socially.

This will require children, between eight and twelve years, from your institution, completing two short Health Related Quality of Life forms, namely the EQ-5D-Y and PedsQL, as well as a P/D measure, Faces Pain Scale, all of which are self-reporting and will take approximately ten minutes each to complete.

Their functional ability will also be recorded. These forms will be administered by the physiotherapists in conjunction with their usual assessment method and by myself, after parental consent has been obtained. Assent will also be sought from each child. The physiotherapists have been approached and are willing to participate in the study. The procedure will be explained in detail to them, by the researcher. Repeat measures will need to be taken three times over a seven month period, from January 2014 to July 2014.

There are no risks to the children participating in this study, nor will it affect any treatment the child may be receiving. The research could result in a better understanding of HRQoL in children within the institution. This information could be incorporated into guidelines regarding the use of the EQ-5D-Y as an instrument that can be used on a routine basis to improve the management of the child.

All information obtained will be confidential and keep in a secure place and only used for the purposes of this study. The information will only be made available to the research personnel.

Should you have any questions or concerns about the study you may contact the researcher or supervisor.

RESEARCHER
Des Scott at:
University of Cape Town
Department of Health and Rehabilitation Sciences
Division of physiotherapy
F45 Old Main Building
Groote Schuur Hospital
Observatory
Tel: 021-4066401/6628
Fax: 021 4066323
Cell: 083 949 8333

SUPERVISOR
Professor J. Jelsma at:
University of Cape Town
Department of Health and Rehabilitation Sciences
Division of physiotherapy
F45 Old Main Building
Groote Schuur Hospital
Observatory
Tel: 021-4066401/6628/6595
Fax: 021 4066323

Or

Professor M. Blockman at:
University of Cape Town
Faculty of Health Sciences Human Research Ethics committee
Tel: 021 406 6492
Fax: 021 406 6411
Room: E52.24 Old Main Building

CONSENT FORM

Declaration	**Yes**	**No**
I give consent from the institution allowing the researchers to conduct the above mentioned study at..... (Name of institution).		

Signed: _____ _____/_____/_____
Authorising signature Date

_____ _____/_____/_____
Researcher Date

UNIVERSITY OF CAPE TOWN

Faculty of Health Sciences

Department of Health and Rehabilitation Sciences

Divisions of Communication Sciences and Disorders; Nursing and Midwifery; Occupational Therapy; Physiotherapy; Disability Studies

F45 Old Main Building, Groote Schuur Hospital
Observatory, Cape Town, W Cape, 7925
Tel: +27 (0) 21 406 6401/ 6428/ 6628/ 6534
Fax: +27 (0) 21 406 6323

Informed Consent Form

INFORMED CONSENT FOR THERAPISTS TAKING PART IN THE STUDY

INFORMATION SHEET

Investigator: Des Scott – BSc Physiotherapy
University of Cape Town
Department of Health and Rehabilitation Sciences
Division of physiotherapy

Title of study: The association between Health Related Quality of Life, health condition and function in South African children.

I am a physiotherapist from the University of Cape Town and am conducting a research project investigating quality of life in children with a health condition, for a Master's Degree. I am interested in finding out in what way a child's health condition affects his/her ability to move, look after him/her self, go about his/her UA, P/D or feels WSU.

I am inviting you to participate in the study by administering the EQ-5D-Y, a health related quality of life (HRQoL) outcome measure to your patients, which will take approximately five minutes and to complete a contextual information sheet on the child. The study will continue for seven months. At the conclusion of the study I would like you to complete a questionnaire on whether the data gathered when using the EQ-5D-Y were useful and whether you intend to include the outcome measure as part of your routine assessment tools in the future.

You are not obliged to participate, it is voluntary and the choice is yours alone. The study will not affect the management of patients, which will continue as usual.
There are no risks or benefits to participating in this study, but you might find the use of the EQ-5D-Y useful in planning the management of your patients.

Should you have any questions or concerns about the study you may contact the researcher or the supervisor:

RESEARCHER
Des Scott at:
University of Cape Town
Department of Health and Rehabilitation Sciences
Division of physiotherapy
F45 Old Main Building
Groote Schuur Hospital
Observatory
Tel: 021-4066401/6628 **Cell:** 083 949 8333
Fax: 021 4066323

SUPERVISOR
Professor J. Jelsma at:
University of Cape Town
Department of Health and Rehabilitation Sciences
Division of physiotherapy
F45 Old Main Building
Groote Schuur Hospital
Observatory
Tel: 021-4066401/6628/6595
Fax: 021 4066323

Or
Professor M. Blockman at:
The University of Cape Town
Faculty of Health Sciences Human Research Ethics committee
Tel: 021 406 6492
Fax: 021 406 6411
Room: E52.24 Old Main Building

CONSENT FORM

Declaration	**Yes**	**No**
I have read through the information sheet and understand it's content		
I understand that my consent is required		
I understand that participation is voluntary and I can withhold my consent without any consequences		
I understand that I will not be personally identified should this research study be published		
I consent to participating in this research study of my own free will		

Signed: _____ _____/_____/_____

Participant

_____ _____/_____/_____

Researcher

UNIVERSITY OF CAPE TOWN

Faculty of Health Sciences

Department of Health and Rehabilitation Sciences

Divisions of Communication Sciences and Disorders; Nursing and Midwifery; Occupational Therapy; Physiotherapy; Disability Studies

F45 Old Main Building, Groote Schuur Hospital
Observatory, Cape Town, W Cape, 7925
Tel: +27 (0) 21 406 6401/ 6428/ 6628/ 6534
Fax: +27 (0) 21 406 6323

INFORMED CONSENT FROM PARENTS OR LEGAL GUARDIANS OF CHILDREN TAKING PART IN THE STUDY

INFORMATION SHEET

Investigator: Des Scott – BSc Physiotherapy
University of Cape Town
Department of Health and Rehabilitation Sciences
Division of physiotherapy

Title of study: The association between Health Related Quality of Life, health condition and function in South African children.

I am a physiotherapist from the University of Cape Town and am conducting a research project investigating quality of life in children with a health condition, for a Master's Degree. I am asking you to allow the physiotherapists and myself to conduct interviews with your child asking specific questions about his/her quality of life.

I am interested in finding out in what way your child's health condition affects his/her everyday life and to follow them up for seven months to see what changes occur in their health condition and what it is related to. In order to do this the physiotherapist will ask your child to fill in a short form asking about his/her ability to move, look after him/her self, go about his/her UA, P/D, feels WSU. This should take about ten minutes. The physiotherapist will also ask you and /or the nurses about anything else that may be happening in the child's life, affecting his/her condition. I will then be asking your child to fill in two other short forms asking the same type of questions and about his/her P/D. I will also be looking at how your child moves about. This should take about half an hour. We will be following your child's health condition for seven months and will be asking him/her to fill these short forms in again at three months and at seven months. All children between 8 and 12 years at this facility and others are being invited to participate.

Please understand that you do not have to give consent allowing us to interview your child. It is voluntary and the choice is yours alone. If you choose not to consent, there will be no negative consequence to you or your child. If you decide to consent now, but at a later date decide that you no longer want your child to be part of the study, you may withdraw your consent and your child's information will be removed from the study. However we would be grateful if you would assist us by allowing us to interview your child.

By allowing your child to be part of the study, the treatment he/she is receiving will not be affected at all. This will continue as usual.

All information you or your child gives us will be confidential and only the researcher and the child's therapist will see the information and know the name of your child, which will be removed once your child has completed all the forms. The completed forms will be kept in a file in a locked cupboard in the physiotherapy department until the information has been recorded onto a secure computer in a password protected file. The information will only be used for the purposes of this study. When the study is written up your child's name will not be used.

There are no risks to participating in this study. There are also no financial benefits to participating. I will be comparing the information your child gives us, with other children who are ill, in the hope of improving the management of children with a health condition.

Should you have any questions or concerns about the study you may contact the researcher or the supervisor:

RESEARCHER
Des Scott at:
University of Cape Town
Department of Health and Rehabilitation Sciences
Division of physiotherapy
F45 Old Main Building
Groote Schuur Hospital
Observatory
Tel: 021-4066401/6628
Fax: 021 4066323
Cell: 083 949 8333

SUPERVISOR
Professor J. Jelsma at:
University of Cape Town
Department of Health and Rehabilitation Sciences
Division of physiotherapy
F45 Old Main Building
Groote Schuur Hospital
Observatory
Tel: 021-4066401/6628/6595
Fax: 021 4066323

Or if you have any questions about your rights or welfare as a parent of a participant please contact

Professor M. Blockman at:
The University of Cape Town
Faculty of Health Sciences Human Research Ethics committee
Tel: 021 406 6492
Fax: 021 406 6411
Room: E52.24 Old Main Building

CONSENT FORM

Declaration	**Yes**	**No**
I have read through the information sheet and understand it's content		
I understand that my consent is required		
I understand that participation is voluntary and I can withhold my consent without any consequences to myself or my child		
I understand that refusal to give consent will not affect the current or future health care of my child.		
I understand that I nor my child will be personally identified should this research study be published		
I consent to my child participating in this research study of my own free will		

Signed:

_____ _____/_____/_____

Parent's signature Date

_____

Child's name

_____ _____/_____/_____

Researcher/Physiotherapist Date

UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Department of Health and Rehabilitation Sciences

Divisions of Communication Sciences and Disorders; Nursing and Midwifery; Occupational Therapy; Physiotherapy; Disability Studies

F45 Old Main Building, Groote Schuur Hospital
Observatory, Cape Town, W Cape, 7925
Tel: +27 (0) 21 406 6401/ 6428/ 6628/ 6534
Fax: +27 (0) 21 406 6323

INFORMED ASSENT FROM CHILD TAKING PART IN THE STUDY

INFORMATION SHEET

Investigator: Des Scott – BSc Physiotherapy
University of Cape Town
Department of Health and Rehabilitation Sciences
Division of physiotherapy

Title of study: The association between Health Related Quality of Life, health condition and function in South African children

I am a physiotherapist from the University of Cape Town and am trying to learn more about how you would describe your health and your satisfaction with your life. I would like you to think about being in my study. I am inviting other children between 8 and 12 years, at this place and at other places to join in, if they want to.

If you decide you want to be in my study, you will be asked by the physiotherapist to fill in a short form asking you how easily you move about, look after yourself, go about your UA and whether you feel P/D or feel WSU. This should take about ten minutes to fill in. The physiotherapist will also ask your parents and/or the nurses if anything is changing in your life. I will have another two forms for

you to fill in later the same day, asking similar questions and about how much P/D you might have. I will also be watching how you move about. This will take about half an hour.

I am interested in following up whether your health changes over seven months, so you will be asked to fill in these three short forms again three months later and again three months after that.

By choosing to be in the study I may learn from you how to improve the way we take care of you.

Other people will not know if you are in the study. I will put what we learn about you together with things we learn about other children, so no one can tell what things came from you. When I tell other people about the research study, I will not use your name, so no one can tell who we are talking about.

Your parents or guardian have to say it's OK for you to be in the study. After they decide, you get to choose if you want to do it too. If you don't want to be in the study, no one will be angry at you. If you want to be in the study now and change your mind later, that's OK. You can stop at any time.

You can call me or my supervisor if you have any questions about the study:

RESEARCHER
Des Scott at:
University of Cape Town
Department of Health and Rehabilitation Sciences
Division of physiotherapy
F45 Old Main Building
Groote Schuur Hospital
Observatory
Tel: 021-4066401/6628
Fax: 021 4066323
Cell: 083 949 8333

SUPERVISOR
Professor J. Jelsma at:
University of Cape Town
Department of Health and Rehabilitation Sciences
Division of physiotherapy
F45 Old Main Building
Groote Schuur Hospital
Observatory
Tel: 021-4066401/6628/6595
Fax: 021 4066323

Or if you have any questions about your rights or welfare as someone taking part in the study, you can contact **Professor M. Blockman** at:

The University of Cape Town

Faculty of Health Sciences Human Research Ethics committee

Tel: 021 406 6492

Fax: 021 406 6411

Room: E52.24 Old Main Building

Agreement

Declaration	**Yes**	**No**
I have read through the information sheet and understand it.		
I know that I don't have to take part in this study if I don't want to.		
I know that my name will not be used when you tell people about the study		
I have decided I want to be in the study		

_____ _____

Signature of Study Participant (child) Date

_____

Name of child

_____ _____

Researcher/Physiotherapist Date

UNIVERSITY OF CAPE TOWN

Faculty of Health Sciences

Department of Health and Rehabilitation Sciences

Divisions of Communication Sciences and Disorders; Nursing and Midwifery; Occupational Therapy; Physiotherapy; Disability Studies

F45 Old Main Building, Groote Schuur Hospital
Observatory, Cape Town, W Cape, 7925
Tel: +27 (0) 21 406 6401/ 6428/ 6628/ 6534
Fax: +27 (0) 21 406 6323

INFORMED CONSENT FROM PARENTS OR LEGAL GUARDIANS OF CHILDREN AT A SCHOOL ALREADY TAKING PART IN THE STUDY TO:

- **COMPLETE AN EQ-5D-Y PROXY FORM**

INFORMATION SHEET

Investigator: Des Scott – BSc Physiotherapy
University of Cape Town
Department of Health and Rehabilitation Sciences
Division of physiotherapy

Title of study: The association between Health Related Quality of Life, health condition and function in South African children.

I am the physiotherapist from the University of Cape Town conducting the research project investigating quality of life in children, for a Master's Degree. You have already kindly given your consent, allowing me to conduct interviews with your child asking specific questions about his/her quality of life. I completed the first set of interviews in February this year.

As I am preparing for the second set of interviews, I am also interested in finding out the relationship between **your child's** reported Health Related Quality of Life and what **you** think of your child's Health Related Quality of Life.

In order to do this I am asking you to fill in a short form asking what **you think** about your child's ability to move, look after him/her self, go about his/her UA, P/D, feels WSU. This should take about ten minutes. I would be grateful if you would assist me by completing the form and allowing me to use this information in my research study.

The information on the form will be confidential and only I will see it. Your name and your child's name will be removed once all the forms have been completed. The completed forms will be kept in a

file in a locked cupboard in the physiotherapy department until the information has been recorded onto a secure computer in a password protected file. The information will only be used for the purposes of this study. When the study is written up neither your name nor your child's name will be used.

Please understand that you do not have to consent to completing this form. It is voluntary and the choice is yours alone. If you choose not to consent, there will be no negative consequence to you or your child. If you decide to consent now, but at a later date decide that you no longer want your information to be part of the study, you may withdraw your consent and this information will be removed from the study.

There are no risks to participating in this study. There are also no financial benefits to participating.

Should you have any questions or concerns about the study you may contact the researcher or the supervisor:

RESEARCHER
Des Scott at:
University of Cape Town
Department of Health and Rehabilitation Sciences
Division of physiotherapy
F45 Old Main Building
Groote Schuur Hospital
Observatory
Tel: 021-4066401/6628
Fax: 021 4066323
Cell: 083 949 8333

SUPERVISOR
Professor J. Jelsma at:
University of Cape Town
Department of Health and Rehabilitation Sciences
Division of physiotherapy
F45 Old Main Building
Groote Schuur Hospital
Observatory
Tel: 021-4066401/6628/6595
Fax: 021 4066323

Or if you have any questions about your rights or welfare as a parent of a participant please contact
Professor M. Blockman at:
The University of Cape Town
Faculty of Health Sciences Human Research Ethics committee
Tel: 021 406 6492
Fax: 021 406 6411
Room: E52.24 Old Main Building

CONSENT FORM

Declaration	Yes	No
I have read through the information sheet and understand it's content		
I understand that my consent is required		
I understand that participation is voluntary and I can withhold my consent without any consequences to myself or my child		
I understand that I nor my child will be personally identified should this research study be published		
I consent to participating in this research study of my own free will		

Signed:

_____ _____/_____/_____

Parent's signature

_____

Child's name

_____ _____/_____/_____

Researcher/Physiotherapist

		MS (n=105) count	%	SS (n=35) count	%	CI (n=32) count	%	AI (n=52) count	%
ACTIVITY: Hard to walk more than one block	Never a problem	85	81.0	11	31.4	17	53.1	18	34.6
	Almost never a problem	2	1.9	0	0.0	1	3.1	2	3.8
	Sometimes a problem	10	9.5	5	14.3	6	18.8	7	13.5
	Often a problem	1	1.0	2	5.7	1	3.1	5	9.6
	Almost always a problem	7	6.7	17	48.6	6	18.8	20	38.5
Hard to run	Never a problem	62	59.0	6	17.1	17	53.1	13	25.0
	Almost never a problem	18	17.1	0	0.0	1	3.1	0	0.0
	Sometimes a problem	20	19.0	7	20.0	6	18.8	11	21.2
	Often a problem	2	1.9	1	2.9	0	0.0	4	7.7
	Almost always a problem	3	2.9	21	60.0	7	21.9	24	46.2
Hard to do sport or exercise	Never a problem	72	68.6	9	25.7	12	37.5	12	23.1
	Almost never a problem	11	10.5	0	0.0	1	3.1	1	1.9
	Sometimes a problem	12	11.4	19	54.3	11	34.4	10	19.2
	Often a problem	4	3.8	1	2.9	0	0.0	5	9.6
	Almost always a problem	6	5.7	6	17.1	7	21.9	24	46.2
Hard to lift something heavy	Never a problem	38	36.2	7	20.0	6	18.8	11	21.2
	Almost never a problem	13	12.4	1	2.9	2	6.3	2	3.8
	Sometimes a problem	31	29.5	13	37.1	11	34.4	14	26.9
	Often a problem	3	2.9	0	0.0	4	12.5	7	13.5
	Almost always a problem	20	19.0	14	40.0	8	25.0	18	34.6
Hard to bath or shower by myself	Never a problem	94	89.5	22	62.9	24	75.0	28	53.8
	Almost never a problem	2	1.9	0	0.0	0	0.0	0	0.0
	Sometimes a problem	4	3.8	7	20.0	3	9.4	9	17.3
	Often a problem	1	1.0	0	0.0	1	3.1	8	15.4
	Almost always a problem	4	3.8	6	17.1	3	9.4	7	13.5
Hard to do chores around house	Never a problem	74	70.5	15	42.9	14	43.8	19	36.5

	Almost never a problem	4	3.8	0	0.0	4	12.5	3	5.8
	Sometimes a problem	15	14.3	14	40.0	7	21.9	6	11.5
	Often a problem	3	2.9	1	2.9	1	3.1	5	9.6
	Almost always a problem	9	8.6	5	14.3	5	15.6	19	36.5
I hurt	Never a problem	35	33.3	15	42.9	6	18.8	19	36.5
	Almost never a problem	13	12.4	0	0.0	2	6.3	6	11.5
	Sometimes a problem	40	38.1	17	48.6	16	50.0	11	21.2
	Often a problem	5	4.8	1	2.9	6	18.8	3	5.8
	Almost always a problem	12	11.4	2	5.7	1	3.1	13	25.0
Have low energy	Never a problem	46	43.8	14	40.0	12	37.5	17	32.7
	Almost never a problem	11	10.5	1	2.9	1	3.1	1	1.9
	Sometimes a problem	34	32.4	10	28.6	12	37.5	13	25.0
	Often a problem	2	1.9	3	8.6	2	6.3	10	19.2
	Almost always a problem	12	11.4	7	20.0	4	12.5	11	21.2
FEELINGS: Feel afraid or scared	Never a problem	44	41.9	18	51.4	7	21.9	22	42.3
	Almost never a problem	9	8.6	2	5.7	4	12.5	3	5.8
	Sometimes a problem	30	28.6	8	22.9	13	40.6	19	36.5
	Often a problem	3	2.9	2	5.7	0	0.0	4	7.7
	Almost always a problem	19	18.1	5	14.3	7	21.9	4	7.7
Feel sad	Never a problem	43	41.0	14	40.0	9	28.1	25	48.1
	Almost never a problem	13	12.4	5	14.3	2	6.3	2	3.8
	Sometimes a problem	38	36.2	15	42.9	12	37.5	15	28.8
	Often a problem	2	1.9	0	0.0	5	15.6	5	9.6
	Almost always a problem	9	8.6	1	2.9	3	9.4	5	9.6
Feel angry	Never a problem	27	25.7	10	28.6	3	9.4	35	67.3
	Almost never a problem	11	10.5	3	8.6	4	12.5	1	1.9
	Sometimes a problem	36	34.3	15	42.9	12	37.5	10	19.2
	Often a problem	8	7.6	2	5.7	7	21.9	3	5.8
	Almost always a problem	23	21.9	5	14.3	5	15.6	3	5.8
Have trouble sleeping	Never a problem	42	40.0	21	60.0	12	37.5	30	57.7
	Almost never a problem	8	7.6	1	2.9	0	0.0	1	1.9
	Sometimes a problem	34	32.4	11	31.4	12	37.5	9	17.3

	Often a problem	3	2.9	1	2.9	2	6.3	8	15.4
	Almost always a problem	18	17.1	1	2.9	5	15.6	3	5.8
Worry what will happen to me	Never a problem	32	30.5	15	42.9	7	21.9	15	28.8
	Almost never a problem	8	7.6	1	2.9	2	6.3	2	3.8
	Sometimes a problem	25	23.8	9	25.7	13	40.6	19	36.5
	Often a problem	2	1.9	4	11.4	3	9.4	7	13.5
	Almost always a problem	38	36.2	6	17.1	6	18.8	9	17.3
SOCIALISING: Have trouble getting along with other kids	Never a problem	44	41.9	27	77.1	13	40.6	28	53.8
	Almost never a problem	10	9.5	1	2.9	2	6.3	0	0.0
	Sometimes a problem	33	31.4	6	17.1	7	21.9	21	40.4
	Often a problem	4	3.8	1	2.9	2	6.3	1	1.9
	Almost always a problem	14	13.3	0	0.0	7	21.9	2	3.8
Other kids do not want to be my friends	Never a problem	53	50.5	16	45.7	16	50.0	30	57.7
	Almost never a problem	14	13.3	2	5.7	3	9.4	1	1.9
	Sometimes a problem	23	21.9	13	37.1	10	31.3	19	36.5
	Often a problem	1	1.0	3	8.6	1	3.1	2	3.8
	Almost always a problem	14	13.3	1	2.9	1	3.1	0	0.0
Other kids tease me	Never a problem	39	37.1	10	28.6	8	25.0	26	50.0
	Almost never a problem	11	10.5	2	5.7	3	9.4	2	3.8
	Sometimes a problem	28	26.7	12	34.3	8	25.0	18	34.6
	Often a problem	5	4.8	0	0.0	5	15.6	4	7.7
	Almost always a problem	22	21.0	11	31.4	7	21.9	2	3.8
Cannot do things other kids my age can do	Never a problem	71	67.6	12	34.3	19	59.4	20	38.5
	Almost never a problem	10	9.5	0	0.0	0	0.0	2	3.8
	Sometimes a problem	9	8.6	7	20.0	5	15.6	17	32.7
	Often a problem	2	1.9	5	14.3	5	15.6	6	11.5
	Almost always a problem	13	12.4	11	31.4	2	6.3	7	13.5
Hard to keep up when I play with other kids	Never a problem	55	52.4	10	28.6	18	56.3	18	34.6
	Almost never a problem	14	13.3	3	8.6	1	3.1	1	1.9

	Sometimes a problem	26	24.8	8	22.9	8	25.0	19	36.5
	Often a problem	2	1.9	6	17.1	1	3.1	8	15.4
	Almost always a problem	8	7.6	8	22.9	3	9.4	6	11.5
SCHOOLING: Hard to pay attention in class	Never a problem	74	70.5	28	80.0	16	50.0	31	59.6
	Almost never a problem	7	6.7	2	5.7	3	9.4	1	1.9
	Sometimes a problem	14	13.3	5	14.3	7	21.9	17	32.7
	Often a problem	2	1.9	0	0.0	2	6.3	3	5.8
	Almost always a problem	8	7.6	0	0.0	3	9.4	0	0.0
I forget things	Never a problem	36	34.3	12	34.3	11	34.4	18	34.6
	Almost never a problem	14	13.3	3	8.6	5	15.6	3	5.8
	Sometimes a problem	35	33.3	17	48.6	9	28.1	28	53.8
	Often a problem	4	3.8	1	2.9	3	9.4	3	5.8
	Almost always a problem	16	15.2	2	5.7	3	9.4	0	0.0
Have trouble keeping up with schoolwork	Never a problem	71	67.6	23	65.7	15	46.9	24	46.2
	Almost never a problem	9	8.6	3	8.6	1	3.1	2	3.8
	Sometimes a problem	21	20.0	8	22.9	7	21.9	21	40.4
	Often a problem	1	1.0	1	2.9	5	15.6	4	7.7
	Almost always a problem	3	2.9	0	0.0	3	9.4	1	1.9
Miss school because not feeling well	Never a problem	52	49.5	10	28.6	8	25.0	4	7.7
	Almost never a problem	9	8.6	0	0.0	7	21.9	0	0.0
	Sometimes a problem	26	24.8	18	51.4	12	37.5	32	61.5
	Often a problem	5	4.8	3	8.6	3	9.4	12	23.1
	Almost always a problem	13	12.4	4	11.4	1	3.1	4	7.7
Miss school to go to Dr or hospital	Never a problem	46	43.8	11	31.4	4	12.5	0	0.0
	Almost never a problem	12	11.4	6	17.1	3	9.4	0	0.0
	Sometimes a problem	35	33.3	12	34.3	21	65.6	32	61.5
	Often a problem	3	2.9	3	8.6	3	9.4	16	30.8
	Almost always a problem	9	8.6	3	8.6	0	0.0	4	7.7

Appendix 22: Baseline WeeFIM results, complete (n=135)

		MS (n=16)		SS (n=35)		CI (n=32)		AI (n=52)	
		Count	%	Count	%	Count	%	Count	%
SELF-CARE: W1 eat	Total assistance	0	0.0	0	0.0	0	0.0	12	23.1
	Mod assistance	0	0.0	0	0.0	0	0.0	2	3.8
	Min assistance	0	0.0	0	0.0	1	3.1	5	9.6
	Modified assistance	0	0.0	1	2.9	0	0.0	0	0.0
	Complete independence	16	100.0	34	97.1	31	96.9	33	63.5
W1 groom	Total assistance	0	0.0	0	0.0	0	0.0	11	21.2
	Max assistance	0	0.0	2	5.7	0	0.0	1	1.9
	Mod assistance	0	0.0	1	2.9	1	3.1	2	3.8
	Min assistance	2	12.5	2	5.7	3	9.4	4	7.7
	Supervision	2	12.5	0	0.0	1	3.1	0	0.0
	Complete independence	12	75.0	30	85.7	27	84.4	34	65.4
W1 bath	Total assistance	0	0.0	0	0.0	0	0.0	1	1.9
	Max assistance	0	0.0	2	5.7	1	3.1	3	5.8
	Mod assistance	0	0.0	2	5.7	2	6.3	8	15.4
	Min assistance	1	6.3	3	8.6	1	3.1	23	44.2
	Supervision	0	0.0	0	0.0	0	0.0	1	1.9
	Complete independence	15	93.8	28	80.0	28	87.5	16	30.8
W1 dress up	Total assistance	0	0.0	0	0.0	0	0.0	4	7.7
	Max assistance	0	0.0	2	5.7	1	3.1	1	1.9
	Mod assistance	0	0.0	3	8.6	1	3.1	5	9.6
	Min assistance	0	0.0	2	5.7	2	6.3	16	30.8
	Supervision	0	0.0	0	0.0	0	0.0	2	3.8
	Complete independence	16	100.0	28	80.0	28	87.5	24	46.2
W1 dress low	Total assistance	0	0.0	0	0.0	0	0.0	19	36.5

	Max assistance	0	0.0	3	8.6	1	3.1	1	1.9
	Mod assistance	0	0.0	3	8.6	2	6.3	3	5.8
	Min assistance	0	0.0	7	20.0	2	6.3	12	23.1
	Modified assistance	0	0.0	1	2.9	0	0.0	0	0.0
	Complete independence	16	100.0	21	60.0	27	84.4	17	32.7
W1 Toileting	Total assistance	0	0.0	0	0.0	0	0.0	2	3.8
	Max assistance	0	0.0	3	8.6	1	3.1	2	3.8
	Mod assistance	0	0.0	0	0.0	0	0.0	4	7.7
	Min assistance	0	0.0	1	2.9	1	3.1	0	0.0
	Modified assistance	0	0.0	8	22.9	2	6.3	8	15.4
	Complete independence	16	100.0	23	65.7	28	87.5	36	69.2
W1 Bladder management	Total assistance	0	0.0	0	0.0	0	0.0	3	5.8
	Max assistance	0	0.0	1	2.9	0	0.0	1	1.9
	Mod assistance	0	0.0	1	2.9	0	0.0	2	3.8
	Min assistance	0	0.0	1	2.9	0	0.0	2	3.8
	Supervision	0	0.0	4	11.4	0	0.0	0	0.0
	Modified assistance	0	0.0	9	25.7	3	9.4	21	40.4
	Complete independence	16	100.0	19	54.3	29	90.6	23	44.2
W1 Bowel management	Total assistance	0	0.0	0	0.0	0	0.0	3	5.8
	Max assistance	0	0.0	1	2.9	0	0.0	3	5.8
	Mod assistance	0	0.0	6	17.1	1	3.1	3	5.8
	Min assistance	0	0.0	1	2.9	0	0.0	2	3.8
	Supervision	0	0.0	3	8.6	0	0.0	0	0.0
	Modified assistance	0	0.0	3	8.6	2	6.3	18	34.6
	Complete independence	16	100.0	21	60.0	29	90.6	23	44.2

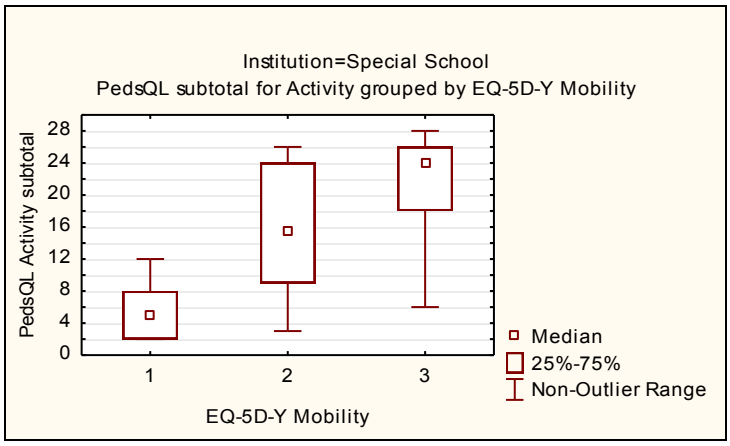
MOBILITY: W1 Transfer chair/wchair	Total assistance	0	0.0	0	0.0	0	0.0	18	34.6
	Max assistance	0	0.0	3	8.6	0	0.0	2	3.8
	Mod assistance	0	0.0	2	5.7	1	3.1	2	3.8
	Min assistance	0	0.0	4	11.4	0	0.0	5	9.6
	Supervision	0	0.0	0	0.0	0	0.0	2	3.8
	Modified assistance	0	0.0	14	40.0	5	15.6	0	0.0
	Complete independence	16	100.0	12	34.3	26	81.3	23	44.2
W1Transfers toilet	Total assistance	0	0.0	0	0.0	0	0.0	25	48.1
	Max assistance	0	0.0	4	11.4	0	0.0	1	1.9
	Mod assistance	0	0.0	5	14.3	3	9.4	0	0.0
	Min assistance	0	0.0	2	5.7	0	0.0	2	3.8
	Supervision	0	0.0	0	0.0	0	0.0	2	3.8
	Modified assistance	0	0.0	13	37.1	2	6.3	0	0.0
	Complete independence	16	100.0	11	31.4	27	84.4	22	42.3
W1 transfers tub	Total assistance	0	0.0	0	0.0	0	0.0	25	48.1
	Max assistance	0	0.0	4	11.4	0	0.0	1	1.9
	Mod assistance	0	0.0	7	20.0	3	9.4	1	1.9
	Min assistance	0	0.0	2	5.7	1	3.1	3	5.8
	Supervision	0	0.0	0	0.0	0	0.0	2	3.8
	Modified assistance	0	0.0	11	31.4	1	3.1	0	0.0
	Complete independence	16	100.0	11	31.4	27	84.4	20	38.5
Walk/Wheel chair	Total assistance	0	0.0	1	2.9	0	0.0	24	46.2
	Max assistance	0	0.0	0	0.0	0	0.0	1	1.9
	Mod assistance	0	0.0	0	0.0	2	6.3	0	0.0
	Min assistance	0	0.0	3	8.6	1	3.1	2	3.8

	Supervision	0	0.0	0	0.0	0	0.0	4	7.7
	Modified assistance	0	0.0	20	57.1	3	9.4	1	1.9
	Complete independence	16	100.0	11	31.4	26	81.3	20	38.5
Stairs	Total assistance	0	0.0	15	42.9	3	9.4	37	71.2
	Max assistance	0	0.0	0	0.0	1	3.1	0	0.0
	Mod assistance	0	0.0	1	2.9	0	0.0	1	1.9
	Min assistance	0	0.0	0	0.0	0	0.0	1	1.9
	Supervision	0	0.0	1	2.9	0	0.0	0	0.0
	Modified assistance	1	6.3	9	25.7	7	21.9	10	19.2
	Complete independence	15	93.8	9	25.7	21	65.6	3	5.8
COGNITION: W1 Comprehension	Mod assistance	0	0.0	4	11.4	0	0.0	0	0.0
	Min assistance	3	18.8	1	2.9	3	9.4	1	1.9
	Supervision	9	56.3	11	31.4	13	40.6	7	13.5
	Modified assistance	0	0.0	0	0.0	0	0.0	4	7.7
	Complete independence	4	25.0	19	54.3	16	50.0	40	76.9
W1 Expression	Mod assistance	0	0.0	2	5.7	0	0.0	0	0.0
	Min assistance	2	12.5	1	2.9	3	9.4	2	3.8
	Supervision	3	18.8	9	25.7	6	18.8	6	11.5
	Modified assistance	0	0.0	0	0.0	0	0.0	4	7.7
	Complete independence	11	68.8	23	65.7	23	71.9	40	76.9
W1 Social interaction	Mod assistance	0	0.0	1	2.9	0	0.0	0	0.0
	Min assistance	0	0.0	0	0.0	2	6.3	2	3.8
	Supervision	1	6.3	8	22.9	9	28.1	5	9.6
	Modified assistance	0	0.0	0	0.0	0	0.0	3	5.8
	Complete independe	15	93.8	26	74.3	21	65.6	42	80.8

	nce								
W1 Problem solving	Mod assistance	0	0.0	8	22.9	0	0.0	1	1.9
	Min assistance	2	12.5	2	5.7	3	9.4	1	1.9
	Supervision	3	18.8	9	25.7	11	34.4	6	11.5
	Modified assistance	1	6.3	1	2.9	0	0.0	1	1.9
	Complete independence	10	62.5	15	42.9	18	56.3	43	82.7
W1 Memory	Mod assistance	0	0.0	8	22.9	0	0.0	1	1.9
	Min assistance	1	6.3	2	5.7	3	9.4	1	1.9
	Supervision	7	43.8	9	25.7	12	37.5	6	11.5
	Modified assistance	2	12.5	3	8.6	2	6.3	1	1.9
	Complete independence	6	37.5	13	37.1	15	46.9	43	82.7

Box –Whisker graphs comparing EQ-5D-Y dimensions with similar dimensions on PedsQL and WeeFIM

Comparing EQ-5D-Y Mobility and PedsQL Activity dimensions at SS

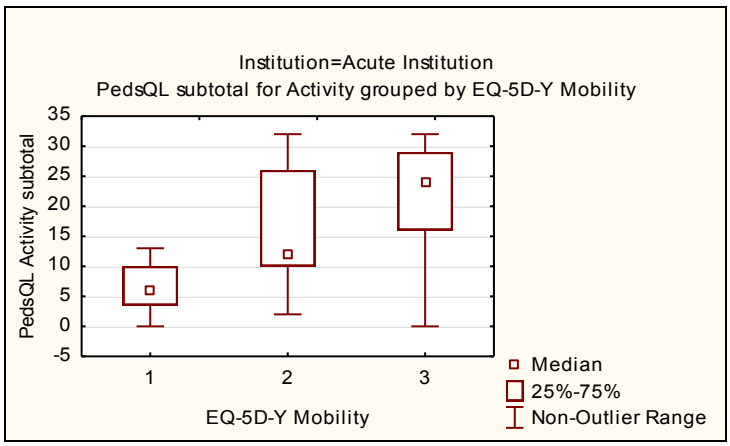


N=35

Appendix Figure 1: Comparing the EQ-5D-Y Mobility and PedsQL Activities dimensions, for SS

There was a significant difference in the ranking of PedsQL scores across the three levels of EQ-5D-Y Mobility, at the SS ($p<0.001$).

Comparing the EQ-5D-Y Mobility and of PedsQL Activities dimension, at AI

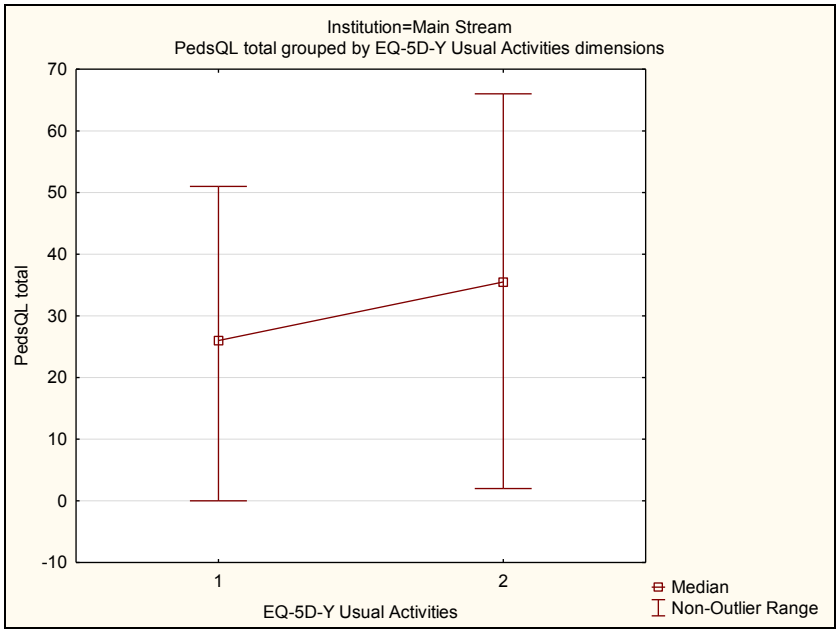


N=52

Appendix Figure 2: Comparing the Mobility EQ-5D-Y and Activities dimension of PedsQL, for AI

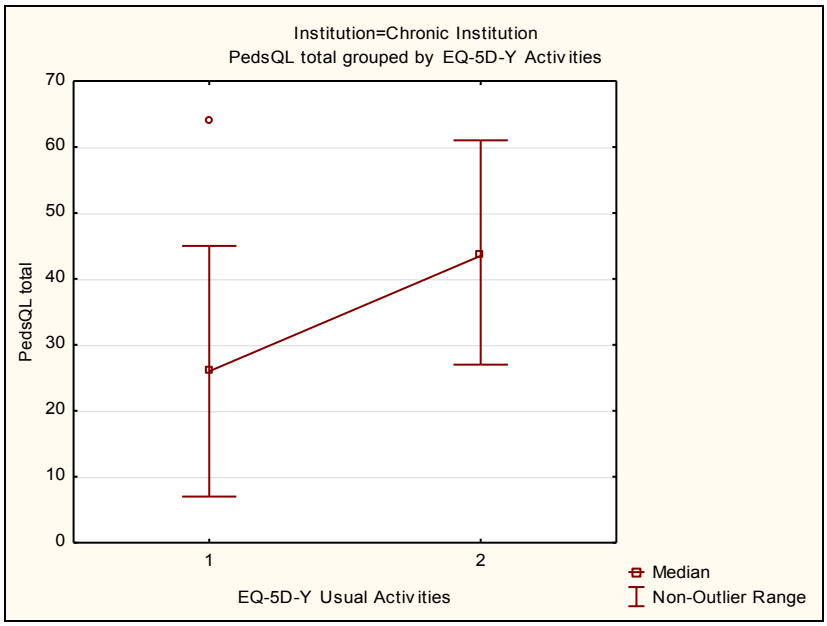
There was a significant difference in ranking of PedsQL Activity score across the three EQ-5D-Y levels for the AI ($p<0.001$)

Comparing EQ-5D-Y UA dimension with PedsQL total, for MS children



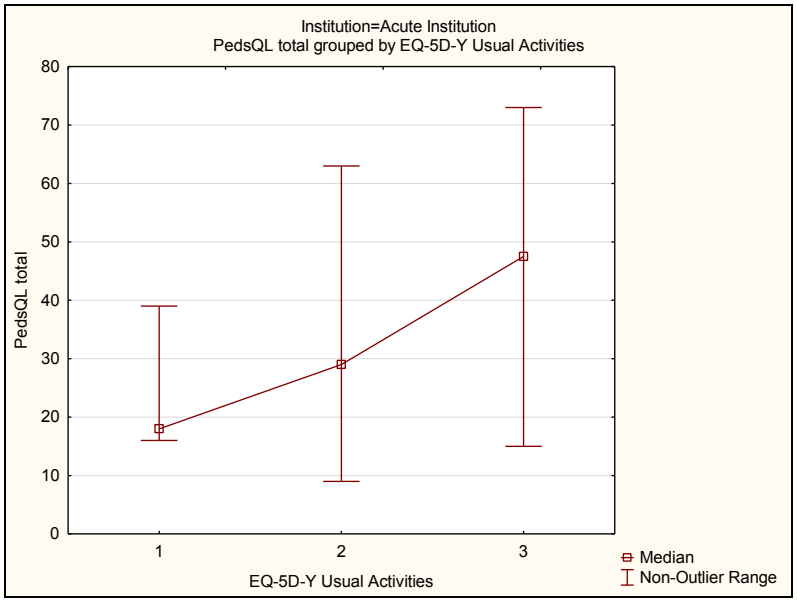
N=105

Appendix Figure 3: Comparing the UA EQ-5D-Y dimension and PedsQL total for MS school



Appendix Figure 4: Comparing the UA EQ-5D-Y dimension and PedsQL total for CI

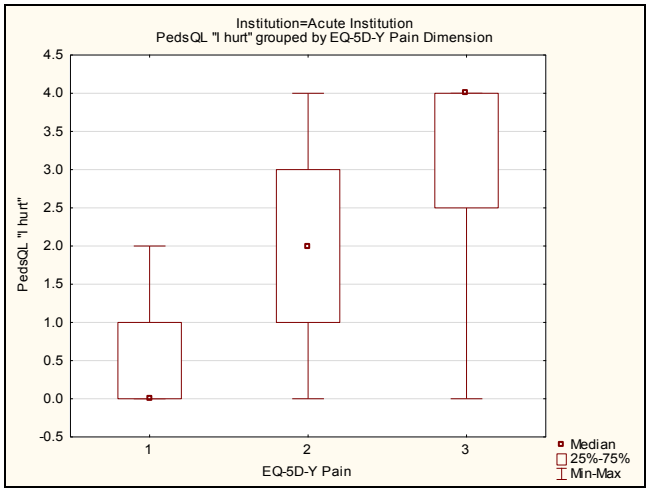
There was a significant difference in ranking of PedsQL total score across two levels of the EQ-5D-Y UA dimension, for MS children (p=0.007) and CI (p=0.002).



Appendix Figure 5: Comparing UA EQ-5D-Y dimension and PedsQL total for AI

There was a significant difference in ranking of PedsQL total score across three levels of the EQ-5D-Y UA dimension, for AI children (p=0.006).

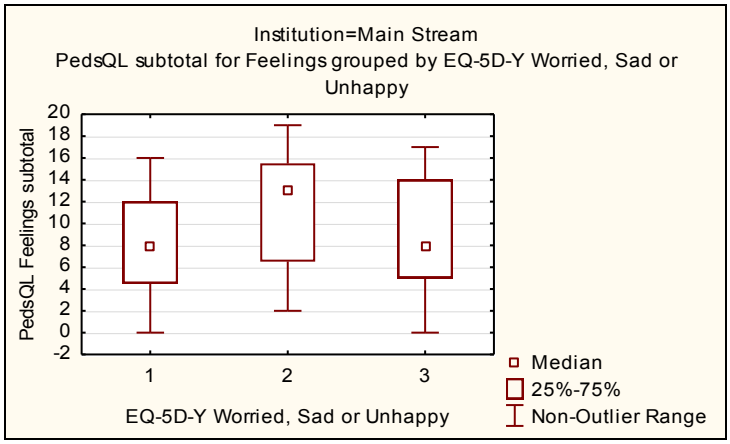
Comparing EQ-5D-Y P/D dimension with PedsQL "I hurt", for AI



N=52

Appendix Figure 6: Comparing EQ-5D-Y P/D dimension with PedsQL "I hurt", for AI

Comparing EQ-5D-Y WSU dimension with PedsQL Feelings dimension, for MS

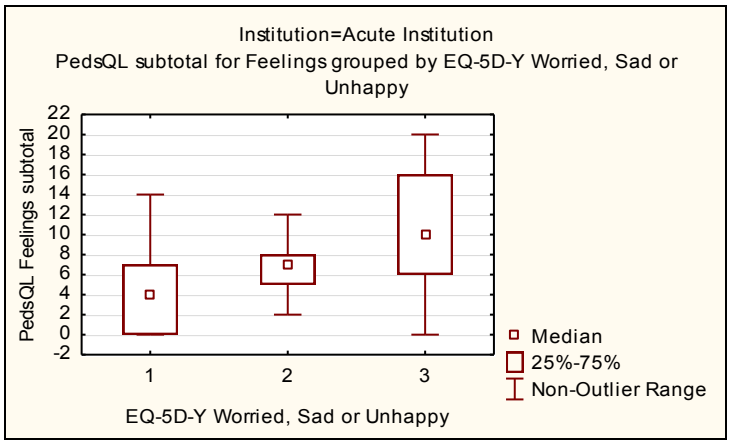


N=105

Appendix Figure 7: Comparing EQ-5D-Y WSU dimension and PedsQL Feelings dimension, for MS

There was a significant difference in ranking of PedsQL Feelings dimension score across two levels of the EQ-5D-Y WSU dimension for MS children (p=0.020).

Comparing EQ-5D-Y WSU dimension with PedsQL Feelings dimension, for AI

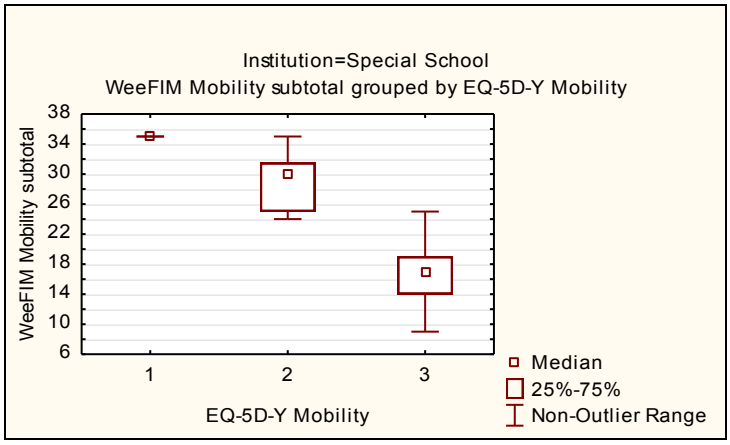


N=52

Appendix Figure 8: Comparing EQ-5D-Y WSU dimension with PedsQL Feelings dimension, for AI

There was a significant difference in ranking of PedsQL Feelings dimension score across the three levels of the EQ-5D-Y WSU dimension for acutely ill children (p=0.007)).

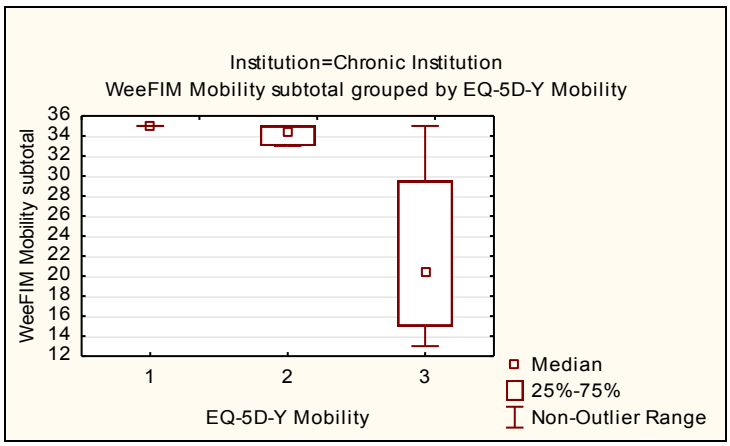
Comparing EQ-5D-Y Mobility and WeeFIM Mobility dimension for SS



N=35

Appendix Figure 9: Comparing the EQ-5D-Y Mobility and WeeFIM Mobility dimension for SS

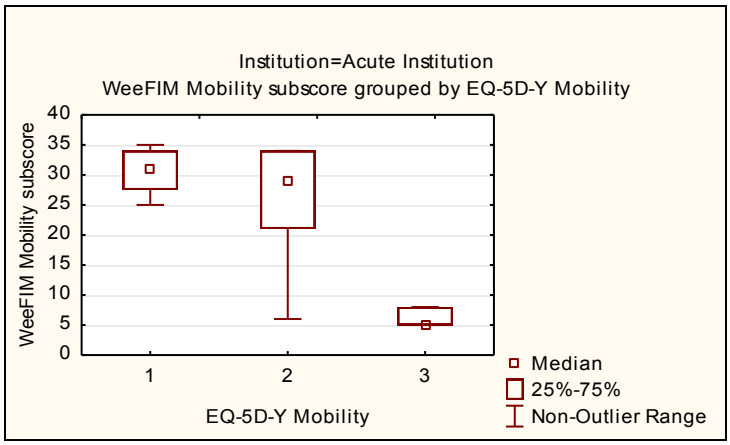
There was a significant difference in ranking of the WeeFIM mobility score across the three levels of the EQ-5D-Y mobility dimension, for the SS ($p<0.001$).



N=32

Appendix Figure 10: Comparing the EQ-5D-Y Mobility and WeeFIM Mobility dimension for CI

There was a significant difference in ranking of the WeeFIM mobility score across the three levels of the EQ-5D-Y mobility dimension, at the CI ($p=0.01$).

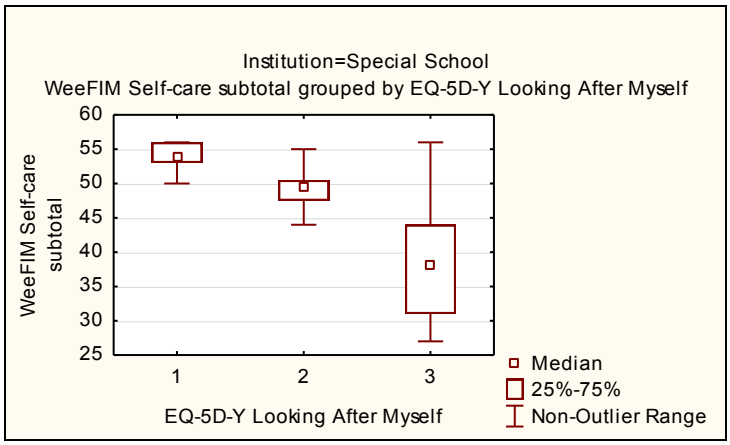


N=52

Appendix Figure 11: Comparing the EQ-5D-Y Mobility and WeeFIM Mobility dimension for AI

There was a significant difference in ranking of the WeeFIM mobility score across the three levels of the EQ-5D-Y mobility dimension, at the AI (p<0.001).

Comparing WeeFIM Self-care subtotal with EQ-5D-Y LAM dimension at SS

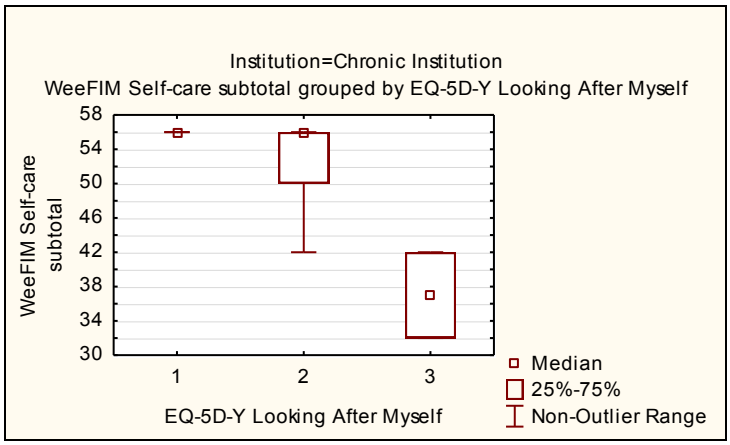


N=35

Appendix Figure 12: Comparing WeeFIM Self-care subtotal with EQ-5D-Y LAM dimension at SS

There was a significant difference in ranking of PedsQL Self-care subtotal across the different levels of the EQ-5D-Y LAM dimension, in the children at the SS (p<0.001).

Comparing WeeFIM Self-care subtotal with EQ-5D-Y LAM dimension at CI

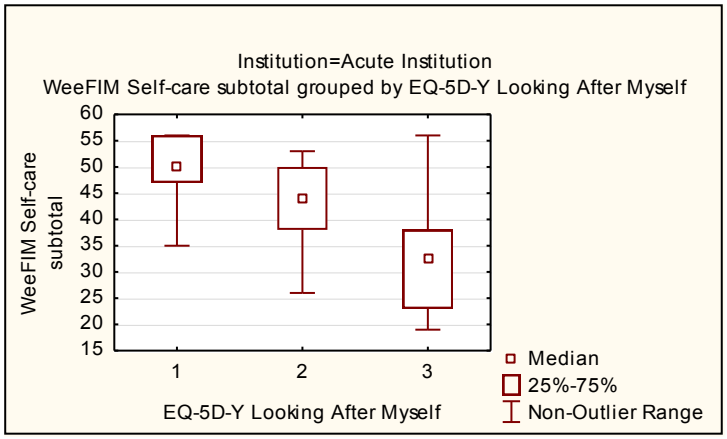


N=32

Appendix Figure 13: Comparing WeeFIM Self-care subtotal with EQ-5D-Y LAM dimension at CI

There was a significant difference in ranking of PedsQL Self-care subtotal across the different levels of the EQ-5D-Y LAM dimension, in the children at the CI (p<0.013).

Comparing WeeFIM Self-care subtotal with EQ-5D-Y LAM dimension at AI

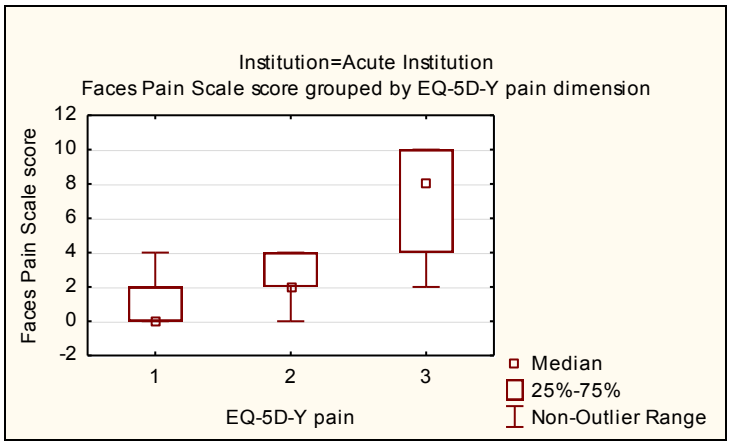


N=52

Appendix Figure 14: Comparing WeeFIM Self-care subtotal with EQ-5D-Y LAM dimension at AI

There was a significant difference in ranking of PedsQL Self-care subtotal across the different levels of the EQ-5D-Y LAM dimension, in the children at the AI ($p<0.001$).

Comparing EQ-5D-Y P/D dimension and Faces Pain Scale score at AI



N=52

Appendix Figure 15: Comparing EQ-5D-Y P/D dimension and Faces Pain Scale score at AI

There was a significant difference in ranking of FPS across the three levels of EQ-5D-Y P/D dimension at the AI ($p<0.001$).