health related quality of life in children

Nicolet C.M. Theunissen

Stellingen

behorende bij het proefschrift
"Health related Quality of Life in Children"
Nicolet C.M. Theunissen, 14 december 1999.

I.- Men verkrijgt onvoldoende inzicht in de Gezondheidsgerelateerde Kwaliteit van Leven van een kind, als men zich beperkt tot het turven van gezondheidsproblemen maar de emotionele waardering ervan buiten beschouwing laat (dit proefschrift).

II.- Het is mogelijk subjectieve opvattingen over de Gezondheidsgerelateerde Kwaliteit van Leven van kinderen objectief te meten (dit proefschrift).

III.- Artsen en ouders zijn slechts beperkt op de hoogte van hoe een kind diens Gezondheidsgerelateerde Kwaliteit van Leven ervaart (dit proefschrift). De verontwaardiging over de eigen stem die 12-jarigen in het recente euthanasie wetsvoorstel kregen is dus ook onterecht.

IV.- De World Health Organization bepleit al vanaf 1948 dat gezondheid fysieke, psychologische en sociale elementen heeft. Desondanks beperkt men zich, tot op heden, ten onrechte tot het verbeteren van de fysieke toestand als men de Gezondheidsgerelateerde Kwaliteit van Leven wil beïnvloeden (dit proefschrift).

V.- "We are still making false assumptions - such as thinking that what we can measure easily is more important than what we cannot" Kelnar, C.J.H. (1990) Pride and prejudice: Stature in perspective, Acta Paediatri.Scand., 370: 5-15.

VI.- "People may not be right about themselves, but their self-evaluations are the ones that most powerfully affect their future behavior." Byrnes, J.F (1984) The Psychology of Religion, New York: The Free Press.

VII.- Het zijn de Nederlanders die aan Engels jargon een speciale gevoelswaarde toewijzen waardoor vertaling onmogelijk wordt geacht. Wij verwarren verschil in klank met verschil in betekenis.

VIII.- Hoe beter het (statistisch) model de werkelijkheid beschrijft, hoe minder we het kunnen volgen.

IX.- De tijd besteed aan het overzichtelijk opbergen van documenten is omgekeerd evenredig aan de tijd nodig voor het terugvinden ervan.

X.- Je kunt iemand alleen veranderen als je die persoon volledig accepteert ook zonder veranderingen.

XI.- "Ofschoon er tienduizend onderwerpen bestaan voor een elegante conversatie, zijn er altijd mensen die geen kreupele kunnen ontmoeten zonder over voeten te praten. (China)" Vanden Berghe, G. (1969) Oosters citaten boek 2. Utrecht: Het spectrum N.V.

XII.- Men kan niet vaak genoeg de laatste hand leggen aan een manuscript.

HEALTH RELATED QUALITY OF LIFE IN CHILDREN

PROEFSCHRIFT

ter verkrijging van de graad van Doctor
aan de Universiteit Leiden
op gezag van de Rector Magnificus Dr. W.A. Wagenaar
hoogleraar in de faculteit der Sociale Wetenschappen,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 14 december 1999
te klokke 16:15 uur

door

Nicolette Catharina Maria Theunissen

geboren te Roosendaal en Nispen in 1964

Promotiecommissie

Promotoren:	Prof. Dr. J.M. Wit Prof. Dr. S.P. Verloove-Vanhorick
Copromotor:	Dr. H.M. Koopman
Referent:	Dr. H.M. Koot (Erasmus Universiteit Rotterdam)
Overige leden:	Prof. Dr. H.M. Berger Prof. Dr. C.M.J.G. Maes

Cover art-work:	Jeanine Theunissen
Cover design:	Linda Bruinenberg, LB Reclame & Presentatie

Financial support for the printing of this thesis has been kindly provided by the Department of Pediatrics of the Leiden University Medical Center, and by Nutricia Nederland B.V.

Printed by: Ridderprint offsetdrukkerij b.v., Ridderkerk, the Netherlands.
ISBN 90-9013294-5

Soms wordt het lichaam gezond door veel ziek te zijn. (Perzië)

Vanden Berghe, G. (1969) Oosters citaten boek 2.
Utrecht: Het spectrum N. V.

Voor mijn familie

Contents

N.C.M. Theunissen*, H.M. Koopman, R.P. Kamphuis, S. Veen and J.M. Wit are from the Department of Pediatrics, Leiden University Medical Center, The Netherlands.

G.H. Verrips, A.G.C. Vogels, M. Fekkes, E. Brugman, A.L. den-Ouden and S.P. Verloove-Vanhorick, are from TNO Prevention and Health, Leiden, The Netherlands.

A.H. Zwinderman is from the Department of Medical Statistics, Leiden University, The Netherlands.

A.A. Kaptein is from the Department of Psychiatry, Leiden University Medical Center, The Netherlands

J.J. Meulman is from the Department of Data Theory, Leiden University, The Netherlands.

*(now at the Department of Health Psychology, Utrecht University, The Netherlands)

1 Introduction

N.C.M. Theunissen

1.1 Background

Until recently, it was difficult to keep children with severe diseases alive. Therefore, mortality was most frequently used as an outcome of paediatric treatment.[1] As medical successes increased, morbidity, health status and the psychological and social consequences of a disease became more important.[2] Although the individual paediatrician already recognised that a certain disease or side effects of treatments can elicit quite different reactions in different children, a systematic outcome measure that could describe these individual differences was missing. This gap was filled by the introduction of the construct 'Health related Quality of Life (HRQoL)'.[3] This construct can be viewed as a method of translating a child's experience of illness into a quantifiable outcome.[4] Without this outcome measure, children's well-being is estimated by researchers and paediatricians who use their own personal reference points and their own experience with similar patients as guidelines.[5,6]

The definition of health provided in 1948 by The World Health Organisation[7] as 'a state of complete physical, mental, and social well being, and not merely the absence of disease or infirmity' has been highly influential in defining the HRQoL construct. By now, there is a growing consensus on four aspects of HRQoL: it is multi-factorial (physical, psychological and social well-being), it is patient self-administered, it is subjective, and its value is variable over time.[3,8] As a widespread definition or theoretical framework is missing, we propose a definition in which these four aspects are accounted for:

HRQoL is the individual's perception of problems in health status, combined with the affective responses to such problems.

In this thesis, the usefulness of this definition is explored in children up to 12 years of age. Furthermore, attention is given to the choice of a certain informant about HRQoL, measuring HRQoL in various age ranges and groups of children with disorders, and longitudinal changes in HRQoL.

1.2 Outline of this thesis

Chapters 2 to 6 concentrate on how to define and obtain HRQoL in children aged 6 to 11 and 1 to 5 years. To start with, *Chapter 2* deals with the construction of the TNO-AZL-Child Quality of Life (TACQOL) questionnaire. Psychometric performance of the TACQOL was studied in a large sample of children (age 6-12 years) from the Dutch open population. The chapter presents the definition of HRQoL that underlie this instrument as well as this thesis. The objective of *Chapter 3* is which informant to choose when studying the child's HRQoL: the child itself or his/her parent. The agreement between the TACQOL child form and parent form is tested in children between 8-11 years of age in the same sample as used in Chapter 2. In *Chapter 4*, agreement between children and parents is tested again, but this time in a sample of children with a chronic disorder. *Chapter 5* presents the psychometric performance of the TNO-AZL Pre-school Quality of Life (TAPQOL) questionnaire. The instrument was constructed according to the same principles as the TACQOL in order to assess HRQoL in children aged 1-5 year. The TAPQOL uses the parent as informant.
Chapter 6 discusses the HRQoL of pre-school children who were born prematurely. The relationship between the TAPQOL (as obtained from the parents), the parental feelings towards the child and the global HRQoL judgement of the neonatologist was studied.

Chapters 7 to 10 deal with HRQoL in a longitudinal perspective. *Chapter 7* consists of a systematic review on longitudinal studies in children between 0 to 12 years of age. The review explores the time variability of children's HRQoL in studies with at least two assessments of HRQoL. It focuses on underlying ideas about what exactly changes HRQoL. *Chapter 8* presents a strategy for dealing with the problem that measurement instruments used in a longitudinal study may differ at different time points. The strategy was applied to a set of longitudinal health status data, collected from preterm born children, but is thought to be very useful in HRQoL studies as well. In *Chapter 9,* the same sample as in Chapter 9 is presented, now with emphasis on the health status development between 5 and 10 years of age in the cohort of preterm children. *Chapter 10* describes a longitudinal study on the HRQoL, health status and self-perception of children with idiopathic short stature. These children participated in a randomised controlled study on the effect of growth hormone treatment.

In *Chapter 11*, the HRQoL results presented in this thesis are summarised and conclusions are drawn.

1.3 Reference List

1. Ebrahim S. Clinical and public health perspectives and applications of health-related quality of life measurement. *Soc Sci Med* 1995; **41**: 1383-1394.
2. Eiser C. Making sense of chronic disease. The eleventh Jack Tizard memorial lecture. *J Child Psychol Psychiat* 1994; **35**: 1373-1389.
3. Leventhal H, Colman S. Quality of life: A process view. *Psychology and Health* 1997; **12**: 753-767.
4. Levi R, Drotar D. Critical issues and needs in health-related quality of life assessment of children and adolescents with chronic health conditions. In: Drotar D, ed. *Measuring health-related quality of life in children and adolescents: implications for research and practice.* Mahwah, New Jersey: Lawrence Erlbaum Associates, 1998: 3-20.
5. Taylor KM, Macdonals KG, Bezjak A, DePetrillo PNG, DePetrillos AD. Physicians' perspective on quality of life: an exploratory study of oncologists. *Qual Life Res* 1996; **5**: 5-14.
6. Vogels T, Verrips GH, Verloove-Vanhorick SP, et al. Measuring health-related quality of life in children: the development of the TACQOL-Parent Form. *Qual Life Res* 1998; **7**: 457-465.
7. World Health Organization. *Constitution of the World Health Organization.* Geneva: World Health Organization, 1948.
8. Schipper H, Clinch JJ, Olweny CLM. Quality of life studies: Definitions and conceptual issues. In: Spilker B, ed. *Quality of life and pharmacoeconomics in clinical trials.* Philadelphia: Lippincott-Raven Publishers, 1996:11-23.

2 Measuring health-related quality of life in a child population

E.G.H. Verrips, A.G.C. Vogels, H.M. Koopman, N.C.M Theunissen,
R.P. Kamphuis, M. Fekkes, J.M. Wit, S.P. Verloove-Vanhorick

2.1 Abstract

The 56-item TNO-AZL-Child-Quality-Of-Life (TACQOL) questionnaire was developed to meet the need for a reliable and valid instrument for measuring Health Related Quality of Life (HRQoL) in children. HRQoL was defined as health status in seven domains plus emotional responses to problems in health status. The TACQOL explicitly offers respondents the possibility of differentiating between their functioning and the way they feel about it. The aims of the study were threefold: to evaluate psychometric performance of the TACQOL in the general population, to evaluate the relationship between Parent Forms and Child Forms and to obtain additional information about validity. A random sample of 1789 parents of 6-11 year olds completed the TACQOL (response rate 71%), as well as 1159 8-11 year olds themselves (response rate 69%). Multiple correspondence analyses showed that item response categories were ordinal and that the TACQOL scales may be regarded as metric. Cronbach's alpha ranged from .65 to .84. Only 57% of reported health status problems were associated with negative emotions. Intraclass correlation coefficients between Parent Forms and Child Forms ranged from .44 to .61. Pearson's correlation coefficients between TACQOL and KINDL ranged from .24 to .59. Univariate analyses of variance showed that children with chronic diseases and children receiving medical treatment had lower TACQOL scores than healthy children. The study showed that with the TACQOL children's HRQoL can be measured in a reliable and valid way.
Key words: children, health status, measurement, quality of life.

2.2 Introduction

For many decades outcome assessment in medicine has focused on mortality, morbidity and, more recently, on health status.[1-6] Although necessary and valuable, such outcome measures do not reflect patients' Health Related Quality of Life (HRQoL). In adults, measuring HRQoL has been the subject of over 500 articles published up to 1991. These were recently reviewed by Gill and Feinstein.[7] Of a sample of 75 of them,

only 17% included patients' personal views. It has been claimed that HRQoL assessments should provide information about capabilities and well-being, and their relevance to the individual concerned.[8-10] Even if health status is self-reported and, therefore, subjective in nature, the patients' own emotional evaluation of their health status often is not explicitly taken into account. In addition, in utility approaches, preferences for health states are elicited typically from the informed general public and not from the individual patient who has to live with the disease.

Only few attempts have been made to develop reliable generic HRQoL instruments for children, including the Health Utilities Index,[11] the Child Health Questionnaire,[12] the KINDL,[13] the 16D,[14] and the FS(II)R.[15] None of these instruments explicitly offers respondents the possibility to differentiate between their functioning and the way they feel about it. Related to this issue, HRQoL research in children presents two additional problems: age specificity and the proxy problem. Many existing questionnaires and scoring systems for adults are not applicable to children because they contain domains such as fertility, sexuality and economic independence. For children, even a domain such as independence in daily life (e.g. toilet use, dressing and tying one's own shoe laces) may be inappropriate. Furthermore, who is going to be the one to give an evaluation of the child's HRQoL: the child him/herself or some proxy such as a parent, a nurse, a doctor or a teacher?

The TNO AZL Child Quality Of Life (TACQOL) questionnaire[16-19] was developed to meet the need for a reliable and valid research tool for measuring HRQoL in children. In the literature, it is claimed that HRQoL should be regarded a multidimensional construct, including at least physical, daily living, social and psychological dimensions.[20-26] In accordance with these claims, the TACQOL covers seven domains: physical complaints and motor functioning (physical), autonomous functioning (daily living), social functioning (social), cognitive functioning, positive moods and negative moods (psychological functioning). Items and domains of the TACQOL were chosen on the grounds of results of focus groups with parents, paediatricians and developmental psychologists.[16,17] HRQoL was defined as health status plus emotional responses to problems in health status or, in other words, as a weighting of health status problems by the emotional impact of such problems. Thus, the TACQOL not only covers health status domains, but also the child's emotional evaluation of reported health status problems. Children's subjective emotional appraisal of their health is explicitly taken into account. The psychometric performance of the questionnaire in a sample of chronically ill children was satisfactory.[16-18]

The TACQOL Parent Form (TACQOL-PF) questionnaire is completed by parents of children aged 5-15. In the PF, parents are asked to answer the questions from the perspective of their child. For children aged 8-15, a parallel TACQOL Child Form (TACQOL-CF) is available.

The aims of the present study were threefold: to evaluate the psychometric performance of the TACQOL-PF and CF in the general population, to evaluate the relationship between the TACQOL-PF and CF and to obtain information about validity. The study was approved by the TNO Medical Ethics Committee.

2.3 Material and methods

2.3.1 Sample

Twelve GGDs (Municipal Health Services) spread over the Netherlands each selected a random sample of 210 children, stratified by gender and age (6-7, 8-9 and 10-11 years of age), resulting in a sample of 2520 children. Letters were sent to parents explaining the aims of the study and stressing the parents' right not to participate. If necessary, a reminder was sent after three weeks. The parent response rate was 71% (n=1789). The response rate from parents of children aged 8-11 was somewhat lower than from parents of children aged 6-7 (67% and 78% respectively). Fourteen % of parents had not been born in the Netherlands. A recent representative survey suggested that approximately 18% of parents of children in the relevant age groups are not of Dutch origin.[27] Ethnic minorities would therefore seem to be somewhat under-represented. Parents of children aged 8-11 (n=1680) were asked to have their children complete a questionnaire as well. The response rate from children was 69% (n=1159) and not related to gender or age.

2.3.2 Questionnaires

The parent questionnaire contained the TACQOL-PF as well as questions about demographic details, children's chronic illnesses, common illnesses and medical treatments. The child questionnaire contained the TACQOL-CF and the Dutch translation of the German KINDL. The KINDL is a generic, multidimensional instrument for children's HRQoL[13] measuring four dimensions: physical functioning

(Body), practical daily functioning (Daily), social functioning (Social) and psychological functioning (Psyche). The KINDL was translated into Dutch in accordance with internationally agreed guidelines for the translation of HRQoL questionnaires.[28]

Both the TACQOL-PF and the TACQOL-CF contain 56 items, covering seven eight-item domains: physical complaints (BODY), motor functioning (MOTOR), autonomy (AUTO), cognitive functioning (COGNIT), social functioning (SOCIAL), positive emotions (EMOPOS) and negative emotions (EMONEG). The items of the TACQOL-PF are presented in Table 1.

Table 1. Items of the TNO AZL Child Quality Of Life Parent Form (TACQOL-PF)

Physical complaints	Cognitive functioning	
(Has your child had/did your child have...)	(Did your child have difficulty with...)	
ear aches or sore throats	paying attention, concentrating	
stomach-aches or abdominal pain	understanding schoolwork	
headaches	understanding what others said	
dizzy	arithmetic	
felt sick/ nauseous	reading	
tired	writing	
sleepy	learning	
dozy/lethargic	saying what he/she meant	
Motor functioning	**Social functioning**	
(Did your child have difficulty with...)	(My child was...)	
running	able to play or talk happily with other children	
walking	able to stand up for his/herself with other children	
standing	other children asked my child to play with them	
walking downstairs	at ease with other children	
playing	able to play or talk happily with us, parents	
running or walking long distances	incommunicative or quiet with us, parents	
balance	restless or impatient with us, parents	
doing things handily or quickly	deviant with us, parents	
Autonomy	**Positive emotions**	**Negative emotions**
(Did your child have difficulty with...)	(My child felt...)	(My child felt...)
going to school on his/her own	joyful	sad
washing his/herself	relaxed	aggressive
getting dressed on her own	in good spirits	angry
going to the lavatory on his/her own	happy	short tempered
eating or drinking on his/her own	content	worried
sports or going out to play on his/her own	cheerful	jealous
doing hobbies on his/her own	enthusiastic	gloomy
riding a bicycle	confident	anxious

In each item, the frequency of occurrence of health status problems is assessed.

If such a problem is reported, the emotional reaction of the child to this problem is determined. The reference period is formulated as 'the last few weeks'. For example, the question 'did your child have difficulty concentrating in school', with response categories 'never', 'occasionally' or 'often', is followed (only when the response is 'occasionally' or 'often') by the statement 'at that time, my child felt..', with response categories 'fine', 'not so good', 'quite bad' or 'bad'. Items are scored by assigning a value of 4 to the 'never' response, a value of 3 to a 'feeling fine' response, a value of 2 to a 'feeling not so good' response, a value of 1 to a 'feeling quite bad' response and a value of 0 to a 'feeling bad' response. Item scores within a scale are added to a scale score, with a range of 0 to 32. If more than two items are missing in a scale, the scale score is declared missing. In case of one or two missing items, the mean of the non-missing items is imputed. Higher scale scores indicate better HRQoL. No emotional responses to items in the positive or negative emotions scales (EMOPOS and EMONEG) are assessed since this would lead to nonsensical items. A typical emotions item is: 'was your child happy', with response categories 'never', 'occasionally' or 'often' . Emotions item categories are assigned a value of 0 to 'never', a value of 1 to 'occasionally' and a value of 2 to 'often'. Consequently, EMOPOS and EMONEG range from 0 to16.

2.3.3 Statistical analyses

In coding pairs of questions into single scores, ordinality of response categories was assumed. This assumption was evaluated by calculating category quantifications in Multiple Correspondence Analyses (HOMALS) and, subsequently, counting the number of such quantifications in violation of ordinality. Furthermore, HOMALS provides object quantifications (scale scores) that may be considered metric variables. Pearson correlation coefficients (PCCs) were calculated between scales obtained using such object quantifications, on the one hand and the TACQOL scales obtained by the scoring system presented (e.g. unweighted summation of the item-pairs per scale), on the other. High PCCs would indicate that TACQOL scales may be considered as metric too. The advantage of the TACQOL scoring system above a scoring system based on the HOMALS object quantifications is that the former is independent of sample characteristics. Mutual independence of the TACQOL scales was evaluated by calculating item-rest and item-scale PCCs, as well as the TACQOL scales PCCs. Cronbach's alpha was calculated for each scale.

The validity of the distinction between health status problems as such and the

emotional impact of such problems was evaluated by calculating the total number of problems and the percentage of such problems associated with negative emotional reactions. The relationship between PF and CF was evaluated by calculating Intra Class Correlation Coefficients (ICC). Concurrent validity was evaluated by calculating PCCs between the TACQOL scales and the KINDL scales. Criterion validity was evaluated by relating the TACQOL scores to three dichotomous criteria: any common, non-chronic illness during a four week period prior to testing, any medical treatment in the six months prior to testing and any chronic illness. This was done by one-way analyses of variance.

2.4 Results

The percentual distributions of the TACQOL scales are presented in Table 2, along with the percentage of missing items per scale. Scale scores were categorised for ease of interpretation. Table 2 shows that the percentage of missings was low.

Table 2. Percentage of missing answers, Cronbach's alpha's and categorised percentual distributions of TACQOL-PF (n=1777) and CF scales (n=1116)

		BODY	MOTOR	AUTO	COGNIT	SOCIAL	EMOPOS	EMONEG
Parents								
Percentage of items with missing answers		0.6	0.4	0.4	1.5	0.5	1.4	1.3
Cronbach's alpha		.70	.79	.69	.84	.67	.84	.71
Categorised percentual distribution (%)								
0-15 BODY to SOCIAL	0-7 EMOPOS, EMONEG	1	0	0	0	0	1	5
16-19 BODY to SOCIAL	8-9 EMOPOS, EMONEG	3	1	0	3	0	4	15
20-23 BODY to SOCIAL	10-11 EMOPOS, EMONEG	12	2	1	6	2	3	28
24-27 BODY to SOCIAL	12-13 EMOPOS, EMONEG	28	5	3	18	9	8	31
28-31 BODY to SOCIAL	14-15 EMOPOS, EMONEG	40	29	23	36	57	29	18
32 BODY to SOCIAL	16 EMOPOS, EMONEG	15	63	73	38	32	55	3
Children								
Percentage of items with missing answers		0.5	0.5	0.5	0.9	0.8	1.1	1.1
Cronbach's alpha		.76	.74	.66	.79	.65	.78	.76
Categorised percentual distribution (%)								
0-15 BODY to SOCIAL	0-7 EMOPOS, EMONEG	6	1	0	1	0	3	7
16-19 BODY to SOCIAL	8-9 EMOPOS, EMONEG	10	1	0	3	1	6	15
20-23 BODY to SOCIAL	10-11 EMOPOS, EMONEG	20	4	1	8	2	9	25
24-27 BODY to SOCIAL	12-13 EMOPOS, EMONEG	27	10	3	18	12	22	26
28-31 BODY to SOCIAL	14-15 EMOPOS, EMONEG	28	41	22	43	51	32	22
32 BODY to SOCIAL	16 EMOPOS, EMONEG	10	44	74	27	34	29	6

TACQOL scales scores were skewed. Cronbach's alpha ranged from .65-.84. The percentage of ordinality violations was 6% in the TACQOL-PF and 7% in the TACQOL-CF. Violations occurred mainly in response categories with a prevalence of less than 1%. PCC between HOMALS object quantifications scales and TACQOL scales ranged from .89 to .95; therefore, TACQOL scales may be considered metric variables. Over 98% of the item-rest PCCs were higher than PCCs of items with other scales. The PCCs between TACQOL-PF scales ranged from .21 to .48, with the exception of the PCC between the AUTO and MOTOR scale (.61). The PCCs between TACQOL-CF scales ranged from .22 to .48.

Table 3 presents the number of problems perceived and the percentage of such problems associated with negative emotions. Parents reported a total of 8144 problems, only 57% of which were associated with negative emotions in their child. Children reported 9411 problems, 63% of which were associated with negative emotions. Children reported significantly more problems on the BODY and MOTOR scales than did their parents . Moreover, on all scales but the SOCIAL scale, children reported more problems associated with negative emotions than did their parents.

Table 3. Number of problems perceived (NP) and number (NE) and percentage (%NE) of such perceived problems associated with negative emotional reactions; Complete PF/CF pairs only (8-11-year-olds; n=1054)

	Parent Form			Child Form		
Scale	NP	NE	% NE	NP	NE	% NE
BODY	2886	2191	76	3721	2960	80
MOTOR	875	495	57	1313	791	60
AUTO	455	208	46	481	279	58
COGNIT	2372	968	41	2416	1116	46
SOCIAL	1556	775	50	1480	796	54
Total	8144	4637	57	9411	5942	63

Table 4 (see next page) shows means and 95% confidence intervals for the scale scores, together with ICCs between the TACQOL-PF and TACQOL-CF scale scores. Table 4 shows that ICCs between PF and CF scales ranged from .44-.61. The magnitude of PF means did not differ very much from CF means. PCCs between TACQOL and KINDL scales ranged from .24 to .59. PCCs of the four KINDL scales among each other ranged between .54 and .74.

Table 4. Mean TACQOL-PF and CF scores, 95% confidence interval of mean (95%CI), and ICC by TACQOL scale; complete PF/CF pairs only (8-11-year-olds; n=1054)

	Parent Form		Child Form		
Scale	Mean	95%CI	Mean	95%CI	ICC
BODY	26.9	26.7 - 27.2	24.9	24.6 - 25.2	.54
MOTOR	30.6	30.5 - 30.8	29.8	29.6 - 30.0	.48
COGNIT	28.7	28.5 - 29.0	28.5	28.2 - 28.7	.61
AUTO	31.3	30.8 - 31.4	31.2	31.1 - 31.3	.45
SOCIAL	29.7	29.6 - 29.9	29.7	29.5 - 29.9	.51
EMOPOS	14.7	14.6 - 14.8	13.6	13.4 - 13.7	.39
EMONEG	11.5	11.4 - 11.7	11.6	11.4 - 11.8	.55

One-way analyses of variance showed that children with a chronic illness and children who had undergone medical treatment had lower scores on all TACQOL-PF and CF scales ($p<.001$), except for the COGNIT scale ($p<.01$). Effects sizes ranged from 2% to 7% of the maximum scale range. Differences on the CF scale were in the expected direction, but were only significant ($p<.05$) for BODY, MOTOR, AUTO and EMONEG. Having had a common illness showed small though statistically significant effects on the BODY, MOTOR, AUTO and SOCIAL scales ($p<.05$).

2.5 Discussion

In this study, HRQoL was defined as health status plus the emotional responses to problems in health status, the latter being the way children react to their health status problems emotionally. Children's HRQoL was measured with the TACQOL, a recently developed generic instrument. The aims of the study were to evaluate the psychometric performance of the TACQOL-PF and TACQOL-CF in the general population, to evaluate the relationship between TACQOL-PF and CF and to obtain additional information on validity.

The psychometric performance of both the TACQOL-PF and the TACQOL-CF was satisfactory. The TACQOL scales were skewed, which is to be expected in a general population. However, the parametric techniques that were applied are quite robust against skewness, and have been demonstrated to be adequate in analysing skewed data if sample size is large enough.[29] Cronbach's alpha ranged from .65 to .84,

which is regarded as satisfactory for use of the TACQOL in comparing groups.[30-32] However, when the value of a scale for an individual is of interest, the TACQOL cannot be used safely; for use in the clinical situation much higher levels of Cronbach's alpha are mandatory. In addition, the stability of the TACQOL, as well as sensitivity to change, still need to be ascertained. The validity of the scale structure, i.e. the domains that were distinguished, is supported by the finding that item-rest PCCs were almost always higher than PCCs for items from other scales. Furthermore, PCCs between TACQOL scales were low to moderate. The construct validity of the TACQOL is therefore good.

PF scales were moderately correlated to CF scales. This implies that, at an individual level, a parent differed considerably from her child in their judgement about the child's HRQoL. This is a common finding that has been described extensively in the literature on proxy ratings.[33,34,35] In commenting on this phenomenon, Saigal et al.[36] state that " ..the viewpoint of those who have to live with a certain health state is what matters". As no golden standard exists, and both parents' and children's opinions may be valuable in evaluating treatment effects, it seems best to obtain both parents' and children's evaluations whenever possible. As PF and CF scale means did not differ greatly, at a group level the TACQOL-PF may be regarded as a satisfactory proxy for the TACQOL-CF. The relationship between TACQOL-PF and TACQOL-CF has been studied in detail elsewhere.[19]

Concurrent validity was evaluated by relating TACQOL-CF scales to KINDL scales. PCCs were low. This held true even for scales intended to measure comparable concepts. The lack of relations between the TACQOL and the KINDL may partly be caused by a different time frame: the last weeks for the TACQOL, and the last week for the KINDL. In addition, since the PCCs between the KINDL scales were high, the Dutch KINDL scales may predominantly reflect a single quality of life dimension. In contrast, the TACQOL-CF scales were only moderately interrelated, indicating high domain specificity, with each domain only moderately related to a common, single quality of life factor. If these findings are replicated in future research on concurrent validity of the TACQOL-CF and the Dutch KINDL, the TACQOL-CF may be more consistent with a multi-dimensional definition of HRQoL than the Dutch KINDL.

To evaluate criterion validity, the TACQOL scales were related to three criteria: common illnesses, medical treatment and chronic illnesses. As expected, these criteria had negative effects on the TACQOL-PF and CF scores, although effect sizes were not very large. As has been reported in the literature, children's HRQoL may be influenced by other factors than their health status. For instance, Saigal et al.[37] found that

even severely handicapped children rated their health status as highly as healthy controls did. At the same time, parents may underestimate both the prevalence and emotional impact of health status problems.[19]

The validity of the distinction between health status and HRQoL was supported by the finding that only approximately half of the health status problems reported were associated with negative emotional reactions in the children. The TACQOL explicitly offers respondents the possibility of differentiating between their functioning and the way they feel about their functioning. The possibility that patients have a health problem, but do not feel bad about it may bias patients' self-reporting in typical health status questionnaires. Patients may wish to incorporate the fact that they do not feel bad about a certain health status problem by rating their health status problem as less severe than a proxy rater such as a doctor, a parent or a spouse would. This mechanism may explain the often reported lack of difference between patients and healthy controls in self-reported health status.[37,38] If it matters how children feel about their functioning rather than how they are functioning,[7] measuring health status alone does not provide all relevant information. Clearly, the present study has shown that 'subjectivity can be made scientific'.[8] With the TACQOL, HRQoL in children can be measured in a reliable and valid way. The instrument may therefore be suitable for evaluating the effects of medical intervention in chronically ill children. Curing health problems is often impossible in chronic diseases such as diabetes mellitus or congenital heart disease, but negative emotional feelings resulting from these diseases may be prevented or reduced.

2.6 Reference List

1. Saigal S, Rosenbaum P, Stoskopf B, et al. Comprehensive assessment of the health status of extremely low birth weight children at eight years of age: Comparison with a reference group. *J Pediatr* 1994; **125:** 411-417.
2. Saigal S, Feeny D, Furlong W, Rosenbaum P, Burrows E, Torrance G. Comparison of the health-related quality of life of extremely low birth weight children and a reference group of children at age eight years. *J Pediatr* 1994; **125:** 418-425.
3. Gemke RJ, Bonsel GJ, van-Vught AJ. Long-term survival and state of health after paediatric intensive care. *Arch Dis Child* 1995; **73:** 196-201.
4. Veen S, Ens-Dokkum MH, Schreuder AM, Verloove-Vanhorick SP, Brand R, Ruys JH. Impairments, disabilities, and handicaps of very preterm and very-low-birthweight infants at five years of age. *Lancet* 1991; **338:** 33-36.
5. Feeny D, Furlong W, Barr RD, Torrance GW, Rosenbaum P, Weitzman S. A comprehensive multiattribute system for classifying the health status of survivors of childhood cancer. *J Clin Oncol* 1992; **10:** 923-928.
6. World Health Organization. *International classification of impairments, disabilities, and handicaps.* Geneva: World Health Organization, 1980: 207.

7. Gill TM, Feinstein AR. A critical appraisal of the quality of quality-of-life measurements. *JAMA* 1994; **272:** 619-626.
8. O'Boyle C. Making subjectivity scientific. *Lancet* 1995; **345:** 602-602.
9. Collings JA. Epilepsy and well-being. *Soc Sci Med* 1990; **31:** 165-170.
10. Collings JA. Psychosocial well-being and epilepsy: an empirical study. *Epilepsia* 1990; **31:** 418-426.
11. Feeny DH, Torrance GW, Furlong WJ. Health utilities index. In: Spilker B, ed. *Quality of life and pharmacoeconomics in clinical trials.* Philadelphia: Lippincott-Raven Publishers, 1996: 239-252.
12. Landgraf JM, Abetz L, Ware JE. *The CHQ user's manual.* (1st edition edition) Boston, MA: The Health Institute, New England Medical Centre, 1996.
13. Bullinger M, von-Mackensen S, Kirchberger I. Ein Fragebogen zur Erfassung der Gesundheitsbezogene Lebensqualitaet von Kindern [German: a questionnaire for assessing the health related quality of life of children]. *Z Gesundheitspsychol* 1994; **2:** 64-77.
14. Apajasalo M, Sintonen H, Holmberg C, et al. Quality of life in early adolescence: a sixteen-dimensional health-related measure (16D). *Qual Life Res* 1996; **5:** 205-211.
15. Stein RE, Jessop DJ. Functional status II(R). A measure of child health status. *Med Care* 1990; **28:** 1041-1055.
16. Vogels T, Theunissen NCM, Verrips GH, Koopman HM, Verloove-Vanhorick SP, Kamphuis RP. Het meten van kwaliteit van leven bij kinderen met chronische aandoeningen [Dutch: Assessing health related quality of life in chronically ill children]. *TIAZ* 1996; **3:** 104-111.
17. Verrips GH, Vogels AGC, Verloove-Vanhorick SP, et al. Health-related quality of life measure for children -- the TACQOL. *J Appl Therapeutics* 1997; **1:** 357-360.
18. Vogels T, Verrips GH, Verloove-Vanhorick SP, et al. Measuring health-related quality of life in children: the development of the TACQOL-Parent Form. *Qual Life Res* 1998; **7:** 457-465.
19. Theunissen NCM, Vogels AGC, Koopman HM, et al. The proxy problem: child report versus parent report in health-related quality of life research. *Qual Life Res* 1998; **7:** 387-397.
20. Aaronson NK. Quality of life: what is it? How should it be measured? *Oncology Huntingt* 1988; **2:** 69-76,64.
21. Eisen M, Ware Jr. JE, Donald CA, Brook RH. Measuring components of children's health status. *Med Care* 1979; **17:** 902-921.
22. Fitzpatrick R, Fletcher A, Gore S, Jones D, Spiegelhalter D, Cox D. Quality of life measures in health care. I: Applications and issues in assessment. *BMJ* 1992; **305:** 1074-1077.
23. Bradburn NM. *The structure of psychological well-being.* Chicago: Aldine Publishing Company, 1969.
24. Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: Development and final revision of a health status measure. *Med Care* 1981; **19:** 787-805.
25. Bullinger M, Hasford J. Evaluating quality-of-life measures for clinical trials in Germany. *Control Clin Trials* 1991; **12:** 91S-105S.
26. Christie MJ, French D, Weatherstone L, West A. The patients' perception of chronic disease and its management: psychosomatics, holism and quality of life in contemporary management of childhood asthma. *Psychother Psychosom* 1991; **56:** 197-203.
27. Brugman E, Goedhart H, Vogels T, van-Zessen G. *Jeugd en seks [Dutch: Youth and sex].* Utrecht: Uitgeverij SWP, 1995.
28. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol* 1993; **46:** 1417-1432.
29. Worthington HV. The suitability of the statistical techniques currently used to describe and analyse cross-sectional caries data. *Community Dent Health* 1984; **1:** 125-130.
30. Nunnally JC. *Psychometric theory.* New York: McGraw-Hill, 1967.
31. Bland JM, Altman DG. Cronbach's alpha. *BMJ* 1997; **314:** 572.
32. McDonald RP. The dimensionality of tests and items. *Britisch Journal of Mathematical and Statistical Psychology* 1981; **34:** 100-117.
33. Sprangers MAG, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: A review. *J Clin Epidemiol* 1992; **45:** 743-760.

34. Verhulst FC, Prince J, Vervuurt PC, de-Jong J. Mental health in Dutch adolescents: self-reported competencies and problems for ages 11-18. *Acta Psychiatr Scand Suppl* 1989; **356:** 1-48.
35. Verhulst FC, Akkerhuis GW. Agreement between parents' and teachers' ratings of behavioral/emotional problems of children aged 4-12. *J Child Psychol Psychiat* 1989; **30:** 123-136.
36. Saigal S, Feeny D, Rosenbaum P, Furlong W, Stoskopf B, Hoult L. Extremely low-birth-weigth infants at adolescence: Health status and quality of life: reply to a letter to the editor. *JAMA* 1996; **276:** 1722-1723.
37. Saigal S, Feeny D, Rosenbaum P, Furlong W, Burrows E, Stoskopf B. Self-perceived health status and health-related quality of life of extremely low-birth-weight infants at adolescence. *JAMA* 1996; **276:** 453-459.
38. Post M. *Living with spinal cord injury.* Utrecht: Elinkwijk, 1997.

3 The Proxy Problem: Child Report versus Parent Report in Health Related Quality of Life Research

N.C.M. Theunissen, A.G.C. Vogels, H.M. Koopman, E.G.H. Verrips,
K.A.H. Zwinderman, S.P. Verloove-Vanhorick, J.M. Wit

3.1 Abstract

This study evaluates the agreement between child and parent reports on children's Health related Quality of Life (HRQoL) in a representative sample of 1105 Dutch children (age 8-11 years old). Both children and their parents completed a 56 item questionnaire (TACQOL). The questionnaire contains seven eight-item scales: physical complaints, motor functioning, autonomy, cognitive functioning, social functioning, positive emotions and negative emotions. The Pearson correlations between child and parent reports were between .44 to .61 ($p<.001$). The intraclass correlations were between .39 to .62. On average the children reported a significantly lower HRQoL than parents on the physical complaints, motor functioning, autonomy, cognitive functioning and positive emotions scales (paired t-test: $p\leq.05$). Agreement on all of the scales was related to the magnitude of the HRQoL scores and to some background variables (gender, age, temporary illness and visiting a physician). According to Multitrait-Multimethod analyses, both the child and parent reports proved to be valid. *Key words: health related quality of life, health status, proxy, child report, parent report.*

3.2 Introduction

Until recently, it was very difficult to keep children with severe diseases alive. Consequently, mortality was most frequently used as an outcome of paediatric treatment.[1] As medical treatments improved, outcome measures such as morbidity, health status, and the psychological and social consequences of medical treatments were increasingly used to evaluate paediatric treatment.[2-9] A systematic outcome measurement that combines perceptions of physical, psychological and social functioning is 'Quality of Life'.[1,3,10,11-21] This article focuses on Health related Quality of Life (HRQoL). The 'health' component refers to Quality of Life as a result of a certain health. According to the definition of the World Health Organization (1948), health

involves physical and psychological as well as social functioning. The assessment by a person of his or her own health functioning is called Health Status (HS). To asses the HRQoL of a person not only is HS relevant but also the emotional evaluation of this HS by the person him- or herself. Therefore, HRQoL is defined as a combination of health status and the affective response to problems in health status.[1,12,22,23] The definition incorporates individual and culturally determined differences in coping with HS problems and reflects internal standards about HRQoL, factors which are emphasised by several authors.[1,3,12,23-25]

Self-report questionnaires are regarded as the primary method of assessing HRQoL.[21] However, their use with children is problematic because children may lack the necessary language skills, the cognitive abilities to interpret the questions, and a long-term view of events.[26] Therefore, proxy-by-parent may be a useful alternative. This explorative study evaluates to what extent parents and children agree on the child's HRQoL and at what age child reports can be reliably and validly obtained. The research population consisted of Dutch children aged 8 to 11 years and their parents. Since agreement between child and parent reports has been investigated only rarely,[27] we are obliged to compare our results with studies about proxy on adults. However, these comparisons will be of limited value, because children do not have the same internalised standard as adults for judging their current level of HRQoL.

Special attention was given to assessing agreement on HS, since most previous research on agreement was done on the HS component of HRQoL and thus, without affective evaluation by the patient.[2,4,13,21,28] Furthermore, the degree of agreement between child and parent for several background variables was assessed.

3.3 Method

3.3.1 Sample

Data were collected from a Dutch sample of 1122 children between 8 to 11 years old and 1127 parents. The analytical sample consisted of 1105 child-parent pairs (17 children and 22 parents were excluded in order to obtain complete child-parent pairs). The children's sample consisted of 523 children of 8 to 9 years old and 582 children of 10 to 11 years old (with equal numbers of boys and girls). Most children attended regular school (97%), and a minority (3%) attended a special school for children with learning disabilities. Fifty-six (5.1%) children were an only-child, 378

(34.5%) were the firstborn, 242 (22.1%) were middle child, and 421 (38.4) were youngest of the children in family. The parent forms were completed either by a female caretaker, mainly mothers (84%), a male caretaker, mainly fathers (11%) or by mother and father together (5%). The parents' ages ranged from 21 to 55 years (Mean = 38.5; SD = 4.4). Their levels of education were 29% low, 43% moderate and 28% high according to the Dutch school system.[29] Most of the families were native Dutch speakers (96%) and most of the parents were born in The Netherlands (90%). Eighty percent of the parents were divorced or single. The sample was selected by 12 representative regional community health services in The Netherlands. They each drew a sample of 140 children, stratified according to gender and age (8-9 and 10-11 years old). Both the children and their parents received a questionnaire by mail. The criteria about informed consent and anonymity were met according to the TNO Medical Ethics Committee. A reminder was sent if questionnaires were not returned within three weeks. The result was a response rate of 67%. Three hundred (27%) children were completely healthy; 603 (54%) children had temporary illnesses such as a cold or influenza due to a winter epidemic; 40 (4%) children had at least one chronic illness; and 162 (15%) children had both a chronic and temporary illness at the time of data collection.

3.3.2 Measures

The data on HRQoL were collected using the TNO-AZL Children's Quality of Life (TACQOL) questionnaire. Two parallel questionnaires for children's HRQoL were available with identical items: a child form (CF) and a parent form (PF). The items were adjusted to the type of informant: "Has your child had..." in the PF and "Have you had..." in the CF.[30] The instrument was developed as a generic instrument intended for HRQoL assessment in medical research and clinical trials.[26] The TACQOL contains seven scales of eight items each: physical complaints, motor functioning, autonomy, cognitive functioning, social functioning, positive emotions and negative emotions. The instrument was constructed to measure HRQoL, defined as the combination of HS and the affective evaluation of problems in HS. A concretely and specifically formulated HS problem, if reported, leads to a question about the emotional response. Figure 1 shows an example of such a question. The items were scored on a 0-4 scale (added into Figure 1 in italics, between brackets). The eight item scales were scored on a 0-32 scale, with higher scores representing better HRQoL.

Figure 1. An example of a TACQOL-PF* question translated from the Dutch original

Has your child had difficulty running?	❑ never *[4]*	❑ sometimes	❑ often		
		During this my child felt:			
		❑(very)good *[3]*	❑ not so well *[2]*	❑ rather bad *[1]*	❑ bad *[0]*

* The TACQOL-CF equivalent of the question is 'Have you had difficulty running?' and 'During this I felt: '.

No affective responses about (positive and negative) emotional functioning were assessed, since this would have led to illogical items. The item scores on emotional problems were on a 0-2 scale (0 = often, 1 = sometimes, 2 = never) and the scale scores ranged from 0-16. The TACQOL-CF and TACQOL-PF were supplemented with questions assessing -- among others -- gender and age, chronic illnesses and temporary illnesses. The item scores were also encoded excluding the affective evaluation to obtain HS scores (0 = often, 1 = sometimes, 2 = never). The HS scales were scored on a 0-16 scale, with higher scores representing better HS.

3.3.3 Analytic Strategy

All of the following computations were done for both HRQoL and HS. Overall, a p-value of .05 or less was assumed to represent a significant result.

The results of the children's and parents' reports were summarised with mean scores and standard deviations. The mean, expressed as a percentage of the maximum score, was added to enable a comparison between the HRQoL and HS group means. The differences in the means between the child and parent reports were tested with paired Student's t-tests. Agreement between the children and parents was further quantified using the means of the absolute differences, the means of the differences (mean bias) and the associated standard deviations,[31] Pearson correlation coefficients, and the intraclass correlations (ICC).[32,33] Differences between HRQoL and HS agreement correlations were tested using computations of Fisher's Z-scores.[34]

The reliability of the HS and HRQoL scales was tested using Cronbach's alpha coefficient. The convergent and discriminant validity was evaluated using Multitrait-Multimethod analyses (MTMM).[35,36] A model was fitted onto the MTMM data using the computer program EQS (BMDP Statistical Software, Los Angeles).[33]

Goodness-of-fit indices (NFI: Bentler-Bonnett Normed fit Index, NNFI: Bentler-Bonnett NonNormed fit Index, CFI: Comparative Fit Index) above 0.90 were assumed to represent a good fit. Seven latent variables for the seven scales were constructed, with no restrictions on their correlations and two additional sources of variations: child versus parent reports and measurement error.[37,38]

Possible relationships between the magnitude of the scores and the amount of proxy agreement could give scatter bias or random fluctuation. The possible occurrence of this phenomena was estimated by regressing the child report against the parent report.[21,33] Forward, stepwise, linear regression analyses were computed for each scale, to reveal possible relationships between absolute proxy agreement and the background variables of gender, age, health of the child (chronically ill or temporary ill), visits to a physician, life events, place in children's row in family, age of the parents and education level of the parents. The child score was considered as a dependent variable, whereas parent score and background variables were considered as independent variables.[21]

Table 1. Central tendency of child and parent reports of HRQoL and HS

Scale		M child (%)[b]	SD child	M parent (%)	SD parent
Physical complaints	HRQoL	24.95 (78)	5.12	26.97 (84)	4.01 *
	HS	12.09 (76)	2.62	13.09 (82)	2.17 *
Motor functioning	HRQoL	29.82 (93)	3.22	30.66 (96)	2.71 *
	HS	14.61 (91)	1.87	15.08 (94)	1.75 *
Autonomy	HRQoL	31.21 (98)	1.94	31.35 (98)	1.61 *
	HS	15.48 (97)	1.21	15.52 (97)	1.17
Cognitive functioning	HRQoL	28.45 (89)	3.90	28.71 (90)	3.89 *
	HS	13.49 (84)	2.47	13.48 (84)	2.74
Social functioning	HRQoL	29.70 (93)	2.78	29.71 (93)	2.63
	HS	14.42 (90)	1.88	14.44 (90)	1.59
Positive emotions	HRQoL/HS[a]	13.56 (85)	2.55	14.69 (92)	2.13 *
Negative emotions	HRQoL/HS[a]	11.57 (72)	2.72	11.52 (72)	2.48

Range HRQoL: 0-32; range HS 0-16; high score = high HRQoL or HS ; [a]on these scales HS= HRQoL, [b]percentage of maximum score, *paired student's t-test: $p \leq .05$

3.4 Results

The TACQOL questionnaires proved to be reliable. The Cronbach's alpha coefficient range of HRQoL scales was .65 to .84 the alpha range of HS scales was .64 to .84. Table 1 (see previous page) presents the means of the HRQoL and HS scales. The children reported lower HRQoL than parents for the scales relating to physical complaints, motor functioning, autonomy, cognitive functioning, and positive emotions (paired t-tests: $p \leq .05$). The situation for HS was slightly different: the children reported a significantly lower HS than their parents for the physical complaints, motor functioning, and positive emotions scales (paired t-tests: $p \leq .05$). Without affective evaluation (=HS), the group means of the autonomy and cognitive functioning scales were alike, whereas the group means differed when affective evaluation was included (=HRQoL).

Table 2. Agreement between child reports and parent reports of HRQoL and HS

Scale		Mean bias[b]	(%)[c]	Mean abs. diff.[d]	(%)	SD diff.[e]	r **	Z-scores	ICC
Physical complaints	HRQoL	-2.01*	(6)	3.10	(10)	4.15	.61		.54
	HS	-1.00*	(6)	4.89	(31)	2.16	.61	-0.07	.56
Motor functioning	HRQoL	-0.84*	(3)	1.67	(5)	3.00	.50		.48
	HS	-0.47*	(3)	1.02	(6)	1.77	.54	-2.82*	.52
Autonomy	HRQoL	-0.14*	(0)	0.75	(2)	1.84	.48		.47
	HS	-0.04	(0)	0.50	(3)	1.17	.53	-4.30*	.53
Cognitive functioning	HRQoL	-0.26*	(1)	2.21	(7)	3.45	.61		.61
	HS	0.01	(0)	1.49	(9)	2.30	.62	-0.85	.62
Social functioning	HRQoL	-0.01	(0)	1.68	(5)	2.69	.51		.51
	HS	-0.02	(0)	1.04	(7)	1.78	.48	1.61	.48
Positive emotions	HRQoL/ HS[a]	-1.13*	(7)	1.81	(11)	2.49	.44		.39
Negative emotions	HRQoL/ HS[a]	0.05	(0)	1.78	(11)	2.47	.55		.55

Range HRQoL: 0-32; range HS 0-16; high score = high HRQoL or HS ; [a] on these scales HS = HRQoL; [b] Child group mean - parent group mean; [c] percentage of maximum score; [d] Mean of absolute difference between parent-child pairs; [e] SD of difference between parent-child pairs; *$p \leq .05$; **Pearson correlations all $p < .001$

Several indices quantifying the agreement between child and parent reports are shown in Table 2. Mean absolute difference and mean bias are largest for physical complaints, which indicates less agreement on these scales than on the other scales. This holds for both the HS and HRQoL scales. The Pearson correlations on HRQoL ranged from .44 to .61 ($p<.001$). The absolute agreement mimicked the relative agreement, ICC and Pearson correlations were alike. Although the differences between the correlations were small, the HRQoL correlations were significantly lower than the HS correlations on motor functioning and autonomy. Physical complaints, social and cognitive functioning correlations were not significantly different.

Figure 3 at the next page shows the MTMM Pearson correlation matrix of HRQoL together with the four validity tests. Convergent validity was completely confirmed and discriminant validity for the most part. Discriminant test number 4 (see the legend of Figure 3) was met except for the traits Autonomy (AUTO) and trait Motor functioning (MOTOR). These traits correlated more with each other within both methods (CAUTO-CMOTOR: .61 and PAUTO-PMOTOR: .55) than with itself between methods (CAUTO-PAUTO: .48 and CMOTOR-PMOTOR: .50). The MTMM Pearson correlation matrix of HS is not shown in this article. The convergent validity for HS was confirmed completely and discriminant validity for the most part. Discriminant test number 4 was met with two exceptions: First, autonomy and motor functioning correlated more with each other within both methods than with themselves between methods. Second, positive emotions and social functioning correlated more with each other within method 'parent report' than with themselves between methods. The model to be fitted onto the MTMM data in HRQoL by the EQS program is represented in Figure 2.

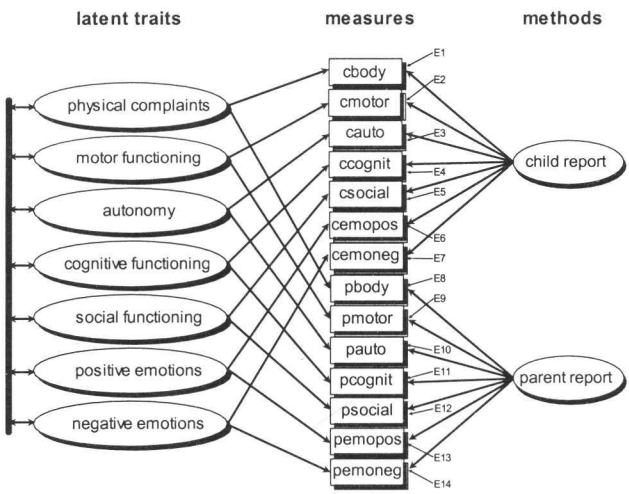


Figure 2. The MTMM model on HRQoL

Figure 3. Multitrait-Multimethod Pearson correlations between child report and parent report (methods) for 7 HRQoL scales (traits)

(methods)		Child report							Parent report						
	(traits)	CBODY	CMOTOR	CAUTO	CCOGNIT	CSOCIAL	CEMOPOS	CEMONEG	PBODY	PMOTOR	PAUTO	PCOGNIT	PSOCIAL	PEMOPOS	PEMONEG
Child report	CBODY	*1.00*													
	CMOTOR	0.47	*1.00*												
	CAUTO	0.32	0.61	*1.00*											
	CCOGNIT	0.39	0.46	0.38	*1.00*										
	CSOCIAL	0.30	0.36	0.33	0.45	*1.00*									
	CEMOPOS	0.23	0.31	0.29	0.28	0.37	*1.00*								
	CEMONEG	0.38	0.34	0.26	0.43	0.48	0.29	*1.00*							
Parent report	PBODY	***0.61***	0.29	0.23	0.25	0.21	0.16	0.27	*1.00*						
	PMOTOR	0.24	***0.50***	0.38	0.24	0.19	0.22	0.19	0.35	*1.00*					
	PAUTO	0.18	0.39	***0.48***	0.20	0.20	0.16	0.14	0.26	0.55	*1.00*				
	PCOGNIT	0.23	0.24	0.14	***0.61***	0.21	0.17	0.24	0.27	0.31	0.28	*1.00*			
	PSOCIAL	0.21	0.24	0.26	0.28	***0.51***	0.29	0.33	0.23	0.30	0.31	0.33	*1.00*		
	PEMOPOS	0.17	0.22	0.19	0.23	0.31	***0.44***	0.27	0.22	0.33	0.25	0.26	0.40	*1.00*	
	PEMONEG	0.29	0.23	0.19	0.27	0.36	0.31	***0.55***	0.32	0.24	0.20	0.27	0.49	0.38	*1.00*

LEGEND

bold+italics = monotrait-heteromethod (validity diagonal)

italics = monotrait-monomethod (testretest reliability diagonal)

= heterotrait-heteromethod triangles

= heterotrait-monomethod triangles

VALIDITY TESTS

1. convergent validity = monotrait-heteromethod > 0 (sign.)
2. discriminant validity= monotrait-heteromethod > heterotrait-heteromethod columns + rows (example: trait BODY)
3. discriminant validity= same paterns for all heterotrait triangles
4. discriminant validity= monotrait-heteromethod > heterotrait-monomethod (example underlined:trait AUTO)

Table 3. Percentage explained variance in a Multitrait-Multimethod model of HRQoL

Measure	Method	Latent trait	Method	Error
Physical complaints	child	68	8	24
	parent	65	14	21
Motor functioning	child	59	10	30
	parent	67	24	9
Autonomy	child	41	22	37
	parent	38	5	57
Cognitive functioning	child	42	17	40
	parent	54	6	40
Social functioning	child	38	30	32
	parent	39	21	40
Positive emotions	child	55	6	39
	parent	65	2	32
Negative emotions	child	50	5	45
	parent	73	0	26
Total (M%)	child	51	14	35
	parent	57	10	32

Goodness of fit (a good fit: >0.9): NFI=0.984, NNFI=0.981, CFI=0.991 (NFI: Bentler-Bonnett Normed fit Index, NNFI: Bentler-Bonnett NonNormed fit Index, CFI: Comparative Fit Index)

As can be seen in Table 3, the model did have a good fit. Children and parent scores were determined primarily by the latent traits (51 and 57% respectively) and much less by the method (14 and 10% respectively) and error (35 and 32% respectively). Therefore, the validity of both methods seemed to be high. The children's scores are explained well by the latent traits physical complaints, motor functioning, positive emotions and negative emotions. The parent's scores are explained well by the latent traits physical complaints, motor functioning, cognitive functioning, positive emotions and negative emotions. The performance was limited for children and parent scores on the autonomy and social functioning scales.

The MTMM model on HS (see Table 4 next page) followed a pattern similar to the pattern of the HRQoL model: good fit, scores determined primarily by the latent trait, and a high overall validity. The children's scores are explained well by the latent trait physical complaints, motor functioning, and positive emotions. The parent's scores are explained well by the latent traits physical complaints, motor functioning, cognitive functioning, positive emotions, and negative emotions. Again, performance was limited for children and parent scores on the autonomy and social functioning scales. HRQoL physical complaints seemed to be explained best by the children's report, HS physical complaints by the parent's report. To summarise, results of MTMM modelling revealed that the TACQOL was valid as a HRQoL questionnaire as well as a HS questionnaire.

Table 4. Percentage explained variance in a Multitrait-Multimethod model of HS

Measure	Method	Latent trait	Method	Error
Physical complaints	child	67	11	22
	parent	70	13	17
Motor functioning	child	70	9	22
	parent	76	16	8
Autonomy	child	36	13	50
	parent	40	6	54
Cognitive functioning	child	43	18	39
	parent	55	6	39
Social functioning	child	41	25	34
	parent	41	16	44
Positive emotions	child	50	5	45
	parent	63	3	34
Negative emotions	child	49	3	48
	parent	73	0	26
Total (M%)	child	51	12	37
	parent	60	9	32

Goodness of fit (a good fit: >0.9): NFI=0.977, NNFI=0.965, CFI=0.984 (NFI: Bentler-Bonnett Normed fit Index, NNFI: Bentler-Bonnett NonNormed fit Index, CFI: Comparative Fit Index)

The extent of agreement was related to the level of HRQoL or HS. The regression on positive emotions is given as an example in Figure 4a . If a child and parent both reported low HRQoL, the child reported relatively higher HRQoL than their parent. If a child and parent both reported a high HRQoL, then the child scores were relatively lower than the parent scores. All HRQoL and HS scales followed the same pattern. Overall, child scores were less extreme than the parent scores.

The child's age was related to agreement on autonomy and positive emotions scales (HRQoL and HS: $p \leq .05$). With low autonomy or positive emotions scores children aged 10 to 11 years were less in agreement with their parents than children aged from 8 to 9 years. Conversely, with high autonomy or positive emotions scores, children aged 10 to 11 years were more in agreement with their parents than children aged 8 to 9 years. In Figure 4b, this age difference with positive emotions is shown as an example. The greater the distance is between the regression line and the Y=X line, the lower the agreement between parents and children.

The child's gender was related to agreement on autonomy (HRQoL and HS: $p \leq .05$). If autonomy scores were low, boys were less in agreement with their parents than girls. Conversely, if autonomy scores were high, boys were more in agreement with their parents than girls.

Figure 4a. Regression of the child's positive emotions on the parent's positive emotions (A higher score represents a better HRQoL)

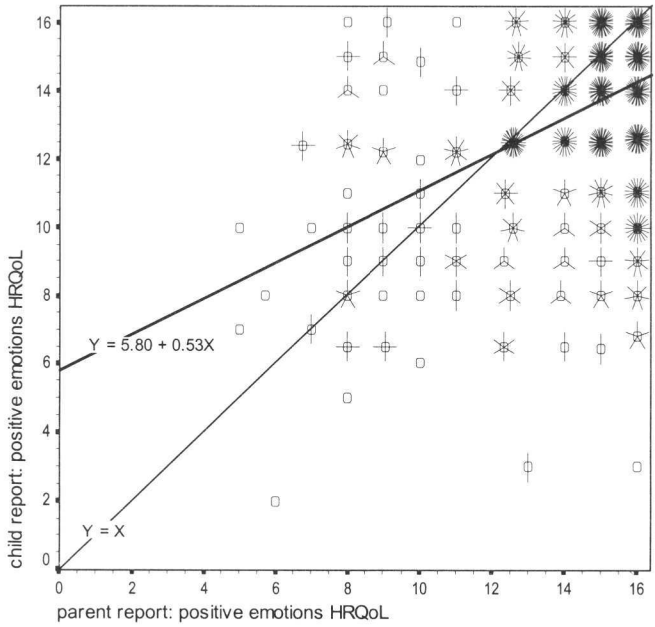
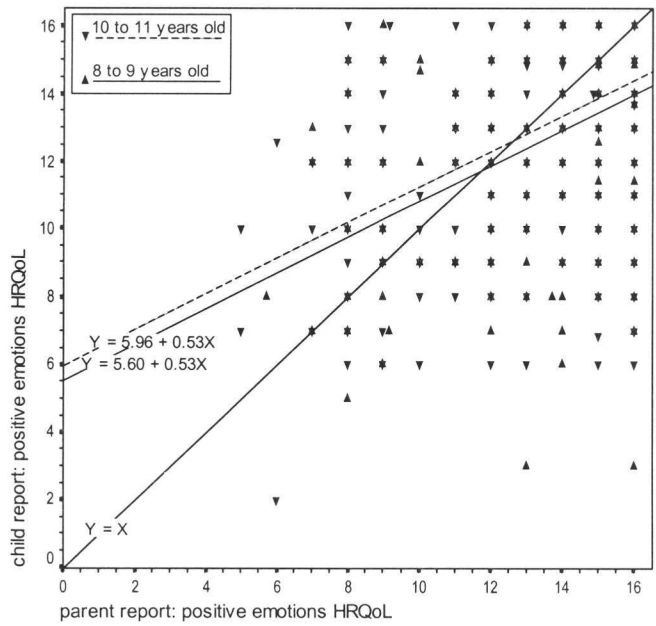


Figure 4b. Regression of the child's positive emotions on the parent's positive emotions by age (A higher score represents a better HRQoL)



The presence of a temporary illness was related to agreement on the physical complaints and social functioning scales (HS: $p \leq .05$). With low HS physical complaints or social functioning scores, children with a temporary illness had more agreement with their parents than children without a temporary illness. Conversely, with high HS physical complaints or social functioning scores, children with a temporary illness had less agreement with their parents than children without a temporary illness.

The occurrence of visits to a physician was related to agreement on the autonomy and motor functioning scales (HRQoL and HS: $p \leq .05$). With low autonomy or motor functioning scores children who visited a physician had more agreement with their parents than children who did not. Conversely, with high autonomy or motor functioning scores, children who visited a physician had less agreement with their parents than children who did not.

The age of the parents was related to agreement on social functioning (HRQoL and HS: $p \leq .05$). With low social functioning scores younger parents had more agreement with their children than older parents. Conversely, with high social functioning scores, younger parents had less agreement with their children than older parents. Chronic illness, life events, place in children's row in family and the education level of the parents were not related to agreement between child and parent reports.

3.5 Discussion

Our results showed that measuring HS as well as HRQoL (combining HS with affective evaluation) is possible. The TACQOL-PF and CF had satisfactory internal consistency with regard to HRQoL as well as HS. The convergent validity between the child and parent reports was acknowledged completely and the discriminant validity for the most part. The construct validity of the HRQoL and HS questionnaires were supported by MTMM models. Both the parent and the child report are valid, but the parent report seemed to perform best on most scales (HRQoL and HS). The child report performed best on HRQoL physical complaints. The performance of both child and parent reports was limited on the autonomy and social functioning scales. Consequently, the questionnaires do not fully meet the criteria on these scales. It is unknown however, if these scales have worse quality then other instruments because MTMM modelling performance of other HRQoL or HS instruments has not yet been studied, and our instrument had good reliability, convergent and discriminant validity and overall construct validity.

As far as generalisation is concerned, some reservations have to be made. Many children were suffering from minor and temporary illnesses such as a cold or influenza. This appeared to have a relation with agreement on HS physical complaints and social functioning scales. Furthermore, the parent's level of education in our study was a little higher than in the general Dutch population. This was not related to agreement between children and parents. Moreover, the sample did not include children with mental disorders. Apart from these reservations, the study's sample represented ordinary Dutch families. However, being a sample drawn from the normal population, any generalisation to clinical populations may be limited. There is reason to believe that this is not a problem. Firstly, the study contained 202 (19%) children with at least one chronic illness. The effect of chronic illness on proxy agreements was therefore investigated in this study and appeared to be absent. Secondly, the TACQOL questionnaires were developed and studied by means of a clinical sample. The questions included turned out to be relevant to a clinical population.[26]

Conceptualisation of HRQoL as a combination of HS and affective evaluation has consequences for agreement. According to previous studies, agreement seemed to be relatively good for observable measures.[2,4,21] Affective evaluations are probably less observable for parents than health status. Therefore, it could be expected that agreement between parent and child on HRQoL were lower than agreement on HS. Indeed, HRQoL agreement on motor functioning and autonomy was significantly lower than HS agreement. However, HRQoL agreement matched HS agreement on the other scales, which indicates that adding affective evaluation does not influence observability. According to other studies, social and psychological HRQoL seem to be less observable than physical HRQoL.[13,21] Perhaps this kind of HS scales already have a strong subjective component. This could explain the lack of difference in agreement between HRQoL and HS on cognitive and social functioning, positive and negative emotions, but not on physical complaints. The physical complaints scale included questions concerning pain. These are probably less observable than the more visible motor problems. Therefore, adding affective evaluation may not increase the subjectiveness of the before mentioned scales. By adding affective evaluation to HRQoL, agreement on motor functioning and autonomy became poor, but the level of subjectiveness was more evenly distributed among the HRQoL scales. It would be interesting to see the effect of adding affective evaluation to adult HRQoL research.

In the following paragraph the discussion of the results about proxy agreement is grouped around the main domains physical, psychological, and social functioning. Our results will be compared with studies about proxy on adults (HRQoL as well as HS) and with studies assessing agreement between children and their parents.

The TACQOL scales, physical complaints, motor functioning and autonomy cover physical HRQoL. The HRQoL and HS results revealed that on average children had more pessimistic ideas about their own physical functioning than their parents did. A similar result was found in a study assessing pain amongst hospitalised children.[28] Conversely, in several studies children, adults or elderly people on average agreed with their proxies or were more optimistic about their physical functioning.[2,4,13,21,39] The Pearson correlations between self- and proxy reports varied from low to high in the studies previously mentioned amongst children, adults and elderly people (range: .18 to .75). The wide variation in correlations may be due to observability of the items assessed, next to age, sample size or differences in health.

The TACQOL scales, cognitive functioning, positive emotions and negative emotions covered psychological HRQoL. The results indicate that children on average agree with their parents or are more pessimistic. Similar results were also found in other studies, along with contrasting results which indicated children as being more optimistic than their proxies about psychological functioning.[13,21,39,40] In several studies including ours, the Pearson correlations were close to .50.[4,21,27] According to some other studies, correlation was close to .25 or not significant at all.[40,41,42] Agreement was irrespective of patient's age or health, and of psychological phenomena (cognitive or emotional). However, our study revealed that with low positive emotions scores older children agreed less with their parents than younger children. Conversely, with high positive emotions scores, the older children agreed more with their parents.

The TACQOL social functioning scale is intended to cover social HRQoL. According to the HRQoL as well as the HS tests in this study, mean of children's scores and parent scores were similar. A comparable result was reported in other studies amongst children, adults and elderly people.[4,21,39] However, in our study a relation was found between the level of HRQoL scores and agreement. Furthermore, the correlations in all studies were modest. The Pearson correlation of these studies is close to .50.[4,21,27] Age of the parents related to agreement on social HRQoL. With low scores younger parents agreed more with their children, with high scores older parents agreed more with their children.

The results of this explorative study indicate that parent reports cannot be substituted for child reports. One might consider parents failing as informants about the child and children as lacking time perspective. Yet, both the child and parent reports proved to be valid. An important result is that agreement relates to the magnitude of the HRQoL scores. For several scales agreement relates to background variables like gender, age, temporary illness, visits to a physician and age of the parents. The child scores appear to be less extreme than the parent scores. When parents are very pessimistic, children seem to say "it isn't so bad", and when parents are very optimistic, children seem to say "it isn't that good". Children with disturbed health seem to place more emphasis on making these statements. Future studies using a longitudinal design should enable the following objects to be investigated: individual child-parent differences, the impact of the proxy's gender in combination with the gender of the child, the HRQoL agreement stability, the effect of changing health status, the effect of rating experience, and the effect of feedback to parent and child about their disagreement. Furthermore, if the observability is influencing agreement, we need to study what exactly parent and child do observe. Until these studies are performed and a consensus is reached about the consequences of the MTMM modelling results, techniques could be developed to help clinical decision making. In conclusion, the parent report may provide a substitute for children's HRQoL at a group level, but we have shown that large differences can exist in proxy agreement on individual child-parent level.

3.6 Reference List

1. Ebrahim S. Clinical and public health perspectives and applications of health-related quality of life measurement. Soc Sci Med 1995; 41: 1383-1394.
2. Magaziner J, Simonsick EM, Kashner TM, Hebel JR. Patient-proxy response comparability on measures of patient health and functional status. J Clin Epidemiol 1988; 41: 1065-1074.
3. Bullinger M, Ravens-Sieberer U. Stand der Forschung zur gesundheitsbezogenen Lebensqualität von Kindern: eine Literaturanalyse [German: State of the art of health-related quality of life research in children: a literature review]. Präv Rehab 1995; 7: 106-121.
4. Epstein AM, Hall JA, Tognetti J, Son LH, Conant L. Using proxies to evaluate quality of life. Med Care 1989; 27: S91-S98.
5. Theunissen NCM. Living with end-stage renal failure: A developmental task for the adolescent. In: Ter-Laak JJF, ed. Developmental tasks: Research program developmental psychology Utrecht-Moscow. Utrecht: Utrecht University, 1994: 95-114.
6. Huisman J, Slijper FME, Sinnema G, et al. Psychosocial functioning of short prepubertal children and girls with Turner's syndrome. Acta Paediatr Scand 1991; 377: 162-162.
7. Siegel PT, Clopper R, Stabler B. Psychological impact of significantly short stature. Acta Paediatr Scand 1991; 377: 14-18.

8. Eiser C. Making sense of chronic disease. The eleventh Jack Tizard memorial lecture. J Child Psychol Psychiat 1994; 35: 1373-1389.
9. Kaplan RM, Anderson JP. A general health policy model: update and applications. Health Serv Res 1988; 23: 203-235.
10. Rogerson RJ. Environmental and health-related quality of life: conceptual and methodological similarities. Soc Sci Med 1995; 41: 1373-1382.
11. Jordan TE. Estimating the quality of life for children around the world: NICQL '92. *Soc Indicat Res* 1993; **30**: 17-38.
12. Lindström B. Measuring and improving quality of life for children. In: Lindström B, Spencer N, eds. *Social Paediatrics.* Oxford: Oxford University Press, 1995: 571-585.
13. Sprangers MAG, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: A review. *J Clin Epidemiol* 1992; **45**: 743-760.
14. WHOQOL Group. The World Health Organization Quality of Life Assessment (WHOQOL): Position paper from the World Health Organization. *Soc Sci Med* 1995; **41**: 1403-1409.
15. Torrance GW. Utility approach to measuring health-related quality of life. *J Chronic Dis* 1987; **40**: 593-603.
16. Cella DF, Tulsky DS. Quality of life in cancer: Definition, purpose, and method of measurement. *Cancer Invest* 1993; **11**: 327-336.
17. Sweeny AJ, Creer TL. Health-related-quality-of-life assessment in medical care. *Dis Month* 1995; **16**: 5-64.
18. Bowling A, Brazier J. 'Quality of Life' in social science and medicine: Introduction. *Soc Sci Med* 1995; **41**: 1337-1338.
19. Baker GA, Smith DF, Dewey M, Jacoby A, Chadwick DW. The initial development of a health-related quality of life model as an outcome measure in epilepsy. *Epilepsy Res* 1993; **16**: 65-81.
20. Ferell BR, Hassey Dow K, Grant M. Measurement of the quality of life in cancer survivors. *Qual Life Res* 1995; **4**: 523-531.
21. Hays RD, Vickrey BG, Hermann BP, et al. Agreement between self reports and proxy reports of quality of life in epilepsy patients. *Qual Life Res* 1995; **4**: 159-168.
22. Vogels T, Theunissen NCM, Verrips GH, Koopman HM, Verloove-Vanhorick SP, Kamphuis RP. Het meten van kwaliteit van leven bij kinderen met chronische aandoeningen (Assessing health related quality of life in chronically ill children). *TIAZ* 1996; **3**: 104-111.
23. Rosenberg R. Health-related quality of life between naturalism and hermeneutics. [Review]. *Soc Sci Med* 1995; **41**: 1411-1415.
24. Amir M, Bar-On DM, Penso R. Positive-Negative Evaluation (PNE) scale: A new dimension of the subjective domains of quality of life measure. *Qual Life Res* 1996; **5**: 73-80.
25. Sprangers MAG, Rozemuller N, van-den-Berk MBP, Boven SV, van-Dam FSAM. Response shift bias in longitudinal quality of life research. *Qual Life Res* 1994; **3**: 49-49.
26. Vogels T, Verrips GH, Verloove-Vanhorick SP, et al. Measuring health-related quality of life in children: the development of the TACQOL-Parent Form. *Qual Life Res* 1998; **7**: 457-465.
27. Bruil J, Maes S. Assessing quality of life among children with a chronic illness: The development of a questionnaire. Poster presented at the first Dutch conference on Psychology & Health, Kerkrade, The Netherlands, 6-8 November, 1995; 1-1.
28. Bellman MH, Paley CE. Pain control in children: Parents underestimate children's pain. *Brit Med J* 1993; **307**: 1563-1563.
29. van-Westerlaak JM, Kropman JA, Collaris JWM. *Beroepenklapper [Dutch: Professional index].* Nijmegen: Instituut voor Toegepaste Sociologie [Institution for Applied Sociology], 1975.
30. Verrips GH, Vogels AGC, Koopman HM, et al. Measuring health-related quality of life in a child population. *Eur J Public Health* 1999; **9**: 188-193.
31. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 307-310.
32. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures: Statistics and strategies for evaluation. *Control Clin Trials* 1991; **12**: 142S-158S.
33. Marshall GN, Hays RD, Nicholas R. Evaluating agreement between clinical assesment methods. *Inter J Methods Psychiatr Res* 1994; **4**: 249-257.

34. Steiger JH. Tests for comparing elements of a correlation matrix. *Psychol Bull* 1980; **87:** 245-251.
35. Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol Bull* 1959; **56:** 81-105.
36. Schmitt N, Stults DM. Methodology review: Analysis of multitrait-multimethod matrices. *Appl Psychol Measure* 1986; **10:** 1-22.
37. Stacy AW, Widaman KF, Hays R, DiMatteo MR. Validity of self-reports of alcohol and other drug use: A multitrait-multimethod assessment. *J Personality Soc Psychol* 1985; **49:** 219-232.
38. Dunn G, Everitt B, Pickles A. Multitrait-multimethod and multiple indicator multiple cause models. In: Dunn G, Everitt B, Pickles A, eds. *Modelling, covariances and latent variables using EQS.* London: Chapman & Hall, 1994: 77-87.
39. Ennett ST, DeVellis BM, Earp JA, Kredich D, Warren RW, Wilhelm CL. Disease experience and psychosocial adjustment in children with juvenile rheumatoid arthritis: Children's versus mothers' reports. *J Pediatr Psychol* 1991; **16:** 557-568.
40. Epkins CC. Teachers' ratings of inpatient children's depression, anxiety, and aggression: A preliminary comparison between inpatient-facility and community-based teachers' ratings and their correspondence with children's self-reports. *J Clin Child Psychol* 1995; **24:** 63-70.
41. Achenbach TM, McConaughy SH, Howell CT. Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychol Bull* 1987; **101:** 213-232.
42. Zivcic I. Emotional reactions of children to war stress in Croatia. *J Am Acad Child Adolesc Psychiatry* 1993; **32:** 709-713.

4 Health related quality of life of children with a chronic illness: Parent versus child report

H.M. Koopman, N.C.M. Theunissen, T.G.C. Vogels, K.A.H. Zwinderman, S.P. Verloove-Vanhorick, J. M. Wit

4.1 Abstract

This study evaluates the agreement between child and parent reports on children's Health Related Quality of Life (HRQoL) in a sample of 416 Dutch children (8 to 15 years) with a chronic disease. Both children and their parents completed a 56 item questionnaire (TACQOL) with seven eight-item scales: physical complaints, motor functioning, autonomy, cognitive and social functioning, positive and negative emotions. The correlations between child and parent reports varied from -.10 to .99 amongst the various chronic conditions. Children reported lower HRQoL on the physical complaints, motor functioning and positive emotions scales. Parents reported lower HRQoL on the social, and negative emotions scales. The child and the parent provide different information on HRQoL. Knowledge of both judgements seems necessary in the care of children with a chronic illness and their parents. *Key words: health related quality of life, proxy, children, chronic illness.*

4.2 Introduction

The definition of health provided by The World Health Organisation (WHO) as "a state of complete physical, mental, and social well being, and not merely the absence of disease or infirmity" [1] has been highly influential in defining the quality of life (QoL) construct. The physical, mental and social dimensions have remained central to the construct of QoL. Health-related quality of life (HRQoL) was initially developed and operationalised in adult illness populations. The concept refers to the specific impact of an illness, injury or medical treatment on an individual's QoL. A general criticism of both adult and paediatric HRQoL is the absence of theoretically driven frameworks to guide research. The multidimensional definition of health proposed by WHO[1] is consistent with a biopsychosocial model of functioning. This model emphasise the interdependent relationships among biological, psychological, and social functioning of a child. The HRQoL construct can be viewed here as a method of translating a child's

experience of illness into a quantifiable outcome. This outcome can be integrated into medical practice and research.[2]

One of the important questions in the field of HRQoL assessment with children is the issue of which informant to ask. Parents are commonly used as the major informants in paediatric HRQoL assessments.[3] However, parental reports may be significantly affected by their own anxiety and adjustment level. Furthermore, because parents are not with their children in every setting in which the child functions, their reports may be incomplete. It may overlook the child's subjective experience and perceptions of life quality.[2] However, the use of HRQoL assessment in (young) children can be problematic because children may lack the necessary language skills, as well as the cognitive abilities to interpret the questions, and the long-term view on events.[4] Agreement between child and parent HRQoL reports has already been investigated in a normal population.[5] Therefore, this study evaluates in a sample of children with chronic disorders to what extent parents and children agree on the child's HRQoL.

4.3 Method

4.3.1 Sample

The analytical sample consisted of 416 children with a chronic illness between 8 to 15 years old (213 boys and 203 girls), and their parents. Both children and their parents received a questionnaire by mail. The children with coeliac disease and metabolic disorder were randomly selected from the membership records of the Dutch Coeliac Patients Society and the Dutch Society of Children with Metabolic Diseases. The children with idiopathic short stature were randomly selected from a group of children that formerly participated in a multicenter growth hormone study. The other groups of children with a chronic illness in this study (asthma, rotation plasty for osteosarcoma, diabetes, juvenile chronic arthritis, leukaemia and spinal cord injury) received clinical care in the Leiden University Medical Centre (LUMC). Criteria about informed consent and anonymity were met in accordance with the LUMC Medical Ethics Committee.

4.3.2 Measures

Data on HRQoL were collected using the TNO-AZL Children's Quality of Life questionnaire (TACQOL). The instrument was developed as a generic instrument

intended for broad use with many types of diseases, treatments, and groups of individuals across the core HRQoL domains.[6] Two parallel questionnaires for child's HRQoL were available with identical items: a child form (CF) and a parent form (PF). The items were adjusted to the type of informant: 'Has your child had...' in the PF and 'Have you had...' in the CF. The TACQOL contains seven scales of eight items each: physical complaints, motor functioning, autonomy, cognitive functioning, social functioning, positive and negative emotions. A concretely and specifically formulated problem, if reported, leads to a question about the emotional response. Examples of items are given in Table 1. Items were scored on a 0-4 scale. The 8 item scales were scored on a 0-32 scale, with higher scores representing better HRQoL. Items on (positive and negative) emotional functioning were scored on a 0-2 scale (0 = often, 1 = sometimes, 2 = never) and scale scores ranged from 0-16.

Table 1. Examples of scale-items of the TACQOL Parent form

Scale	Item example (Parent form)
Physical complaints	Has your child had a headache?
Motor functioning	Has your child had difficulty walking?
Autonomy	Did your child have difficulty dressing him/her self?
Cognitive functioning	Did your child have problems understanding his/her schoolwork?
Social functioning	My child was quiet and not talkative with us, parents
Positive emotions	Enthusiastic
Negative emotions	Worried

Response categories: 'Never' / 'Sometimes' / 'Often' . If the answer is 'Sometimes' or 'Often', the next question in the first five scales: is: 'During this my child felt'(PF)/'During this I felt'(CF): '(Very) good' / 'Not so well' / 'Rather bad' / 'Bad'

4.3.3 Analytic Strategy

The results of the children's and parents' reports on group level were summarised with the mean expressed as a percentage of maximum score (to enable comparison between HRQoL group means). The differences in mean between child and parent reports were tested with paired student's t-tests. Agreement between children and parents was further quantified using Pearson correlation coefficients (PCC) and intraclass correlations (ICC). Partial correlations between child and parent reports were calculated for each chronic condition. The correlations were controlled for gender and age because these variables differed substantially between disorders. Linear regression analyses were computed for every scale, to reveal possible relationships between child and parent reports and the variables gender and age.

Table 2. Agreement between child and parent reports on HrQOL (TAQOL-CF and PF)

Scale	M Child	SD Child	M Parent	SD Parent	t	PCC	ICC
Physical complaints	77	(16)	82	(14)	**	.64 *	.60
Motor functioning	89	(14)	90	(14)	**	.64 *	.64
Autonomy	96	(9)	95	(9)		.62 *	.62
Cognitive functioning	87	(13)	86	(14)		.62 *	.62
Social functioning	91	(10)	90	(11)	**	.57 *	.57
Positive emotions	84	(17)	88	(17)	**	.46 *	.44
Negative emotions	72	(17)	70	(15)	**	.55 *	.54

Mean score and standard deviation score are recoded (range 0-100) to allow for comparison of scales with different range of scores. High score = high HrQOL, PCC = Pearson coefficient correlations, ICC = intraclass correlations; * $p \leq .05$; ** paired student's t-test: $p \leq .05$

4.4 Results

Cronbach's alpha coefficient range of HRQoL scales was .75 to .89 (PF) and .68 to .84 (CF). Only the CF social functioning scale showed an alpha below .70 (.68). In Table 2 the agreement between child and parent reports is shown at group level. The difference is largest on physical complaints and positive emotions, which indicates less agreement between child and parent report. Children reported lower HRQoL on the physical complaints, motor functioning and positive emotions scales. Parents reported lower HRQoL on the social and negative emotions scales. Pearson correlations ranged from .46 to .64 ($p < .001$). Intraclass correlations were alike and ranged from .44 to .64.

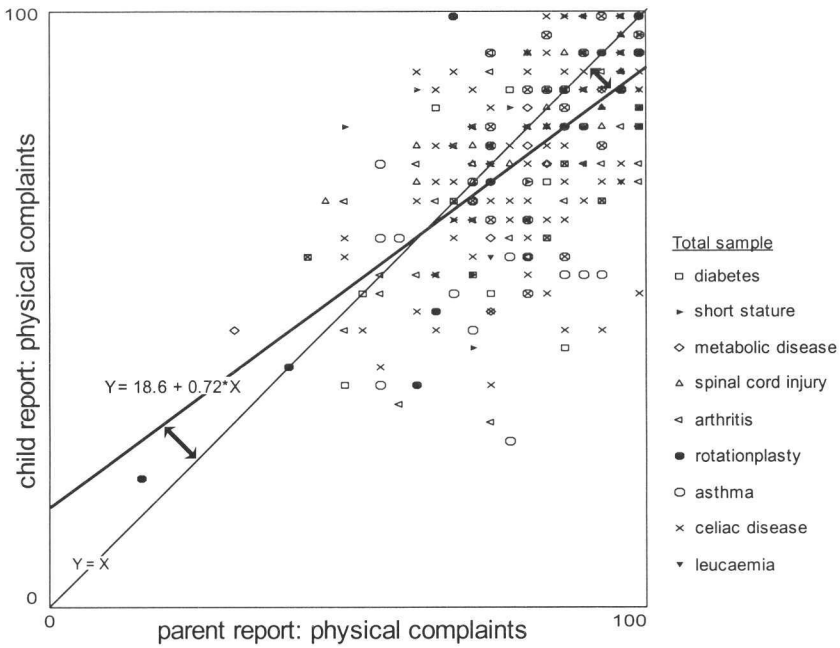
Table 3. Partial correlations (controlled for sex and age) between child and parent HrQOL reports

Chronic Disease	N	PhC	Mot	Aut	Cog	Soc	Pos	Neg
Diabetes Mellitus	(23)	.59 *	.03	.40	.51 *	.14	.51 *	.45
Short stature	(19)	.40	.31	.16	.70 *	.20	-.10	.53 *
Metabolic Disease	(30)	.83 *	.99 *	.86 *	.98 *	.23	.16	.51
Spinal Cord Injury	(22)	.88 *	.54	.94 *	.32	.69 *	.54 *	.74 *
Juvenile Chronic Arthritis	(85)	.68 *	.76 *	.73 *	.51 *	.73 *	.55 *	.41 *
Rotation plasty	(14)	.84 *	.57	.51	.66	.64	.38	.64 *
Asthma	(55)	.46 *	.34	.10	.72 *	.60 *	.67 *	.56 *
Coeliac disease	(153)	.59 *	.55 *	.59 *	.64 *	.61 *	.47 *	.61 *
Leucaemia	(15)	.57 *	.43	.82 *	.70 *	.08	.30	.69 *

PhC = Physical complaints, Mot = Motor functioning, Aut = Autonomy, Cog = Cognitive functioning, Soc = Social functioning, Pos = Positive emotions, Neg = Negative emotions

Table 3 presents the partial correlations (controlled for gender and age) between the children and their parents on the HRQoL scales for every chronic condition. Although PCCs for the whole sample were moderate (.46 - .64), and 43 out of 63 correlation coefficients were above .50, the range of partial correlations was wide. Considering motor functioning in the various chronic conditions, for instance, the child-parent correlation was .03 (ns) in the diabetes group, and .99 ($p<.001$) in the metabolic disease group.

Figure 1. Regression on physical complaints together with the score of individual child-parent pairs



In Figure 1 the regression on physical complaints is given, together with the scores of individual child-parent pairs. The greater the distance between the regression line and the Y=X line, the lower is the agreement between parents and children. If child and parent both reported a high HRQoL (right-hand top corner), then the child scores were relatively lower than the parent scores. If both reported a low HRQoL (left-hand bottom corner) child scores were relatively higher. Overall, child scores were less extreme than the parent scores. The six other TACQOL scales followed the same pattern. The child's gender was related to agreement on physical functioning and

negative emotions scales ($p \leq .05$). If HRQoL scores were high, boys were more in agreement with their parents than girls. The child's age was not related to agreement between child and parent reports.

4.5 Discussion

In this sample of children with a chronic illness and their parents, we found HRQoL patterns comparable with the results of the former study in the open population.[5] Firstly, in both studies the HRQoL scores showed a ceiling effect and were generally good for both children and parents. Secondly, the mean differences between the children and the parents were comparable between studies, although different in detail. In Theunissen et al.[5] children on average reported a significant lower TACQOL score than the parents on the physical complaints, motor functioning, positive emotions, autonomy and cognitive functioning scales. In the present study the same results were found on the first three scales, but not on the autonomy and cognitive functioning scales. On the social functioning and negative emotions scales, Theunissen et al.[5] found no statistical differences, whereas in this study children reported significantly higher HRQoL than their parents. Thirdly, correlations between children and parents were moderate in both studies. Lastly, in both studies child scores were less extreme than parent scores. When parents were pessimistic, children were less despondent, and when parents were optimistic, children were less confident.

According to previous studies, agreement was relatively good for observable measures.[7-9] Evidence of physical and social activity is generally available to the parent. Therefore, the parent could describe the child's status as well as the child would. In contrast, cognitive functioning, for example, is more likely to be experienced as private thoughts and opinions on the part of the child. This is difficult to observe for the parent, and as a result agreement could be low. This study showed different results. Parent and children reports agreed on cognitive functioning and not on social, physical and motor functioning. Parents seemed to be well informed about their child's results at school.

An explanation of the contrasting results on social, physical an motor functioning flows from the context in which parents make inferences about their child's HRQoL. Their interest may be in confirming prevailing beliefs and theories about the future impact of the disease on HRQoL of their developing children. The lack of correspondence between parent and child could be related to the influence of 'halo

effects' on parents' ratings. Parents seem to express the thought 'given the circumstances, my child does reasonably well'. For some scales (physical functioning and negative emotions) this effect is stronger in regard to their daughters. With regard to emotional functioning, negative feelings and opinions may be more likely to be displayed than positive ones.[7] For example, if a doctor's office visit went well, a child might make no comment to a parent, but if it went badly the parent might hear about it in some detail. A parent would therefore infer more overall negativity, than the child actually experiences. Furthermore, there is a tendency for observers to give more weight to negative than to positive information when forming impressions of others. Negative expressions, such as weeping or complaining, are surely more salient and memorable than more positive ones. Therefore, even if children revealed their negative and positive feelings equally, parents' impressions would be more influenced by the negative, with the result that their opinion of the child's condition would be biased negatively. The results of this study were in line with this hypothesis. Children reported less negative emotions.

Correlations of the total group of children with a chronic illness and their parents were modest although the range of the correlations in the various chronic conditions varied. In the studies previously mentioned, correlations between self and proxy reports varied from low to high amongst children, adults and elderly people. When the proxy lived with the subject, correlations tended to be stronger.[8] However, all children lived with their parents, so this does not explain the wide range of correlations we found in the various chronic conditions. Future research in larger groups of children with a chronic condition have to confirm these correlations.

In conclusion, the results of this study indicate that parent reports cannot be substituted by child reports. There is a large variation in parent - child agreement on the various scales and the various chronic conditions. However, child scores appear to be less extreme than parent scores. When parents are pessimistic, children seem to say "it isn't so bad", and when parents are optimistic, children seem to say "it isn't that good". Knowledge of the feelings and opinions of both the parent and the child may be helpful to improve the care of children with a chronic illness.

4.6 Reference List

1. World Health Organization. *Constitution of the World Health Organization.* Geneva: World Health Organization, 1948.
2. Drotar D. *Measuring health-related quality of life in children and adolescents: Implications for research and practice.* Mahwah, New Jersey: Lawrence Erlbaum Associates, 1998.
3. Haverkamp F, Noeker M. 'Short stature in children - a questionnaire for parents': a new instrument for growth disorder-specific psychosocial adaptation in children. *Qual Life Res* 1998; **7:** 447-455.
4. Kamphuis RP. The concept of quality of life in pediatric oncology. In: Aaronson NK, Beckman J, eds. *The quality of life of cancer patients.* New York: Ravenpress, 1987.
5. Theunissen NCM, Vogels AGC, Koopman HM, et al. The proxy problem: child report versus parent report in health-related quality of life research. *Qual Life Res* 1998; **7:** 387-397.
6. Vogels T, Verrips GH, Verloove-Vanhorick SP, et al. Measuring health-related quality of life in children: the development of the TACQOL-Parent Form. *Qual Life Res* 1998; **7:** 457-465.
7. Ennett ST, DeVellis BM, Earp JA, Kredich D, Warren RW, Wilhelm CL. Disease experience and psychosocial adjustment in children with juvenile rheumatoid arthritis: Children's versus mothers' reports. *J Pediatr Psychol* 1991; **16:** 557-568.
8. Epstein AM, Hall JA, Tognetti J, Son LH, Conant L. Using proxies to evaluate quality of life. *Med Care* 1989; **27:** S91-S98.
9. Sprangers MAG, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: A review. *J Clin Epidemiol* 1992; **45:** 743-760.

5 The TAPQOL. A Health Related Quality of Life instrument for 1 to 5-year old children

M. Fekkes, N.C.M. Theunissen, E. Brugman, S. Veen, E.G.H. Verrips, H.M. Koopman, A.G.C. Vogels, J. M. Wit, S. P. Verloove-Vanhorick

5.1 Abstract

The 46-item TNO-AZL Pre-school children Quality of Life (TAPQOL) questionnaire was developed to meet the need for a reliable and valid instrument for measuring Health-Related Quality of Life (HRQoL) in pre-school children. HRQoL was defined as health status in 13 domains plus the emotional responses to problems in health status. The TAPQOL has to be filled in by the parents. The aim of this study was to evaluate the psychometric performance of the TAPQOL.

A sample of 121 parents of preterm children completed the TAPQOL questionnaire (response rate 88%) as well as 362 parents of children form the open population (response rate 60%). On the base of Cronbach's alpha, item-rest correlation, and principal component analysis, the TAPQOL scales were constructed from the data of the preterm children sample.

The psychometric performance of these scales was evaluated for both the preterm children sample and the open population sample. Cronbach's alpha ranged from .66-.88 for the preterm children sample and from .26-.84 for the open population sample. The unidimensionality of the separate scales was confirmed by principal component analysis for both the preterm children sample and the open population sample. Pearson's correlation coefficients between scales were on the average low. T-test analyses showed that very preterm children and children with chronic diseases had lower scores on the TAPQOL scales than healthy children, indicating a worse quality of life. This study shows that the TAPQOL can be a good instrument to measure HRQoL in pre-school children, but more research is needed to evaluate the psychometric performance of the TAPQOL in different clinical populations. *Key words: health related quality of life, pre-school children, preterm birth, quality of life instruments.*

5.2 Introduction

The application of Health Related Quality of Life (HRQoL) as an outcome measure in medical care has increased enormously during the last decades. For children however, less HRQoL-instruments are developed than for adults. Especially for pre-school children only a few instruments are available, such as for example the RAND and the FS(II)R.[1,2] These instruments are certainly useful, although they do not measure HRQoL as a multi-dimensional construct. It is suggested by several authors that a multi-dimensional approach is preferred when measuring health outcomes, in order to capture the full range of manifestations of health and illness.[3-6] This lack of multi-dimensional HRQoL instruments for children may be due to two major methodological problems. The first one is the proxy problem. Young children are not capable of filling in a questionnaire themselves, so the help of a proxy is needed. For children's HRQoL, the parents are in general the best proxies but the use of a proxy can decrease the reliability, validity and accuracy of the information.[3,7,8] The second problem is the age-specific development of children in several health related domains. It is impossible to include the whole development of 0-18 year old children in one HRQoL-instrument. Age-specific instruments are needed when measuring HRQoL in children.

The need for a multi-dimensional HRQoL instrument for pre-school children is great. Such an instrument will be useful in measuring the impact of diseases and treatments on the different domains of young children's lives. To meet this need for a reliable and valid instrument measuring HRQoL in very young children, the TNO-AZL Pre-school children Quality of Life-questionnaire (TAPQOL) was developed: a generic, multi-dimensional HRQoL instrument for 1 to 5-year old children. The TAPQOL covers 13 domains, thought to be relevant for children between 1 and 5.

It has been pointed out that, when measuring HRQoL, it is important not only to include health status but also the negative emotions that are elicited by limitations in health status.[9-11] In accordance with these suggestions, in developing the TAPQOL, HRQoL is defined in relation to but clearly distinguished from the concept of Health Status. Health Status refers to all kind of actual problems and limitations in functioning which are directly or indirectly related to one's disease. HRQoL includes the person's emotional responses to such problems and limitations. Consequently, the TAPQOL assesses functional problems weighted by the degree to which a child shows negative emotions to such problems. The TAPQOL has to be filled in by one of the parents of the child.

Aim of the present study was to evaluate the psychometric performance of the TAPQOL. A set of items was administered to parents of a group of preterm children, and to parents of a group of children from the general population. On base of the data from the preterm children sample we selected items to form a final measure. We evaluated that measure on psychometric criteria, including reliability, construct validity, concurrent validity and criterion validity.

5.3 Material and methods

5.3.1 Procedure sample 1

During an 8 month period, all parents of 138 preterm children, aged 1-5, who visited the out-patient neonatology clinic of the Leiden University Medical Centre in the Netherlands, were invited to complete a set of questions. Completed questionnaires could be returned by mail.

5.3.2 Procedure sample 2

Six Well Baby Clinics in various parts of the Netherlands drew a random sample of 100 children each, from the open population. These 600 children were 1 to 5-years old. The children's parents were invited to complete a set of questions. Completed questionnaires could be returned by mail.

5.3.3 TAPQOL-questionnaire

Scales

On the basis of literature and discussions with HRQoL-experts, paediatricians, psychologists and mothers of 1-5-year old children, 13 domains thought to be relevant for the HRQoL of 1-5-year old children were included in the TAPQOL. For every domain a number of items was constructed. These domains are: sleeping, appetite, lung problems, stomach problems, skin problems, eating disorders, motor functioning, social functioning, problem behaviour, communication, anxiety, positive mood and Liveliness. The three domains motor functioning, social functioning and communication are applicable only to children of 1.5 years and older. Feasibility of a preliminary version of the TAPQOL questionnaire was evaluated in a group of mothers of 1-5-year old

children. Several items were deleted or rephrased after this feasibility pilot. The final TAPQOL consists of 35 items for children younger than 1.5 years and 46 items for children aged 1.5 - 5 years. The period of recall is 3 months. An English version, translated from Dutch in accordance with international guidelines[12], is available on request. Examples of items are given in Table 1.

Table 1. Examples of scale-items of the TAPQOL

Scale	Item example (In the last three months...)
Sleeping	Did your child sleep restlessly?
Appetite	Was your child's appetite poor?
Lungs	Has your child been short of breath?
Stomach	Has your child had stomach-ache or abdominal pain?
Skin	Has your child had dry skin?
Eating disorders	Did your child have difficulty swallowing food?
Motor functioning	Did your child have difficulty with walking?
Social functioning	Was your child at ease with other children?
Problem behaviour	Was your child aggressive?
Communication	Did your child have difficulty in making it clear what he/she wanted?
Anxiety	Was your child anxious?
Positive mood	Was your child in good spirits?
Liveliness	Was your child energetic?

Item-construction

In the scales measuring sleeping, appetite, lung problems, stomach problems, skin problems, eating disorders, motor functioning, and communication, items consist of two questions. In these items, first the frequency of a specific complaint or limitation during the last three months is recorded. If such a problem has occurred, the emotional response of the child to this problem is assessed. An example is presented in Figure 1. Each item is encoded into one single score, ranging from 0 to 4. A score of 0 is given if the child has no problem; a score of 1 if the child has a problem (i.e. occasionally or often) but the child feels "fine" during the problem; a score of 2 if the child has a problem and feels "not so good"; a score of 3 if the child has a problem and feels "quite bad"; and a score of 4 if the child has a problem and feels "bad". Thus, functional problems are weighted by their emotional impact.

Figure 1. Item example TAPQOL

In the last three months, has your child been ...

Short of breath?	❑ never	❑ occasionally	❑ often		
		At that time, my child felt:			
		❑ fine	❑ not so good	❑ quite bad	❑ bad

In the scales measuring social functioning, problem behaviour, anxiety, positive mood and Liveliness, the items measure only the frequency of a specific complaint or limitation during the last three months. In the items of these domains it is not asked how the child feels, because items of these domains already imply a positive or negative emotional state. In these scales, each item score is ranging from 0 to 2. An example is presented in Figure 2. A score of 0 is given if the child "never" has a problem; a score of 1 if the child "occasionally" has a problem; and a score of 2 if the child "often" has a problem.

Figure 2. Item example TAPQOL

How was your child in the last three months?

In good spirits	❑ never	❑ occasionally	❑ often

5.3.4 FS(II)R and other variables

Apart from the TAPQOL-questionnaire, parents were asked to fill in questions on demographics, children's chronic diseases, common illnesses and recent medical treatment of their child. Further, the parents version of the FS(II)R, a 14-item measure of children's health status[1,2] was used, to evaluate concurrent validity.

5.3.5 Analyses

Using the data of the preterm-children (sample 1) the final TAPQOL-scales were constructed, based on Cronbach's alpha, item-rest correlation, and principal component analysis. The number of items was reduced from 99 to 46 items in the final TAPQOL-questionnaire. The number of items per scale now varies from 3 to 7. Scale scores were obtained by adding item scores within scales, and transforming crude scale scores linearly to a 0-100 scale, with higher scores indicating better quality of life. Reliability and construct validity were evaluated for these scales.

The psychometric performance of the scales was evaluated in the sample from the open population (sample 2). Reliability was evaluated by calculating Cronbach's alpha's for each scale.

Construct validity was evaluated by means of several procedures. Because the item scores are a combination of functional problems weighted by the degree to which a child shows negative emotions to such problems, the ordinality of these final item scores was checked. Homogeneity Analysis (HOMALS) was performed to check the ordinality of the scores. We assumed that the category quantifications should show a distinct order in the categories. The unidimensionality of the separate scales was evaluated by Principal Component Analysis and by calculating item-rest and item-scale Pearson's correlation coefficients. The relationships between the 13 different domains of the TAPQOL was evaluated by calculating Pearson's correlation coefficients between the TAPQOL-scales. The validity of the distinction between health status problems and the emotional experience of such problems was evaluated by calculating the total number of health status problems and the percentage of such problems that led to a negative feeling for the child. Concurrent validity was evaluated by calculating Pearson's correlation coefficients between the TAPQOL-scales and the FS(II)R-score. Criterion validity was evaluated by relating the TAPQOL scale scores to two dichotomous criteria by means of T-tests. Firstly, healthy children from the open population sample were compared with preterm children with a gestational age less than 32 weeks, the last being a group with an expected less good health than the open population.[13-15] Secondly, children in the open population sample suffering from any chronic disease were compared with children that did not suffer from a chronic disease.

5.4 Results

5.4.1 Sample 1

Respondents

Response rate in the group of parents of preterm children was 88% (n =121). All children had a gestational age less than 37 weeks. The sample included 65 children with a gestational age less than 32 weeks. The percentage of girls of the sample was 42%. The age of the children ranged from 10 to 60 months: 45% of the children was between 10 and 24 months old, 40% between 24 and 36 months old , 12% between 36 and 48 months old and 13% between 48 and 60 months old.

Distribution and reliability

The categorised percentual distributions presented in Table 2 show that the TAPQOL-scales scores were skewed. The percentage of missing was low: for most scales less than 5%. Cronbach's alpha ranged from .66 to .88. For only one scale, i.e. 'Stomach problems', the Cronbach's alpha was below .70.

Table 2. Cronbach's alpha and categorised percentual distribution of the TAPQOL scales for the preterm children sample (N = 109)

Scale	No. of items	Alpha's	% Missing	Categorised percentual distribution (Scale: 0 – 100)				
				0-24	*25-49*	*50-74*	*75-99*	*100*
Sleeping	4	.88	0	1	5	31	34	29
Appetite	3	.79	1.8	1	1	11	53	34
Lungs	3	.83	2.8	0	6	16	18	60
Stomach	3	.66	1.8	1	4	14	32	49
Skin	3	.79	0	1	0	7	50	42
Eating disorders	3	.70	1.8	0	1	10	27	62
Motor functioning	4	.77	5.1	0	1	8	41	50
Social functioning	3	.75	5.1	1	0	18	27	54
Problem behaviour	7	.86	0.9	6	9	53	28	4
Communication	4	.80	7.7	0	1	17	58	24
Anxiety	3	.71	0.9	2	5	49	21	23
Positive mood	3	.82	0	0	0	4	3	93
Liveliness	3	.87	0	0	0	6	4	91

Construct validity

The sample was too small to perform HOMALS analysis meaningfully, so this analysis was only performed for sample 2. For each set of items belonging to a scale a separate principal component analysis was performed. All these analyses resulted in only one factor being extracted indicating the unidimensionality of the scales. All items had a factor loading above .50 (Table 3 next page). All except two items had a higher corrected correlation with their own scale (item-rest-correlation) than with other scales (Table 3).

Table 3. Results of principal component analyses and item-scale correlations for the preterm children sample (N = 109)

Scale	No. of items	No. of items with factor loading >.50 [1]	No. of items with highest correlation with their own scale [2]
Sleeping	4	4	4
Appetite	3	3	3
Lungs	3	3	3
Stomach	3	3	3
Skin	3	3	3
Eating disorders	3	3	3
Motor functioning	4	4	4
Social functioning	3	3	2
Problem behaviour	7	7	7
Communication	4	4	4
Anxiety	3	3	2
Positive mood	3	3	3
Liveliness	3	3	3

[1] For every scale a separate principal component analysis was performed; [2] Item-rest correlations were compared with item-other scale correlations

Table 4. Inter-scale correlations of the TAPQOL scales for the preterm children sample (N = 109)

Scale	1	2	3	4	5	6	7	8	9	10	11	12
1 Sleeping	--											
2 Appetite	.29	--										
3 Lungs	.26	.17	--									
4 Stomach	.32	.30	.30	--								
5 Skin	.05	.08	.20	.22	--							
6 Eating disorders	.34	.42	.20	.34	-.10	--						
7 Motor functioning	.06	.21	.20	.21	.25	.23	--					
8 Social functioning	-.11	-.08	.09	.08	.10	.03	.14	--				
9 Problem behaviour	.19	.09	.18	.26	.06	.11	-.01	.19	--			
10 Communication	.38	.40	.21	.28	.06	.33	.07	.10	.19	--		
11 Anxiety	.31	.11	.17	.26	.28	.28	.25	.29	.41	.48	--	
12 Positive mood	.00	.00	.18	.16	.06	-.01	.10	.30	.26	.05	.27	--
13 Liveliness	-.05	.05	.12	.05	-.03	.18	.07	.60	.06	.02	.19	.30

The correlation coefficients between the TAPQOL scales were on average low (Table 4). Most correlation coefficients between scales were below .40. The correlation between the Liveliness-scale and the social functioning-scale was the highest, i.e. .60.

5.4.2 Sample 2

Respondents

Response rate in the open population sample was 60% (n=362). The percentage of girls of the sample was 44%. The age of the children ranged from 10 to 60 months: 34% of the children was between 10 and 24 months old, 30% between 24 and 36 months old , 31% between 36 and 48 months old and 5% between 48 and 60 months old. The sample included 62 children with a chronic illness who had visited the doctor in the last six months. The majority of this group had respiratory problems like asthma (12 children), chronic bronchitis (15 children) and whooping-cough (4 children). Another 13 children suffered from allergy problems.

Table 5. Cronbach's alpha and categorised percentual distribution of the TAPQOL scales for the open population sample (N = 340)

Scale	No. of items	Alpha's	% Missing	Categorised percentual distribution (Scale: 0 – 100)				
				0-24	*25-49*	*50-74*	*75-99*	*100*
Sleeping	4	.84	0.3	0	3	18	51	28
Appetite	3	.76	0.3	0	1	8	64	27
Lungs	3	.84	0.3	1	1	8	9	81
Stomach	3	.51	0.6	0	1	9	25	69
Skin	3	.59	0.3	0	0	5	47	48
Eating disorders	3	.26	0	0	0	2	24	74
Motor functioning	4	.43	2.4	0	0	0	12	88
Social functioning	3	.68	1.7	1	1	10	20	68
Problem behaviour	7	.75	0	1	7	61	29	2
Communication	4	.69	4.2	0	0	3	51	46
Anxiety	3	.63	0	0	1	47	21	32
Positive mood	3	.77	0	0	0	3	2	95
Liveliness	3	.75	0.3	0	0	3	4	93

Distribution and reliability

The categorised percentual distributions presented in Table 5 (see previous page) shows that the TAPQOL-scales scores were skewed. The percentage of missing was low: for all scales less than 5%. For the open population, Cronbach's alpha ranged from .26 to .84 (Table 5). For most scales the Cronbach's alpha was above .60. However, for the scales measuring stomach problems, skin problems, eating disorders and motor functioning, the Cronbach's alpha was very low. This low reliability may have been due to the low prevalence and variance of such problems in the open population sample.

Table 6. Results of principal component analyses and item-scale correlations for the open population sample (N = 340)

Scale	No. of items	No. of items with factor loading >.50 [1]	No. of items with highest correlation with their own scale [2]
Sleeping	4	4	4
Appetite	3	3	3
Lungs	3	3	3
Stomach	3	3	3
Skin	3	3	3
Eating disorders	3	3	2
Motor functioning	4	4	4
Social functioning	3	3	3
Problem behaviour	7	7	7
Communication	4	4	4
Anxiety	3	3	1
Positive mood	3	3	3
Liveliness	3	3	3

[1] For every scale a separate principal component analysis was performed; [2] Item-rest correlations were compared with item-other scale correlations

Construct validity

HOMALS analyses indicated that for 40 out of the 46 items the category quantifications did not violate the assumed ordinality, when categories with a frequency of 10 or more were taken into account. For the six items that violated the assumed ordinality, the violations concerned small differences between two categories.

For each set of items belonging to a scale a separate principal component analysis was performed. All these analyses resulted in only one factor being extracted indicating the unidimensionality of the scales. All items had a factor loading above .50 (Table 6). All except three items had a higher corrected correlation with their own scale (item-rest-correlation) than with other scales (Table 6).

The correlation coefficients between the TAPQOL scales were on average low (Table 7). Most correlation coefficients between scales were below .30. The correlation between the anxiety-scale and the problem behaviour-scale was the highest, i.e. .44.

Table 7. Inter-scale correlations of the TAPQOL scales for the open population sample (N = 340)

Scale	1	2	3	4	5	6	7	8	9	10	11	12
1 Sleeping	--											
2 Appetite	.21	--										
3 Lungs	.06	.15	--									
4 Stomach	.12	.21	.04	--								
5 Skin	.07	.05	.05	.11	--							
6 Eating disorders	.09	.25	.04	.20	.09	--						
7 Motor functioning	.19	.08	.02	-.12	.01	.11	--					
8 Social functioning	-.09	-.05	-.12	-.05	-.06	-.02	-.08	--				
9 Problem behaviour	.18	.16	-.01	.12	.03	-.01	-.01	.04	--			
10 Communication	.08	.04	.11	-.02	-.06	-.07	.09	-.24	.14	--		
11 Anxiety	.31	.24	.10	.21	.10	.01	.01	-.13	.44	.20	--	
12 Positive mood	-.15	-.06	-.22	-.07	-.02	-.03	-.06	.24	-.13	-.14	-.20	--
13 Liveliness	-.02	-.06	-.17	.06	.00	-.03	-.06	.10	-.03	-.03	-.09	.30

The validity of the distinction between health status problems and the emotional experience of such problems was evaluated by calculating the total number of health status problems and the percentage of such problems that led to a negative feeling for the child. Table 8 shows that the number of problems in Health Status that elicited a negative feeling in the child varied greatly among the different scales. For example sleeping, lung and stomach problems elicited much more negative feelings in the child than skin problems or problems with motor functioning did.

Table 8. Total number of problems and number of problems that elicited a negative emotion in the child. Both samples together (N = 449)

Scale [1]	No. of items	Total No. of problems (n = 449)	No. of problems that elicited a negative emotion in the child
Sleeping	4	861	418 (49%)
Appetite	3	703	107 (15%)
Lungs	3	215	173 (80%)
Stomach	3	253	214 (85%)
Skin	3	389	70 (18%)
Eating disorders	3	174	73 (42%)
Motor functioning[2]	4	438	58 (13%)
Communication[2]	4	152	40 (26%)
Total	27	3185	1153 (36%)

[1] The scales Liveliness, positive mood, problem behaviour and social functioning are not presented because for the items of these scales there was no distinction between health status and negative emotions elicited by problems in health status, [2] Scales only for children of 1.5 years and older (N = 354)

Table 9. Pearson's product moment correlation coefficients between TAPQOL-scales and the FS(II)R, for children from both the preterm sample and the general population sample (N = 449)

Scale	FS(II)R-score Preterm children	FS(II)R-score General population children
Sleeping	.41	.39
Appetite	.26	.29
Lungs	.24	.27
Stomach	.48	.18
Skin	.18	.10
Eating disorders	.20	.03
Motor functioning	.18	.06
Social functioning	.16	.22
Problem behaviour	.32	.41
Communication	.33	.23
Anxiety	.43	.36
Positive mood	.22	.36
Liveliness	.33	.24

Concurrent validity

The relationships between the TAPQOL-scale scores and the FS(II)R score were calculated by means of Pearson's correlation coefficients. These correlation coefficients are represented in Table 9. The relationships between the TAPQOL-scales and the FS(II)R-score are low. This is the case for both the preterm children sample and the general population sample. When only the health status part of the TAPQOL-scales is compared with the FS(II)R-score the correlations are not substantially higher.

Table 10. Mean TAPQOL-scale scores for healthy, preterm (gestational age < 32 weeks) and chronically ill children

Scale	M (±SD) healthy children (n = 251)	M (±SD) preterm children (n = 65)	M (±SD) chronically ill children (n = 62)	p^a healthy vs. preterm children	p^a healthy vs. chron.ill children
Sleeping	83.1 (±17)	75.9 (±21)	77.2 (±19)	.004	.018
Appetite	85.9 (±12)	82.6 (±17)	79.6 (±16)	NS	.001
Lungs	97.2 (±9)	85.3 (±20)	78.1 (±28)	< .000	< .000
Stomach	92.6 (±13)	84.2 (±21)	87.6 (±17)	< .000	.013
Skin	92.8 (±10)	90.5 (±15)	88.0 (±13)	NS	.002
Eating disorders	96.4 (±7)	90.5 (±15)	95.8 (±9)	< .000	NS
Motor functioning	98.8 (±4)	90.2 (±13)	98.0 (±5)	< .000	NS
Social functioning	91.4 (±15)	83.7 (±22)	90.0 (±18)	.003	NS
Problem behaviour	67.7 (±15)	63.7 (±21)	67.0 (±16)	NS	NS
Communication	91.7 (±10)	80.9 (±15)	92.1 (±11)	< .000	NS
Anxiety	79.2 (±17)	66.9 (±22)	75.8 (±19)	< .000	NS
Positive mood	98.9 (±6)	98.4 (±6)	97.8 (±9)	NS	NS
Liveliness	98.1 (±8)	96.4 (±12)	97.0 (±10)	NS	NS

[a]p-value of T-test of for differences of means; Higher scores indicate better Health-Related Quality of Life; NS = Not significant (p>.05)

Criterion validity

Criterion validity of the TAPQOL was evaluated by comparing the TAPQOL-scale scores of a group of less healthy with healthy children. Firstly, preterm children with a gestational age less than 32 weeks from preterm children sample were compared with healthy children from the open population sample. Secondly, chronically ill children from the open population sample (mostly children with respiratory problems) were compared with the group of healthy children from the open population sample. The TAPQOL-scale scores for each group are represented in Table 10. Table 10 shows

that preterm children and children with a chronic disease had significantly lower scores on several of the TAPQOL-scales, indicating a worse quality of life. Almost all scales, except the scales problem behaviour, positive mood and Liveliness, were significantly related to at least one of the criteria (being chronically ill or preterm born): these scales differentiated between a healthy and a less-healthy group of children.

5.5 Discussion

The aim of this study was to evaluate the psychometric performance and validity of the TAPQOL, a recently developed generic Health-Related Quality of Life instrument for 1-5-year old children. The reliability of most TAPQOL scales was good in the preterm children sample, with Cronbach's alphas ranging from .66 to .89. Reliability in the open population sample was satisfying for most scales. However, the Cronbach's alpha for the scales relating to "eating disorders", "stomach functioning" and "motor functioning" were low. This low reliability may have been due to the low prevalence and variance of problems in the open population sample. This explanation gains plausibility if the psychometric performance of the TAPQOL scales in the preterm children sample is taken into account. The prevalence and variance of problems in this group is higher, and the reliability of the scales is better. More research has to be done to evaluate the reliability of the TAPQOL-scales in different samples of children with chronic illnesses. The unidimensionality of the TAPQOL scales was confirmed by principal component analysis. Almost all item-rest correlations were higher than the item correlations with other scales, which also supported the scale structure. The correlations between the scales were low, confirming the multi-dimensional definition of HRQoL. This is in accordance with the conventional approach to HRQoL as a multi-dimensional construct.[5,16] This approach has been confirmed by psychometric evaluations of HRQoL instruments for adolescents and adults.[17-21] The results of our study show that the multi-dimensional approach of HRQoL is also valid for very young children. Because of this approach the TAPQOL is useful to measure a broad scope of different domains. This broad scope of the TAPQOL may be helpful in measuring the impact of disease and treatment on the lives of young children, especially because it is not always clear which domains will be inflicted by disease or treatment. Concurrent validity was evaluated by relating the TAPQOL-scales with the FS(II)R-score. The correlations were low. The absence of a high correlation could be due to the different concepts that are measured: the FS(II)R measures health status[1,2] and the TAPQOL

assesses functional problems in health status weighted by the degree to which the child shows negative emotions in relation to such problems. However, the correlations were not substantially higher when only the functional status part of the TAPQOL was compared with the FS(II)R. Another explanation might be that the TAPQOL measures 13 different scales that each on their own have no high correlation with the 'general health' measured by the FS(II)R.

The assumption that a complaint or limitation in health status does not automatically lead to a negative feeling, as suggested in the literature,[9-11] is confirmed by the results of our study. Only 36% of the problems in health status elicit a negative feeling in the child. This percentage varies greatly among the different scales. For example, only 15% of the problems in the "appetite-scale" leads to a negative feeling in the child, while this figure is 85% for the "stomach functioning-scale". These results indicate the different impact that problems in health status can have on the child's quality of life. It confirms the relevance of the distinction between health status problems and the emotional impact of these problems. Not every problem in health status automatically triggers a bad feeling in the child. Such information may be of great value when the TAPQOL is used to evaluate different therapies or medical interventions for children. Curing health problems is not always possible in such conditions as diabetes mellitus or congenital heart disease, but negative emotional feelings resulting from these conditions may be prevented or reduced.[21] The TAPQOL can be an instrument to measure such a prevention or reduction. Evaluation of the criterion validity of two other HRQoL instruments for children, the FS(II)R and the RAND, show that these two instruments distinguish between well and ill children.[1,2] In our study, criterion validity was evaluated by relating two criteria to the TAPQOL scales. For both criteria, the group with the less good health (i.e., the preterm children and the group with a chronic illness) had a significantly lower score on most of the TAPQOL scales. These results demonstrate that the TAPQOL scales can detect differences between healthy and less healthy children.

In conclusion, the results of this study show that the TAPQOL has both good validity and psychometric performance for the preterm children sample. The psychometric performance of some scales is less satisfying in the open population sample, which might be due to the low prevalence of health status problems in this group. More research must be done to evaluate the psychometric performance of the TAPQOL in other clinical populations. The instrument may be used to assess group differences in pre-school children's HRQoL, for example in studies evaluating the effects of different treatments.

5.6 Reference List

1. Lewis CC, Pantell RH, Kieckhefer GM. Assessment of children's health status. Field test of new approaches. *Med Care* 1989; **27**: S54-S65.
2. Stein RE, Jessop DJ. Functional status II(R). A measure of child health status. *Med Care* 1990; **28**: 1041-1055.
3. Vivier PM, Bernier JA, Starfield B. Current approaches to measuring health outcomes in pediatric research. *Curr Opin Pediatr* 1994; **6**: 530-537.
4. Eisen M, Ware Jr. JE, Donald CA, Brook RH. Measuring components of children's health status. *Med Care* 1979; **17**: 902-921.
5. Aaronson NK. Quality of life: what is it? How should it be measured? *Oncology Huntingt* 1988; **2**: 69-76,64.
6. Fitzpatrick R, Fletcher A, Gore S, Jones D, Spiegelhalter D, Cox D. Quality of life measures in health care. I: Applications and issues in assessment. *BMJ* 1992; **305**: 1074-1077.
7. Weissman MM, Orvaschel H, Padian N. Children's symptom and social functioning self-report scales. Comparison of mothers' and children's reports. *J Nerv Ment Dis* 1980; **168**: 736-740.
8. Sprangers MAG, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: A review. *J Clin Epidemiol* 1992; **45**: 743-760.
9. Gill TM, Feinstein AR. A critical appraisal of the quality of quality-of-life measurements. *JAMA* 1994; **272**: 619-626.
10. Christie MJ, French D, Weatherstone L, West A. The patients' perception of chronic disease and its management: psychosomatics, holism and quality of life in contemporary management of childhood asthma. *Psychother Psychosom* 1991; **56**: 197-203.
11. Collings JA. Epilepsy and well-being. *Soc Sci Med* 1990; **31**: 165-170.
12. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol* 1993; **46**: 1417-1432.
13. Saigal S, Rosenbaum P, Stoskopf B, et al. Comprehensive assessment of the health status of extremely low birth weight children at eight years of age: Comparison with a reference group. *J Pediatr* 1994; **125**: 411-417.
14. Saigal S, Feeny D, Furlong W, Rosenbaum P, Burrows E, Torrance G. Comparison of the health-related quality of life of extremely low birth weight children and a reference group of children at age eight years. *J Pediatr* 1994; **125**: 418-425.
15. Veen S, Ens-Dokkum MH, Schreuder AM, Verloove-Vanhorick SP, Brand R, Ruys JH. Impairments, disabilities, and handicaps of very preterm and very-low-birthweight infants at five years of age. *Lancet* 1991; **338**: 33-36.
16. Aaronson NK. Quality of life assessment in clinical trials: methodologic issues. *Control Clin Trials* 1989; **10**: 195S-208S.
17. Langfitt JT. Comparison of the psychometric characteristics of three quality of life measures in intractable epilepsy. *Qual Life Res* 1995; **4**: 101-114.
18. Apajasalo M, Sintonen H, Holmberg C, et al. Quality of life in early adolescence: a sixteen-dimensional health-related measure (16D). *Qual Life Res* 1996; **5**: 205-211.
19. Verrips GH, Vogels AGC, Verloove-Vanhorick SP, et al. Health-related quality of life measure for children -- the TACQOL. *J Appl Therapeutics* 1997; **1**: 357-360.
20. Theunissen NCM, Vogels AGC, Koopman HM, et al. The proxy problem: child report versus parent report in health-related quality of life research. *Qual Life Res* 1998; **7**: 387-397.
21. Vogels T, Verrips GH, Verloove-Vanhorick SP, et al. Measuring health-related quality of life in children: the development of the TACQOL-Parent Form. *Qual Life Res* 1998; **7**: 457-465.

6 Quality of life in pre-school children in relation to a preterm birth

N.C.M. Theunissen, S. Veen, M. Fekkes, H.M. Koopman,
K.A.H. Zwinderman, E. Brugman, J.M. Wit

6.1 Abstract

Objective: The relationship of a preterm birth with Health Related Quality of Life (HRQoL) was examined for children aged 1-4. **Sample:** Three gestational age groups were selected, < 32 weeks (n=65), 32-37 weeks (n=41), ≥ 37 weeks (n=54), and a reference group from the open population (n=50). **Design:** Main instrument was the TNO-AZL Pre-school Quality Of Life (TAPQOL) questionnaire, which was completed by the parents. Other outcome measures obtained from parents or neonatologists were investigated in addition. **Main results:** Children born < 32 weeks had significantly lower HRQoL than the reference group in the scales for lungs, stomach, eating disorders, motor functioning, communication and anxiety. The parental feelings towards the child were related to the child's HRQoL. We found differences between the neonatologist and the parent in perception of the child's situation, which can have clinical consequences. **Conclusion:** Neonatal intensive care after birth has HRQoL implications for all children, particularly in children born < 32 weeks of gestation. *Key words: health related quality of life, health status, preterm birth, pre-school children, infants, (very)-low-birth-weight infants.*

6.2 Introduction

In the 1960s neonatal intensive care was introduced, and although the survival rate of preterm born children has improved ever since, the prevalence of major disabilities and handicaps has remained similar.[1-3] Children born preterm are at risk for developmental disorders like cognitive and school performance problems,[2,4-9] emotional and behavioural problems,[5,7,10,11] neurological problems and disorders in motor ability and physical health.[4,5,7,12] This brings about the ethical question whether full neonatal intensive care should be given to all children born preterm. However, many questions involving the quality of life of these children still have to be answered. Most studies up till now have concentrated on isolated physical or psychosocial effects of preterm birth.

Cognitive and motor development appear to be disrupted in early childhood by a preterm birth.[6,12-18] Furthermore, emotional and behavioural problems in pre-school children are related to a preterm birth.[19,20] However, these studies provide a limited picture, focussing as they do, only on specific aspects of health. Health Related Quality of Life (HRQoL) is useful here, because it combines physical, psychological and social wellbeing in one outcome measure.[7,21-23] The 'health' component refers to quality of life as a result of a certain health. According to the definition of the World Health Organization (1948), health involves physical and psychological as well as social functioning.[24] The assessment by a person of his or her own health functioning is called health status (HS). However, a certain quantity of HS does not directly result in a certain quality of life. There are individual and culturally determined differences in adapting to HS problems, factors which are emphasised by several authors.[22,25-29] Therefore we defined HRQoL as HS weighted by the degree to which the person shows negative emotions to such problems.[22,25,26,30] HRQoL measures could assist in clinical decision-making and permit the assessment of the effect of early interventions.[19] Preferably, HRQoL is assessed by a person him or herself. However, pre-school children cannot be used as informants, because HRQoL is usually assessed through paper and pencil questionnaires. Alternatively, a proxy respondent can be used. The closer the relationship between patient and proxy, the higher the agreement is between them. Therefore, the parent is the most preferable informant about the child's HRQoL.[31]

The following hypotheses were formulated concerning the relationship between gestational age at birth and HRQoL in pre-school children. Firstly, it was hypothesised that children born < 32 weeks of gestation would have a lower HRQoL than a reference group from a open population. Secondly, it was hypothesised that children born between 32-37 weeks of gestation would have a higher HRQoL than children born at < 32 weeks of gestation, and a lower HRQoL than children in a reference group. To test the relative effect of other perinatal factors apart from gestational age, a group of term (≥ 37 w.) neonatal intensive care unit (NICU) survivors was included as a clinical control group. It was hypothesised in addition that term NICU survivors would have lower HRQoL than children in a reference group.

Since measuring HRQoL in pre-school children born preterm is relatively new, the relationship of HRQoL with other outcome measures was studied as well. Firstly, special attention was given to HS results compared to HRQoL results, since most previous research on HRQoL was done without the emotional evaluation by the

patient. Secondly, the judgement of the neonatologist, about the seriousness of the outcome and the suffering of the child, was compared with the HRQoL and HS as reported by the parent. Possible agreement or disagreement between parent and neonatologist could have implications for medical decision-making regarding the child. Thirdly, the feelings of the parent towards their child were compared with the HRQoL. Parents were proxy respondents for their children's HRQoL. Therefore we expected an association between feelings toward the child and the reported HRQoL.

6.3 Method

6.3.1 Subjects and data collection procedures

The study contained both a clinical sample and a reference sample of children aged 1-4 years. The clinical sample consisted of the parents of 193 children, consecutively consulting the outpatient neonatology clinic of the Leiden University Medical Center between July and December 1996. They were invited to complete the TNO-AZL Pre-school Quality Of Life (TAPQOL) questionnaire, a pre-school child's HRQoL questionnaire using parents as proxy respondents.[32] Parents were approached during their visit to the clinic or by mail. Response was 86%. In addition, the parents were asked for permission to study their children's medical case history. One-hundred-and-sixty-four parents granted permission. Four of these children were excluded as information on gestational age was lacking. Thus, 160 children were included in the clinical sample. Three clinical groups with a NICU history were constructed: 65 children born after less than 32 completed weeks of gestation, 41 children born between 32 and 37 weeks and 54 children who were term NICU survivors ($\geq$ 37 weeks) with neonatal problems other than a preterm birth (like asphyxia).

A random sample of 50 pre-school children, drawn from six "Well-Baby-Clinics" distributed all over the Netherlands, was used as a reference group alongside the three clinical groups. Criteria about informed consent and anonymity were met according to the TNO Medical Ethics Committee.[32] Demographic data such as the pre-school child's sex, age, number of siblings, maternal and paternal educational level, occupation of the mother, religious denomination and parents' age were not different between the four groups.

6.3.2 Measures

TAPQOL (TNO-AZL Pre-school Quality Of Life)

This is a generic instrument for assessing HRQoL of pre-school children, aged 1 to 5 year, in medical research and clinical trials. It is developed by the TNO Institute of Prevention and Health and the department of paediatrics of the Leiden University Hospital Center (AZL). These institutes co-operated in research among preterm born children since 1983, and in research about HRQoL since 1990. The instrument is constructed as a sequel to the TACQOL, a questionnaire to measure HRQoL of children 6 to 12 years of age.[30,31,33] The TAPQOL has 10 scales which are relevant for children of 1 to 1.5 years old: stomach, skin, lungs, sleeping, appetite, eating disorders, liveliness, positive mood, problem behaviour, anxiety. For children of 1.5 to 5 years old, the TAPQOL includes 3 more: motor functioning, social functioning and communication. Twenty-five percent of the sample was less than 1.5 years of age. Parents were instructed to compare the quality of their child's life with that of peers of the same age. A concretely and specifically formulated Health Status problem, if reported, leads to a question about the emotional response. Examples of items are given in Table 1. Items were scored 0-4. Scale scores were obtained by adding item scores within scales, and transforming crude scale scores linearly to a 0-100 scale, with higher scores indicating better HRQoL. Emotional responses were not assessed in items of the liveliness, positive mood, problem behaviour, anxiety and social functioning scales, because this would have led to illogical items. In this situation, item scores were on a 0-2 scale (0 = often, 1 = sometimes, 2 = never) and scale scores ranged from 0-100. The item scores of the other scales were also encoded excluding the emotional evaluation to obtain HS scores (0 = often, 1 = sometimes, 2 = never) and scale scores ranged from 0-100, with higher scores representing better HS.

In our sample Cronbach's alpha coefficient range of HRQoL was .61 to .88. Stomach, eating disorders and social functioning had alpha's below .70. The range of HS scales was .59 to .89. Stomach, skin functioning, eating disorder and social functioning had alphas below .70. The construction and psychometric performance of the instrument has been described in more detail elsewhere.[32] The instrument has good construct, criterion and concurrent validity.

Pre-school child's and parent's characteristics

In addition, questions were asked which assessed the pre-school child and parent's characteristics. This included general information such as parental education and parents' age.

Table 1. Examples of scale-items of the TAPQOL

Scale	Item example (In the last three months...)
Sleeping	Did your child sleep restlessly?
Appetite	Was your child's appetite poor?
Lungs	Has your child been short of breath?
Stomach	Has your child had stomach-ache or abdominal pain?
Skin	Has your child had dry skin?
Eating disorders	Did your child have difficulty swallowing food?
Motor functioning	Did your child have difficulty with walking?
Social functioning	Was your child at ease with other children?
Problem behaviour	Was your child aggressive?
Communication	Did your child have difficulty in making it clear what he/she wanted?
Anxiety	Was your child anxious?
Positive mood	Was your child in good spirits?
Liveliness	Was your child energetic?

Response categories: 'never' / 'sometimes' / 'often'; If the answer is 'sometimes' or 'often', the next question in the stomach, skin, lungs, sleeping, appetite, eating disorders, motor functioning, and communication scales is: 'At that time, my child felt': 'fine' / 'not so good' / 'quite bad' / 'bad'

Medical Case History Checklist of pre-school children with a NICU history

This is a checklist for systematic retrospective evaluation of the clinical sample's case histories. It was specially developed for this study by two neonatologists (S.V, S.M.) and a developmental psychologist (N.C.M.T). The list assesses NICU diagnoses such as birth factors, pulmonary disorders and circulation disorders and treatments that were considered clinically relevant.

Judgement of the neonatologist: Global HRQoL of the child

The judgement of the neonatologist was used as a confirmatory outcome measure next to the TAPQOL administering. A multidimensional HRQoL questionnaire like the TAPQOL was too time consuming for the neonatologists. Therefore we used a short global measure instead. It contained three questions: present medical complications as a result of NICU period (answers: yes or no), seriousness of outcome ("not applicable" to "very serious" using a 0 to 7 point Likert scale), and suffering of the child ("not applicable" to "very serious" using a 0 to 7 point Likert scale).

CAFEIN ("Caretaker's Feelings towards their Infant with NICU history")

This instrument assesses the feelings of a parent towards their pre-school child with a neonatal intensive care history. It was developed for this study by a developmental psychologist (N.C.M. T.). The content was based on qualitative parent

reports from the follow-up program "Project On Preterm and Small for gestational age infants (POPS)" [1,8,9,34] and on literature about parent-child relations in prematurity.[19,35-38] It contains three scales: Attachment and upbringing (6 items, e.g.: "Did you have difficulties to feel attached to the child", Cronbach's alpha=.64); Positive emotions in relation with the child (6 items, e.g.: "Were you happy in relation to the child?", Cronbach's alpha=.72); and Negative emotions in relation with the child (8 items, e.g.: "Were you worried in relation to the child?", Cronbach's alpha=.80). The questions follows the format of the TACQOL[30,31,33] and TAPQOL[32], with scale scores ranging from 0-100. Construct validity of the questionnaire was good: Item ordinality was confirmed by homogeneity analysis (HOMALS) and the structure of the questionnaire was confirmed by factor analyses.

6.3.3 Data analysis

The HRQoL results of the three clinical groups and the reference group were summarised with the mean scores and the standard errors. Differences between groups with respect to HRQoL were tested by MANOVA. This was done with and without confounding variables (birth weight for gestational age standard deviation score (SDS), pre-school child's sex and age, maternal and paternal educational level) to single out the effect of gestational age.

The relationship of several perinatal factors with HRQoL was investigated in two steps. In step one the relation between the HRQoL scales and the perinatal factors was analysed with canonical correlation analysis. The purpose of this procedure is to determine how similar sets of variables are to one another. Canonical correlation analysis is an extension of multiple regression (MANOVA), with more than one dependent variable. Global tests for no correlation were performed at the nominal level of .05 using the first canonical correlation. Age of the child during HRQoL assessment was added as a covariant, because developmental differences in pre-school children can be large. Age was not corrected for weeks of gestation.[39,40] The motor functioning, social functioning and communication scales were excluded from these canonical correlation analyses, because they were relevant only for children of 1.5 to 5 years.

In step two, multiple regression analyses (MANOVA) of each HRQoL scale separately on clusters of perinatal factors was performed, with age of the pre-school child as covariant.

Table 2. Distribution of clinical variables amongst the three clinical groups (according to information in case histories)

		Very premature < 32 w. N (%)	Premature 32 - 37 w. N (%)	Term NICU survivors ≥37 w. N (%)
Birth factors				
Birthweight (in grams)	≤ 1330	51 (80)	3 (7)	0 (0)
	1330-2500	13 (20)	28 (68)	14 (26)
	> 2500	0 (0)	10 (24)	40 (74) *
Standard deviation score (SDS) for birth weight at given gestational age	< -2 SD	8 (13)	7 (17)	14 (26)
Infant's sex	male	37 (57)	25 (61)	33 (61)
Multiple pregnancy	yes	22 (34)	7 (17)	3 (6) *
Complications during pregnancy	yes	46 (77)	34 (85)	27 (51) *
Apgar score at 5 min.	< 7	14 (22)	6 (15)	11 (22)
Pulmonary disorders and treatments				
Idiopathic respiratory distress syndrome (IRDS)	yes	38 (59)	8 (20)	3 (6) *
Oxygen administration	yes	37 (60)	12 (29)	14 (26) *
Continuous Positive Airway Pressure (CPAP)	yes	53 (84)	10 (24)	4 (8) *
Intermittent Positive Pressure Ventilation (IPPV)	yes	53 (83)	10 (24)	20 (38) *
Pneumothorax	yes	10 (15)	3 (7)	3 (6)
Bronchopulmonary dysplasia (BPD)[41]	yes	12 (19)	1 (2)	0 (0) *
Circulation disorders				
Persistent ductus arteriosus (PDA)[a]	yes	17 (26)	3 (7)	2 (4) *
Hypotension[b]	yes	11 (17)	0 (0)	2 (4) *
Other disorders				
Hyperbilirubinemia	yes	35 (54)	11 (27)	13 (24) *
Intraventricular haemorrhage (IVH)[42]	yes	12 (18)	3 (7)	0 (0) *
Number of diagnoses mentioned in event history (out of a list of 26 possible diagnoses including 'other')	≤ 3 diagnoses	14 (22)	22 (55)	40 (76)
	3-5 diagnoses	22 (34)	13 (33)	10 (19)
	> 5 diagnoses	28 (44)	5 (13)	3 (6) *
Social economic status (SES)				
Maternal educational level	low	17 (29)	13 (34)	18 (35)
	medium	25 (42)	14 (37)	25 (48)
	high	17 (29)	11 (29)	9 (17)
Paternal educational level	low	20 (36)	14 (37)	23 (46)
	medium	13 (23)	11 (29)	16 (32)
	high	23 (41)	13 (34)	11 (22)
Job occupation of mother	no	37 (57)	20 (50)	28 (53)

* Chi-Square (Pearson) p≤.05; [a] diagnosed by ultrasound; [b] mean arterial pressure (30 mmHG)

Furthermore, canonical correlation analyses were used to associate the judgement of the neonatologist and the feelings of the parents (CAFEIN) with HRQoL scales. All of the afore-mentioned computations with HRQoL were done for HS as well. Overall, a p-value of .05 or less was assumed to represent a significant result. Statistical analyses were performed using SPSS 6.1 (SPSS Inc., Chicago U.S.A.).

6.4 Results

Table 2 at the previous page presents the definitions of the perinatal variables and social economical status (SES) and their distribution amongst the three clinical groups. As can be seen most perinatal problems occurred in the very preterm group (Chi-square $p \leq .05$).

6.4.1 Prematurity versus HRQoL

Figure 1 at page 68 presents the means of the HRQoL scales of the four groups. In Table 3, the differences between groups with and without confounding variables are shown. As an example the results of the lungs scale are discussed here.

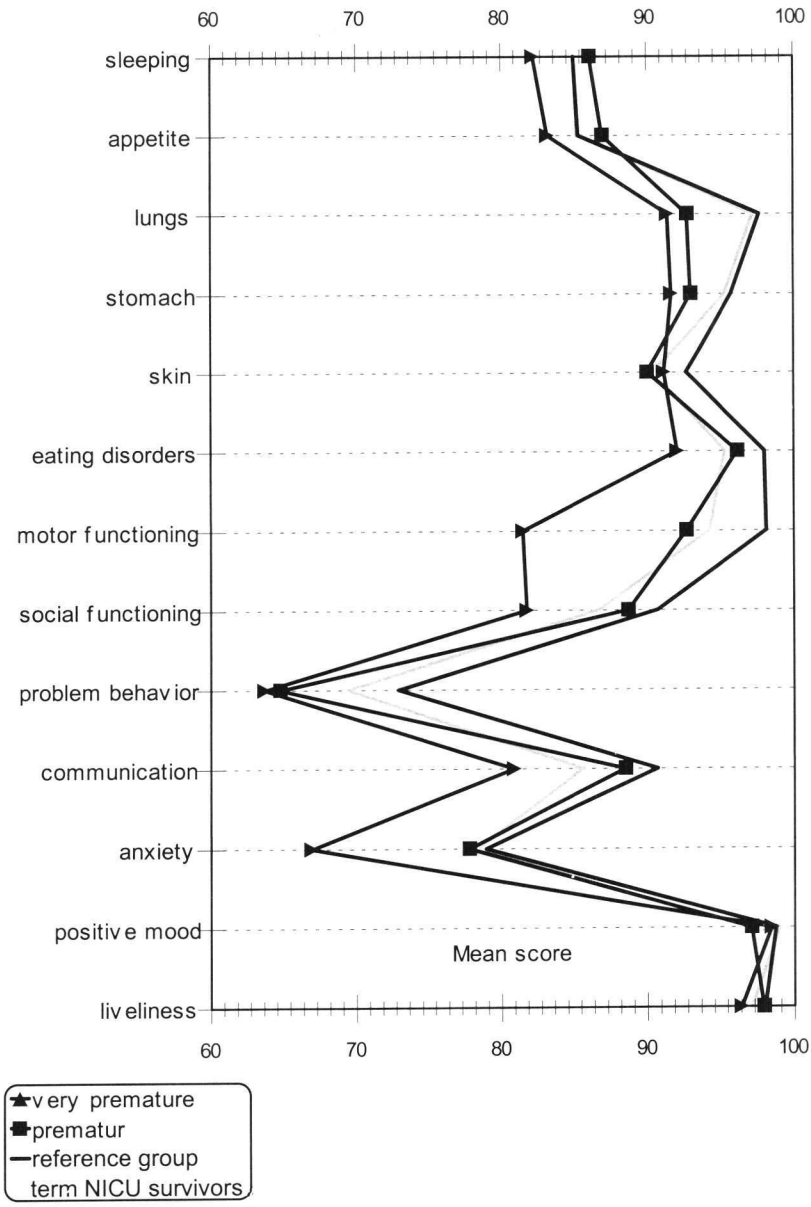
In the lungs scale, children born very preterm (< 32 w.) on average had a HRQoL score that was 10.2 points lower than the HRQoL score of the reference group, with a standard error of 3.5. This difference was significant ($p \leq .05$). When corrected for the confounding variables the difference between the very preterm and the reference group increased a little: 11.2 points (SE=4.5, p=.009). Furthermore, children born very preterm had a HRQoL lungs score that was on average 2.9 points lower than the HRQoL lungs score of the preterm group (32-37 w.). However, this difference was not significant (SE=3.7, p=.44). When corrected for the confounding variables the difference between the very preterm and the preterm group was larger (5.1 points), but remained statistically non-significant (SE=4.2, p=.23). Lung scale differences between pre-school children born between 32-37 w. and the reference group or between the term NICU survivors and the reference group were not significant, with or without confounding variables. Overall, pre-school children born at < 32 w. of gestation had significantly lower scores than the reference group in the scales for lungs, stomach, eating disorders, motor functioning, communication and anxiety. When corrected for confounding variables the difference on the stomach scale disappeared. Therefore, hypothesis 1 was partly supported.

Table 3. Differences in mean HRQoL scores[a] between very premature, premature and reference group and between term NICU survivors and reference group

	Very premature - Reference		Very premature - Premature		Premature - Reference		Term NICU survivors - Reference	
	raw data	adjusted[b]	raw data	adjusted	raw data	adjusted	raw data	adjusted
Sleeping	-5.1 (3.7)	-4.2 (4.5)	-6.3 (4.0)	-6.8 (4.5)	1.5 (4.2)	2.6 (4.9)	0.0 (3.9)	5.5 (4.8)
Appetite	-2.9 (2.9)	-1.5 (3.4)	-3.3 (3.1)	-3.4 (3.5)	0.4 (3.3)	1.9 (3.8)	-0.4 (2.5)	1.1 (3.0)
Lungs	-10.2 (3.5)*	-11.2 (4.2)*	-2.9 (3.7)	-5.1 (4.2)	-7.3 (3.9)	-6.1 (4.6)	-0.7 (2.8)	-1.5 (3.5)
Stomach	-7.1 (3.5)*	-6.3 (4.3)	-2.4 (3.8)	-3.3 (4.3)	-4.7 (4.0)	-3.0 (4.7)	-0.6 (3.2)	0.1 (4.1)
Skin	-1.5 (2.6)	-2.9 (3.2)	-2.0 (2.8)	2.7 (3.2)	-3.5 (2.9)	-5.6 (3.4)	-3.3 (3.0)	-4.1 (3.8)
Eating disorders	-6.8 (2.3)*	-5.8 (2.6)*	-3.3 (2.4)	2.3 (2.6)	-3.5 (2.6)	-3.5 (2.8)	-1.9 (1.2)	-1.8 (1.5)
Motor functioning	-8.5 (2.1)*	-10.5 (2.7)*	-4.5 (2.4)	-6.1 (2.8)*	-4.0 (2.5)	-4.4 (3.2)	-3.7 (1.6)*	-4.0 (2.3)
Social functioning	-3.0 (3.9)	-5.6 (4.5)	-8.0 (4.3)	-6.4 (4.6)	5.0 (4.5)	0.8 (5.1)	2.5 (3.7)	0.1 (5.0)
Problem behaviour	-5.7 (3.7)	-6.1 (4.3)	-1.1 (4.0)	0.0 (4.3)	-4.7 (4.2)	-6.1 (4.7)	3.5 (2.8)	2.1 (3.2)
Communication	-10.2 (2.9)*	-11.1 (3.1)*	-7.6 (3.3)*	-8.9 (3.3)*	-2.5 (3.5)	-2.3 (3.6)	-7.5 (3.6)*	-12.5 (4.8)*
Anxiety	-11.4 (3.9)*	-14.4 (4.6)*	-10.9 (4.2)*	-10.1 (4.6)*	-0.6 (4.4)	-4.3 (5.0)	0.6 (3.7)	-1.7 (4.5)
Positive mood	-0.5 (1.5)	-0.1 (1.9)	-1.4 (1.6)	1.7 (1.9)	-1.9 (1.7)	-1.7 (2.0)	-0.3 (1.2)	-0.2 (1.5)
Liveliness	-1.6 (1.9)	-2.1 (2.1)	-1.5 (2.0)	-1.1 (2.1)	-0.1 (2.2)	-1.0 (2.3)	-0.9 (1.8)	-1.7 (2.1)

Note: The lungs scale is discussed in the Results section as an example; [a] Differences in mean (SE) based on univariate t-value and .95 confidence intervals; [b] Adjusted for birth weight SDS, infant's sex and age, maternal and paternal educational level; * $p \leq .05$

Figure 1. HRQoL group means of the three clinical groups and the reference group



Hypothesis 2 was supported in so far as a significant difference was found between the very preterm and preterm group for motor functioning (adjusted only), communication and anxiety. Hypothesis 3, that the term NICU survivors would have lower HRQoL than the reference group, was supported for motor functioning and communication.

6.4.2 Relationship between perinatal factors and HRQoL

The first canonical correlation and multiple correlations between the HRQoL scales on the one side and perinatal factors on the other side are presented in Table 4. As can be seen, a significant canonical correlation of the HRQoL scales was found with birth factors, pulmonary disorders and treatments, circulation disorders and other disorders, but not with SES. Birth factors correlated with problem behaviour, communication, anxiety and liveliness. The combination of pulmonary disorders and treatments correlated with the appetite, lungs, eating disorders, motor functioning, communication and anxiety scales. The circulation disorders correlated with the appetite, skin, social functioning, problem behaviour, communication and anxiety scales. The other disorders correlated with the appetite, motor functioning, social functioning, problem behaviour, communication and anxiety scales. Although the canonical correlation of the SES variables with all HRQoL scales was not significant, SES variables correlated with the appetite, eating disorders and problem behaviour scales.

Table 4. Multiple correlations between perinatal factors and HrQOL

	All perinatal factors	Birth factors	Pulmonary disorders and treatments	Circulation disorders	Other disorders	SES
All HRQoL scales [a]	.65 *	.55 *	.53 *	.46 *	.47 *	.41
Sleeping	.45	.21	.27	.05	.14	.09
Appetite	.42	.27	.32 *	.28 *	.30 *	.25 *
Lungs	.50	.21	.35 *	.13	.11	.08
Stomach	.45	.24	.29	.10	.17	.11
Skin	.45	.27	.23	.24 *	.15	.15
Eating disorders	.53 *	.31	.30 *	.12	.18	.23 *
Motor functioning	.58	.27	.38 *	.15	.40 *	.12
Social functioning	.47	.34	.30	.32 *	.35 *	.09
Problem behaviour	.51	.34 *	.24	.27 *	.29 *	.25 *
Communication	.55	.44 *	.36 *	.29 *	.38 *	.24
Anxiety	.46	.34 *	.36 *	.27 *	.27 *	.20
Positive moods	.34	.17	.18	.17	.24	.13
Liveliness	.48	.33 *	.28	.18	.11	.08

* significant F: $p \leq .05$; [a] excluding scales only relevant for children > 1.5 years of age (motor functioning, social functioning and communication)

6.4.3 Relation between outcome measures

HRQoL versus HS

The most salient difference between HRQoL and HS scores concerned motor functioning. The adjusted mean HRQoL of the very preterm was 10.5 points lower then the mean HRQoL of the reference group, whereas the HS of the very preterm was 18.7 points lower than the HS of the reference group. For the difference between the very preterm and the preterm groups the same kind of effects were found (–6.1 HRQoL and –9.7 HS respectively). Therefore, as far as the motor functioning of the very preterm is concerned, HS was estimated to be much lower than HRQoL. In contrast, on lungs, stomach and sleeping the reverse is true: The HS of the very preterm was estimated higher than the HRQoL.

The correlations between perinatal factors and HRQoL motor functioning were lower and less often significant than between perinatal factors and HS motor functioning. For instance, the correlation of HRQoL with birth factors was estimated at .27 (p=.532), the correlation of HS with birth factors was estimated .61 (p=.001). Therefore, HRQoL motor functioning related less to the perinatal factors than HS motor functioning. In contrast, however, HRQoL appetite and communication related more to the perinatal factors than HS appetite and communication.

HRQoL versus judgement of the neonatologist

The judgement of the neonatologist, concerned the medical outcome of NICU history, the seriousness of outcome and the suffering of the pre-school child. According to the first canonical correlation, the judgement of the neonatologist correlated with the HS estimated by the parents (.53 ; p=.037), but not with the HRQoL (.42 ; p=.182). Therefore, neonatologists did not include the emotional evaluation of HS problems in their judgement, which the parents did.

HRQoL versus feelings of the parents

According to the first canonical correlation, the CAFEIN, assessing the feelings of a parent towards its child, correlated with both HRQoL (.65 ; p=.000) and HS (.62 ; p=.000) of the child. The expected association between feelings of the parent and the reported HRQoL was found.

6.5 Discussion

This paper describes the relation between HRQoL and perinatal factors accompanying preterm birth. We expected a relation between gestational age and HRQoL. The data of the present study confirm this hypothesis, indicating that pre-school children born at less then 32 weeks of gestation had lower HRQoL than the reference group. We found significant differences on scales with a physical character (lungs, stomach, eating disorders, motor functioning) but also on psychosocial scales (communication and anxiety). The HRQoL of the pre-school children born 32-37 w. and the HRQoL of term NICU survivors resembled the HRQoL of the reference group in most aspects. However, the order of the groups – although not always significant – is often as expected. This means lowest HRQoL in the very preterm (< 32 w.), medium HRQoL in the preterm (32-37 w.) and highest HRQoL in the reference group. Furthermore, the HRQoL of the term NICU survivors is worse than the HRQoL of the reference group. Apparently, the need for neonatal intensive care after birth has implications for HRQoL in early childhood, but to an even greater degree in children born very preterm.

In other studies, relations between various perinatal factors and pre-school child development were found.[4,15,16,43] Comparison with these studies is difficult because of the lack of consensus on predictors, outcome measures (no HRQoL), and methodologies. In our study, association was found between perinatal factors and HRQoL. SES was less associated with HRQoL than birth factors, pulmonary disorders and treatments, circulation disorders and other disorders. According to other studies, development in pre-school children of lower SES families appeared to be worse,[4,6,16,18] but the parents' perception of their pre-school child's HRQoL is less influenced by their SES.

This study is the first attempt to measure the HRQoL of pre-school children born preterm. We defined HRQoL as health status weighted by the emotional response to problems in health status. HRQoL motor functioning of the very preterm children was higher than would have been estimated without emotional evaluation (that is HS). In contrast, HRQoL lungs, stomach and sleeping were lower than would have been estimated by HS alone. In this study, the judgement of the neonatologist concerning the medical outcome of NICU history, seriousness of outcome and suffering of the pre-school child, correlated with the HS but not with the HRQoL. As a result of the above, one could image a clinical situation: The neonatologist is surprised that a parent does not

want to have full treatment for the child's motor functioning problems. The parent simply does not consider the motor problems to be as serious as the neonatologist does. On the other hand the parent evaluated lungs, stomach and sleeping problems as being more of an emotional burden than the neonatologist does. The parent therefore does not understand why these problems in their pre-school child receive less attention from the neonatologist.

As expected we found that the feelings of a parent towards their child related to parental perceptions of their child's HRQoL. Thereby, the data may reveal as much about the parents as about the children. According to previous research, mothers who had higher depression scores themselves reported more behaviour problems in their pre-school children.[19] However, mothers of preterm children believed more in the power of the environment to produce positive outcomes. This could result in them giving their child more stimulation and structure than other children receive from their mothers,[13,14,44] which in turn would result in less reported problems. In other studies the idea of a 'prematurity stereotype' was not supported.[14] Interactions between mother and pre-school child were not influenced by preterm birth.[45] The mother's view about cognition and temperament of preterm children were similar to the views of mothers of full-term children.[13,14] Therefore, the HRQoL of the pre-school child could be both the result or the cause of the parent's feelings towards the pre-school child. However, parents are the main decision-makers in respect of the rearing and medical treatment of their children. This makes their perception of the pre-school child's HRQoL at least clinically relevant.

A limitation of the study is the use of a relatively small hospital-based sample of NICU children. Children who had no complaints after neonatal intensive care, children who were being treated in rehabilitation centres because of severe handicaps, and children who were followed up in other hospitals were not included in the sample. This may have caused both underestimation and overestimation of HRQoL respectively, probably counterbalancing each other.

In conclusion, the study revealed that at 1 to 4 years after birth, the HRQoL of children born after less than 32 weeks of gestation was lower than the HRQoL of children born with more than 32 weeks of gestation. This concerned differences in the lungs, stomach, eating disorders, motor functioning, communication and anxiety scales. There appeared to be a relation between perinatal factors assessed on the basis of case histories and HRQoL in pre-school children as reported by the parents. SES was less associated with HRQoL than were birth factors, pulmonary, circulation and other

disorders. An proportional association is found between parent's feelings toward the child and the HRQoL of the child. Furthermore, we found differences between the neonatologist's and the parent's perception of the child's situation, which can have clinical consequences. Needing neonatal intensive care after birth has HRQoL implications for all children, and, to an increasing degree, for children born very preterm. To determine the stability of HRQoL in pre-school children born preterm, longitudinal research will be needed, preferable with a larger, regional based sample.

Acknowledgements
We would like to express our deepest appreciation to M. de Haan MD (2), S. Martens MD (1) and M. Rijken MD (1) for sharing their neonatal knowledge and taking part in data collection. We also like to thank Prof. Dr. S. Pauline Verloove-Vanhorick (1) (2) for her critical review of the manuscript. We are indebted to the parents for their participation.

6.6 Reference List

1. Veen S, Ens-Dokkum MH, Schreuder AM, Verloove-Vanhorick SP, Brand R, Ruys JH. Impairments, disabilities, and handicaps of very preterm and very-low-birthweight infants at five years of age. *Lancet* 1991; **338:** 33-36.
2. Robertson CM, Etches PC, Kyle JM. Eight-year school performance and growth of preterm, small for gestational age infants: a comparative study with subjects matched for birth weight or for gestational age. *J Pediatr* 1990; **116:** 19-26.
3. Tin W, Wariyar U, Hey E. Changing prognosis for babies of less than 28 weeks' gestation in the north of England between 1983 and 1994. Northern Neonatal Network. *Brit Med J* 1997; **314:** 107-111.
4. Stanton WR, McGee R, Silva PA. Indices of perinatal complications, family background, child rearing, and health as predictors of early cognitive and motor development. *Pediatr* 1991; **88:** 954-959.
5. Middle C, Johnson A, Alderdice F, Petty T. Birthweight and health and development at the age of 7 years. *Child Care Health Dev* 1996; **22:** 55-71.
6. Halsey CL, Collin MF, Anderson CL. Extremely low birth weight children and their peers: a comparison of pre-school performance. *Pediatr* 1993; **91:** 807-811.
7. Saigal S, Rosenbaum P, Stoskopf B, et al. Comprehensive assessment of the health status of extremely low birth weight children at eight years of age: Comparison with a reference group. *J Pediatr* 1994; **125:** 411-417.
8. den-Ouden AL, Hille ETM, Bauer L, Verloove-Vanhorick SP. School Performance in very preterm children: a report from the collaborative project on Preterm and Small for gestational age (POPS) infants in the Netherlands. *Lancet* 1993; **342:** 550-551.
9. Hille ETM, den-Ouden L, Bauer L, Van-den-Oudenrijn C, Brand R, Verloove-Vanhorick SP. School performance at nine years of age in very premature and very low birth weight infants: Perinatal risk factors and predictors at five years of age. *J Pediatr* 1994; **125:** 426-434.
10. Levy-Shiff R, Einat G, Har Even D, et al. Emotional and behavioural adjustment in children born prematurely. *J Clin Child Psychol* 1994; **23:** 323-333.
11. Ross G, Lipper EG, Auld PA. Social competence and behaviour problems in premature children at school age. *Pediatr* 1990; **86:** 391-397.

12. Wolke D, Ratschinski G, Ohrt B, Riegel K. The cognitive outcome of very preterm infants may be poorer than often reported: an empirical investigation of how methodological issues make a big difference. *Eur J Pediatr* 1994; **153:** 906-915.
13. Pearl R, Donahue M. Brief report: Four years after a preterm birth: Children's development and their mothers' beliefs and expectations. *J Pediatr Psychol* 1995; **20:** 363-370.
14. Donahue ML, Pearl R. Conversational interactions of mothers and their pre-school children who had been born preterm. *Journal of Speech and Hearing Research* 1995; **38:** 1117-1125.
15. Siegel LS. Reproductive, perinatal, and environmental variables as predictors of development of preterm (less than 1501 grams) and fullterm children at 5 years. *Semin Perinatol* 1982; **6:** 274-279.
16. Thompson RJ, Gustafson KE, Oehler JM, Catlett AT. Developmental outcome of very low birth weight infants at four years of age as a function of biological risk and psychosocial risk. *J Dev Behav Pediatr* 1997; **18:** 91-96.
17. Meisels SJ, Plunkett JW, Roloff DW, Pasick PL, Stiefel GS. Growth and development of preterm infants with respiratory distress syndrome and bronchopulmonary dysplasia. *Pediatr* 1986; **77:** 345-352.
18. Cohen SE, Parmelee AH. Prediction of five-year Stanford-Binet scores in preterm infants. *Child Dev* 1983; **54:** 1242-1253.
19. Spiker D, Kraemer HC, Constantine NA, Bryant D. Reliability and validity of behaviour problem checklists as measures of stable traits in low birth weight, premature preschoolers. *Child Dev* 1992; **63:** 1481-1496.
20. Grunau RVE, Whitfield MF, Petrie JH. Pain sensitivity and temperament in extremely low-birth-weight premature toddlers and preterm and full-term controls. *Pain* 1994; **58:** 341-346.
21. Saigal S, Feeny D, Furlong W, Rosenbaum P, Burrows E, Torrance G. Comparison of the health-related quality of life of extremely low birth weight children and a reference group of children at age eight years. *J Pediatr* 1994; **125:** 418-425.
22. Ebrahim S. Clinical and public health perspectives and applications of health-related quality of life measurement. *Soc Sci Med* 1995; **41:** 1383-1394.
23. Sweeny AJ, Creer TL. Health-related-quality-of-life assessment in medical care. *Dis Month* 1995; **16:** 5-64.
24. World Health Organization. *Constitution of the World Health Organization.* Geneva: World Health Organization, 1948.
25. Lindström B. Measuring and improving quality of life for children. In: Lindström B, Spencer N, eds. *Social Paediatrics.* Oxford: Oxford University Press, 1995: 571-585.
26. Rosenberg R. Health-related quality of life between naturalism and hermeneutics. *Soc Sci Med* 1995; **41:** 1411-1415.
27. Bullinger M, Ravens-Sieberer U. Stand der Forschung zur gesundheitsbezogenen Lebensqualität von Kindern: eine Literaturanalyse [German: State of the art of health-related quality of life research in children: a literature review]. *Präv Rehab* 1995; **7:** 106-121.
28. Amir M, Bar-On DM, Penso R. Positive-Negative Evaluation (PNE) scale: A new dimension of the subjective domains of quality of life measure. *Qual Life Res* 1996; **5:** 73-80.
29. Sprangers MAG, Rozemuller N, van-den-Berk MBP, Boven SV, van-Dam FSAM. Response shift bias in longitudinal quality of life research. *Qual Life Res* 1994; **3:** 49-49.
30. Vogels T, Theunissen NCM, Verrips GH, Koopman HM, Verloove-Vanhorick SP, Kamphuis RP. Het meten van kwaliteit van leven bij kinderen met chronische aandoeningen [Dutch: Assessing health related quality of life in chronically ill children]. *TIAZ* 1996; **3:** 104-111.
31. Theunissen NCM, Vogels AGC, Koopman HM, et al. The proxy problem: child report versus parent report in health-related quality of life research. *Qual Life Res* 1998; **7:** 387-397.
32. Fekkes M, Theunissen NCM, Brugman E, et al. The TAPQOL. A health related quality of life instrument for 1 to 5-year old children. *submitted for publication* 1998.
33. Verrips GH, Vogels AGC, Verloove-Vanhorick SP, et al. Health-related quality of life measure for children -- the TACQOL. *J Appl Therapeutics* 1997; **1:** 357-360.

34. Verloove-Vanhorick SP, Verwey RA, Ebeling MCA, Brand R, Ruys JH. Mortality in very preterm and very low birth weight infants according to place of birth and level of care: Results of a national collaborative survey of preterm and very low birth weight Infants in the Netherlands. *Pediatr* 1988; **81:** 404-411.
35. Schols M. *Betekenissen van een premature geboorte: Effecten van interventies in de relatie tussen moeder en prematuur kind in de eerste twee levensjaren [Dutch: Meanings of a premature birth: Effects of interventions in the relationship between mother and premature child in the first two years of life].* Enschede: Febo druk bv, 1993: 1-230.
36. van-der-Heyden Z. *Vanuit de couveuse de wereld in: vragen van ouders van een couveusekind beantwoord [Dutch: Out of the incubator, into the world: questions of parents with an incubator-child answered].* Ede: Zomer & Keuning Boeken B.V., 1992: 5-112.
37. Bruinse HW, Cats BP. *Te vroeg op de wereld: Over prematuren in Nederland [Dutch: Too early in this world: about prematures in the Netherlands].* Utrecht: Lemma BV, 1994: 7-126.
38. De-Leeuw R, Bakker L. *Te vroeg geboren [Dutch: Born too early].* Ede: Zomer en Keuning Boeken B.V., 1989: 5-158.
39. Miller G, Dubowitz LM, Palmer P. Follow-up of pre-term infants: Is correction of the developmental quotient for prematurity helpful? *Early Hum Dev* 1984; **9:** 137-144.
40. den-Ouden AL, Rijken M, Brand R, Verloove-Vanhorick SP, Ruys JH. Is it correct to correct? Developmental milestones in 555 "normal" preterm infants compared with term infants. *J Pediatr* 1991; **118:** 399-404.
41. Bancalari E, Abdenour GE, Feller R, Gannon J. Bronchopulmonary dysplasia: clinical presentation. *J Pediatr* 1979; **95:** 819-823.
42. Papile LA, Burstein J, Burstein R, Koffler H. Incidence and evolution of subependymal and intraventricular hemorrhage: a study of infants with birth weights less than 1,500 gm. *J Pediatr* 1978; **92:** 529-534.
43. Goldstein KM, Caputo DV, Taub HB. The effects of prenatal and perinatal complications on development at one year of age. *Child Dev* 1976; **47:** 613-621.
44. Becker PT, Engelhardt KF, Steinmann MF, Kane J. Infant age, context, and family system influences on the interactive behaviour of mothers of infants with mental delay. *Research in Nursing and Health* 1997; **20:** 39-50.
45. Schermann Eizirik L, Hagekull B, Bohlin G, Persson K, Sedin G. Interaction between mothers and infants born at risk during the first six months of corrected age. *Acta Paediatr* 1997; **86:** 864-872.

7 Quality of life in children in a longitudinal perspective: an exploratory review

N.C.M. Theunissen, A.A. Kaptein, H.M. Koopman, T.G.C. Vogels, J.M. Wit

7.1 Abstract

Objective: This paper explores the time variability of quality of life (QoL) in children between 0 to 12 years of age. Design: A systematic review of original studies, with at least two QoL assessments, and published between 1966 and 1998. The publications were identified from medical and psychological sources by computerised searches followed by manual selection. Data synthesis: The 32 selected publications were discussed according to their general characteristics, QoL assessment, longitudinal QoL research design and approaches to what changes QoL. Main results: Only two publications met all QoL assessment requirements (multi-factorial, self-administered, subjective) as well as longitudinal requirements (clear description of assessment period, recall period, sample size at end of study, longitudinal statistics). The approach to change that underlie the 32 publications can be described as: stable physical health gives stable QoL and changes in physical health change QoL. It is rarely acknowledged that psychological, social and situational variables can change QoL as well. Conclusions: More studies are needed that meet QoL assessment requirements as well as longitudinal requirements. Discussion is necessary about what exactly changes QoL, as this influences the planning of the assessments and guides the interpretation of changes. *Key words: quality of life, health-related quality of life, children, paediatric, literature review, longitudinal studies, follow-up.*

7.2 Introduction

As medical successes in keeping children with serious diseases alive increase, the children's quality of life (QoL) receives growing attention. It is recognised that a certain disease or side effects of treatments can elicit quite different reactions in different children.[1] QoL accounts for these individual differences. As a result, QoL is increasingly used as one of the indicators of whether or not medical treatment is successful.[1,2] Although a widespread definition or theoretical framework is missing, there is a growing consensus on four aspects of QoL: it is multi-factorial (physical, psychological and social

well-being), it is patient self-administered, it is subjective, and its value is variable over time.[1,3] These aspects of QoL are considered useful paradigms in adults' as well as children's QoL. It is generally assumed that in childhood development is more rapid and important than at other stages in life. Therefore, time variability is considered particularly relevant in children. In longitudinal research the time variability of QoL is the explicit objective of study. Therefore, the purpose of the present review is to explore children's QoL in a longitudinal perspective.

7.3 Approaching changes in children's QoL

The goal of a longitudinal study reveals implicit ideas of the investigators about the changeability of QoL. Two main approaches to change can be distinguished in advance: The first approach is expressed in studies that are conducted for the following reasons: (a) To describe QoL in a particular group, for instance, to describe the impact of a disease on daily life or on the condition of a patient[4-9]; (b) To describe developmental processes; for instance, to describe age trends in a specific healthy sample or group of patients with a particular illness,[9] or to assess patterns of QoL over time[7]; (c) to identify physical or psychological determinants of QoL[4,9] and predict future QoL from it[5,10]; (d) to predict morbidity and mortality using QoL as baseline data[4,6,7,9]; or (e) to evaluate the test-retest reliability or reproducability of new QoL instruments. Although QoL is said to be variable over time,[1,3] these five research aims all depend on a certain amount of invariability in time, without which QoL cannot be predicted. We call this the *predictability* approach to change, which is defined as the maintenance of a relative position on particular characteristics over time.[10] It can denote both stability (absolute levels of a characteristic remain stable over time), as well as continuity (consistency in relative rank over time on a characteristic). The second approach is expressed in studies conducted for other purposes: (f) to evaluate the effect of an intervention or treatment on QoL[1,4-7,9]; or (g) to evaluate the responsiveness to change of a new QoL instrument. These two research aims emphasise the possibility of QoL to change over time. We call this the *plasticity* approach to change, which is defined as describing the ability of an individual to change in characteristics over time. The plasticity approach is considered to be the changeable and time-variable aspect of development.[10] Predictability (invariability) and plasticity (changeability) intuitively represent opposing characteristics. Therefore, the way change is approached needs to be further elaborated as we endeavour to study longitudinal QoL in children.

7.4 Questions to be answered in the review

In this systematic review the answers on the following questions were searched for.

- How many QoL studies in children used a longitudinal perspective? In what area and in what age ranges were they performed?
- Has QoL in these publications been defined and measured according to the current consensus?
- Did the aim of the study have implications for the definition of QoL, the research design, or the approach to change?
- Is it possible to draw general conclusions from these studies about QoL changes in children?

7.5 Method

7.5.1 Criteria for Selecting Studies

- Only original studies were included; reviews or theoretical papers without new data were excluded.
- The majority of the subjects had to be between the ages of 0 and 12 years at the first QoL assessment.
- At least two assessments of QoL had to be reported
- The authors had to declare that their instruments assessed QoL
- It had to be a prospective study
- Studies with children with a mental retardation or with psychiatric patients were excluded.

7.5.2 Literature Base

We conducted a computer search using the following CD-ROM data bases: MEDLINE Express (1966-7/98), OVID-MEDLINE (1966-7/1998), PsycLIT (1967-6/1998), PsycLIT Journal Articles (1991-12/1997), PsycLIT Chapters&Books (1974-12/1997), EMBASE (1979-1997), CC Search (1995-8/1998) and Pascal BioMed (1990-1997). To identify QoL studies the search terms Quality of life and Life Quality were used. In order to find studies in children between 0-12 years of age the following search terms were used: child, children , childhood, pediatr*, paediatr* or infant.

To include longitudinal studies the following terms were used: longitudinal, longterm, long-term, long term, followup, follow-up, follow up, stability, stable, change, changes, increase, decrease, improve*, develop*. The position of the search terms in the records was not restricted, because longitudinal measurement of QoL in children was allowed to be an incidental or subsidiary aspect of a study. Search hits were captured into a computer database. Doubles were merged in a way that original CD-ROM data bases could be traced. The resulting references were all manually/visually studied to see if the selection criteria were met. If the publication language of the selected reference was English, a reprint of the paper was obtained from the university library. Available papers were studied in full in order to conclude whether they fitted the selection criteria.

7.5.3 Data synthesis

The numeric result of the computer search was presented. Furthermore, the general characteristics of the studies were listed using the following elements: year of publication, country in which the study was performed, years of children's birth, sample size, age at first assessment, description of subjects' characteristics, study aim, importance of QoL in the study (main or subsidiary objective) and variables that were measured in addition to QoL.

The QoL assessment was given using the following elements: name of the QoL instrument(s), generic or disease specific instrument, type of instrument (utility, uni-dimensional or global, multi-dimensional, or battery approach), informant of QoL, objective or subjective evaluations, QoL domains (physical, psychological or social functioning), and QoL definition if provided by the authors.

The longitudinal QoL research design was evaluated using the following elements: research type (experimental, quasi-experimental or observational), an assessment diagram in which number of observations and time between observations were given, total period of assessments, instrument's recall period, sample size at the start of the study (both total size and group sizes), sample size at the end of the study, and longitudinal statistics used in the study to test the longitudinal changes.

An evaluation was made of the approaches to change that underlie the studies. The predictability approach was illustrated by the question: Was the QoL presumed to be stable (or continuous), and was this supported by the results of the study? The plasticity approach was illustrated by the question: What was presumed to elicit changes in QoL, and was this supported by the results of the study?

7.6 Results

7.6.1 The numeric result of the computer search

The search in the CD-ROM data-bases resulted in 4064 hits. After merging the doubles, 2573 references of publications remained. Seventy percent of the references would have been found by solely using the MEDLINE Express data-base. None of the other databases could have given the remaining 30% on its own. Of the 2573 references only 115 met the selection criteria. In this selection, nine publications were non-English (French,[11-14] Russian,[15] Swedish,[16] Italian,[17] Spanish[18] or German[19]) and were therefore excluded. One publication was not available in the Netherlands.[20] It was decided that this publication probably was difficult to obtain in other countries as well, and could therefore be excluded.

The remaining 105 publications were studied in full. Although these 105 publications all seemed to meet the selection criteria according to the information obtained from the data-base, only 32 fully met the selection criteria.[21-52] The general characteristics of these publications will be presented in the next paragraph. If only MEDLINE Express would have been used, 27 publications (84%) from the final selection would have been found anyhow. In addition, four publications (12%) originated from the CC data-base. Although CC is specifically known for its up-to-date information, surprisingly these four were not the most recent publications in the selection.[22,26,27,34] The last article could have been found by using PsycLIT, PsycLIT Journal Articles or EMBASE.[45]

The search terms used resulted in many hits that appeared not to be useful in the end. For example, if somewhere in the reference the word 'child' was used, it was selected even if children were not the objective of the publication. Furthermore, in many references the abstract ended with recommending the study of QoL in future research, although QoL was not the objective of the publication. In other cases, the keyword QoL was given in the data bases to references that did not use the word QoL in the paper at all. Nevertheless, in retrospect it was not possible to use a better search term when conducting the computer search.

Closer inspection revealed that some papers were related to each other. It concerned Juniper et al.[37] with Guyatt et al.[25] and Cleary et al. [46] with Donadieu et al.[32] As they each had somewhat different study aims, they were not removed from the selection.

Table 1. General characteristics of the reviewed studies

Refr. nr.	First author	Publ. Year	Country	Years of birth	N	Age at first assessment(1)	Description of subjects characteristics: children with ...(2)	Study aim	QOL main or side objective? (3)	Other variables measured next to QOL?(4)
21	Juniper	1998	Canada	±1987	75	M=9.8 y, SD=1.9 y.	rhinoconjunctivitis (In retrospect: I1=stable, I2=changed)@2	G	main	phy (#3)
22	Barr	1997	Canada	±1983-1996	18	Med=3y11m, range=11m. to 14y.	standard (I1t) and high risk (I2t) cancer @2	F	main	phy (-)
23	Cantani	1997	Italy	±1983-1989	300	range=3.5-7.5, Med=4.4	asthma(It=treatment, Ic=control) @2	F	side	phy (-)
24	Bartholomew	1997	USA	±1979-1996	199	M=8.6, range= 0-18y	cystic fibrosis (It=treatment group,Ic=control group)@2	F	side	psy (-)
25	Guyatt	1997	Canada	±1979-1989	52	M=12, SD=3.1, range=7-17 (age groups: 7-10 y., 11-17 y. results separately reported)	asthma	G	main	phy (-) or (#3)
26	Gill	1997	Great Britain	±1979-1989	50	range= 6-16	sickle cell disease(I1=sickle cell anaemia, I2=sickle cell disease) and R=healthy children @2	G	main	phy (#3)
27	Juniper	1997	Canada	±1979-1989	52	range=7-17y	asthma (In retrospect: I1=stable, I2=changed)@2	E+G	main	phy(#3), psy (#3), S (#3)
28	Meltzer	1997	USA	±1978-1990	204	range=6-18y	It=perennial rhinitis treatment group, Ip=perennial rhinitis placebo group @1	F	side	phy (-)
29	Parkin	1997	Canada	±1976-1991	35	range=5-20, separate results on 5-12y and 13-20y	spina bifida	E	main	psy (#1)
30	Iorio	1997	Italy	±1976-1987	94	Med=8 y., range: 3-14 y.	chronic viral hepatitis	F	main	phy (-)
31	Rosenfeld	1997	USA	±1975-1996	186	Med=3.4 y., range 6m-12y.	otitis media (In retrospect: I1=stable, I2=changed)@2	E+G	main	phy (#1)
32	Donadieu	1997	France	±1966-1991	19	range= 0.4-23.7, Med=5,25	severe chronic neutropenia	F	side	phy (-)
33	Kazak	1996	USA	±1989	162	M=5.69, SD=4.39	leukemia (It1=medical intervention, It2=medical and psychological intervention, Ic=control) @2	F	side	psy (-)
34	Dossetor	1996	Australia	±1982-1995	28	M=5y5m; range=7m-13y6m	inpatients of children's hospital	G	main	phy (-)
35	Berth-Jones	1996	England	±1979-1993	27	M=9y, range 2-16y	eczema	F	side	phy (-)
36	Carpay	1996	The Netherlands	±1979-1991	80	range=4-16	epilepsy	E	main	phy (#1)
37	Juniper	1996	Canada	±1979-1989	52	M=12, SD=3.1, range: 7-17 (age groups: 7-10 y., 11-14 y., 15-17 y.)	asthma (In retrospect: I1=stable, I2=changed)@2	G	main	phy (#3)
38	Spencer	1996	England	could not be deduced	168	"infants and prescholers"	normal' children, developmental problems, acute illnesses, and chronic illnessess (-)	E	side	phy (-)

Refr. nr.	First author	Publ. Year	Country	Years of birth	N	Age at first assessment(1)	Description of subjects characteristics: children with ...(2)	Study aim	QOL main or side objective? (3)	Other variables measured next to QOL?(4)
39	Konishi	1995	Japan	±1983-1991	5	range=2y5m-10y11m	epilepsy	F	side	phy (-)
40	Marrero	1995	USA	±1981	106	M=13.3y, SD=4.5y	diabetes (It=treatment group, Ic=control group)@2	F	side	phy (-), psy (-) and S (-)
41	Buchanan	1995	Australia	±1980-1991	15	range=3-26y. Seperate results of children 3-14 y and adolescents/young adults 15-26y	epilepsy	F	side	phy (-)
42	Gemke	1995	The Netherlands	±1976-1991	468	M=55mnd=4,58y, range=1m-16y	children from a tertiary paediatric ICU (trauma patients excluded)	A	main	phy (#2)
43	Eiser	1995	England?	±1975-1987	35	M=14,4, range=8-20	cancer or a medical history of cancer	E	main	phy (#1) psy (#1)
44	French	1994	England	±1981-1989	535	range=4-16y. Separate results for 4-7y.(M=5.5), 8-11y(M=9.7), 12-16y (M=13.8).	I=asthma , R=healthy children @2	E	main	phy (-)
45	Aldenkamp	1994	The Netherlands	±1981	200	I: M=12,8 C" M=12.6	I=outgrown epilepsy, R=healthy children @2	C	main	phy (-)
46	Cleary	1994	England, France, Germany and Poland	±1976-1993	130	range=4,5m(0,375y)-18y	congenital agranulocytosis	F	main	not reported
47	Van-Damme-Lombaerts	1994	Switzerland	±1968-1988	107	Med=11,6, range 6m - 20 y	chronic renal failure	F	side	phy (-)
48	Morris	1993	England	±1980-1990	11	Med=6.7y, range=2.3-12.3	end stage renal failure (It1=treatment arm one, It2=treatment arm two) @2	F	side	phy (-)
49	Wray	1992	UK	±1972-1990	28 + ?R	I: M=8.63, range=0.1-16.0y, R: M=8.59, range=0.3-15.8y	I=heart or heart-lung transplantation, R= healthy children @2	F	main	phy (-)
50	Hoppe-Hirsch	1990	France	±1957-1986	120	range=1-15, peak at 5-6 years	operated for medulloblastoma	A	side	phy (-)
51	Becker	1988	USA	±1987	51	range=2w-12w	healthy infants with infant colic	F	side	phy (-)
52	Kekomäki	1981	Finland	±1963-1975	7	range=5-17y	incontinence due to congenital abnormalities (It1=treatment arm one, It2=treatment arm two) (-)	F	side (in introduction), main (in discussion)	phy (-)

(1) M=average,Med=Med, m=months, y=years

(2) @1: between group comparison, not tested; @2 : between group statistics; (-): no test or comparison

(3) A=describing a group, B=describing QOL development, C=identifying determinants, D=predict morbidity/mortality from baseline QOL, E=testing instruments F=treatment evaluation, G=testing instruments responsiveness to change

(4) phy=physical, psy=psycological,S=sociological (#1)= comparison between QOL and other parameter,not tested ; (#2) = statistics between QOL and other paremeter; (#3) = statistics between change in QOL and other parameter; (-): no test or comparison

7.6.2 General characteristics

The general characteristics of the 32 publications are given in Table 1. As can be seen, the selected papers were published between 1981 and 1998. The distribution of the *publication years* is heavily skewed towards the more recent years, illustrating the flourishing development of children's QoL instruments since the nineties. Moreover, 12 of the publications had instrument development as their *aim of study.* Five of these papers tested reproducability (see aim (e) in the introduction),[29,36,38,43,44] five tested responsiveness to change (aim g),[21,25,26,34,37] and two papers tested both.[27,31] Seventeen studies aimed at treatment evaluation (aim f),[22-24,28,30,32,33,35,39-41,46-49,51,52] two at describing a particular group (aim a),[42,50] and one at identifying determinants of QoL (aim c).[45] None of the publications attempted to describe developmental processes (aim b), and none aimed at predicting morbidity and mortality by using QoL as baseline data (aim d).

Seventeen publications had QoL as the *main objective* of the study,[21,22,25-27,29-31,34,36,37,42-46,49] 14 considered QoL as subordinate objective.[23,24,28,32,33,35,38-41,47,48,50,51] One publication referred to QoL as a side issue according to the introduction, but as main objective according to the discussion.[52] Although QoL is considered to include physical, psychological as well as social functioning, most studies used separate instruments to measure one of these aspects apart from the QoL instrument. Twenty-eight papers included extra physical *variables* obtained from physicians or parents.[21-23,25-28,30-32,34-45,47-52] Six papers included extra psychological variables,[24,27,29,33,40,43] two studies included extra social variables,[27,40] and two included no other assessments than QoL assessments.[44,46] Only four publications directly tested the relation between change in QoL and other variables.[21,25,26,37]

Some publications reported several studies or study phases.[21,36,43] The description of the study sample concerned the longitudinal parts of these publication only. The *sample size* ranged from 5 to 535 with a median of 75 children. All papers studied *children with a physical disorder,* although some of them included a healthy reference group in addition.[38,44,45,51] Twelve publications used several sub-groups,[22-24,26,28,33,40,44,45,48,49,52] between group statistics were available in all but one.[52] In addition, four publications started with one group and ended up with two groups in retrospect, with children that changed or were stable.[21,27,31,37]

As a rule, management and treatment of disorders improves or changes during the years, and the conclusions drawn from certain populations could be outdated.[5,9] Therefore, it was considered important to report the *years of children's birth.*

Unfortunately none of the publications reported these. To have at least some indication, the years of birth were estimated using the age of the children, the years of enrolment and the date the papers were received or accepted by the journals. As a result, the real years of birth could be earlier than given in the table, which is stressed by the '±'-sign. One of the selection criteria was that the subjects had to be primarily between the ages of 0 and 12 years. We had to interpret this criterion rather liberally because the cut-off points were rarely in this *age range* and the exact distribution of ages was not always completely clear. Some publications compared several age groups in their study.[25,29,37,41,44]

7.6.3 QoL assessment

In the next section the QoL assessment information in Table 2 is discussed in relation to the general characteristics in Table 1. As can be seen in Table 2, various instruments or techniques are used to measure QoL. Twenty publications presented the measurement properties of their main QoL instrument.[21,22,24-31,33,34,36-38,40,42-44,49] Obviously, all publications that had instrument testing as their aim, belonged to this category. Two studies provided very little information but suggested good measurement properties.[45,46] As many as 10 publications used instruments that had not been validated or tested at all.[23,32,35,39,41,47,48,50-52] These were not necessarily the oldest publications, although the oldest one was amongst them.[52] Instead, these were all publications that used QoL as a subordinate objective. Eleven publications used a generic instrument,[24,26,30,34,38,42,45,47-50] 11 used a disorder specific instrument,[21,23,25,28,32,33,35,36,40,43,52] seven used both[22,27,29,31,37,44,46] and in three publications it was unclear what was used.[39,41,51]

As stated before, QoL should be multi-factorial (physical, psychological and social well-being), patient self-administered, and subjective.[1,3] A multi-factorial measure could be obtained by multi-dimensional instruments as well as by utility instruments, although the last ones use often a final sum score instead of a profile.[8] A battery approach (a combination of instruments) could be multi-factorial but is less useful, because the instruments in the battery usually have different formats that are difficult to combine in a profile.[6] Ten publications used multi-factorial instruments measuring physical, psychological as well as social well-being,[26,28,29,30,33,35,40,43,44,48] but only seven of these reported good measurement properties concerning reliability and validity as well.[26,28,30,33,40,43,44] Eleven publications used the child as informant,[21,25-28,35,37,43-45,47] five of these compared a proxy informant like parents or clinician with the child.[21,25,35,43,45]

Table 2. QOL assessment in the reviewed studies

Refr. nr.	Name QOL instrument (1)	Generic or specific intrument	Type of instrument	Informant (2)	Objective or subjective evaluations	QoL domains (3)	QoL definition
21	Paediatric Rhinoconjunctivitis Quality of Life Questionnaire (PRQLQ) (!)	specific	multi-dimensional	children, clincians, parents ($2)	subjective	O, Ph.	HRQoL not only measure how much patients are bothered by their symptoms, they also measure the impact that the symtoms have on the day-to-day functioning (physical, social, occupational and emotional).
22	A. Overall assessment of Health status (?); B.classification into 4 temporary health states (?); C.Health utilities index mark 2 (HUI2) (!); D.Health utilities index mark 3 (HUI3) (!)	A, C and D are generic; B. is specific	A. global, B. multi-dimensional, C&D. Utility	nurse, physicians, parent ($2)	A,B,C,D :subjective	A: O B: Ph, Psy C&D: Ph,Psy	HRQoL: A: global rating of the subject's health status, B: classification into temporary health states C+D: health state.
23	diary card (?)	specific	multi-dimensional	parent	subjective	Ph	QOL: not reported (limitations of the quality of life per year).
24	"measures of health and quality of life variables" (11 different instruments) (! + ?)	generic	battery	child-or-parent (-)	subjective	Ph,Psy,S	HRQoL: not reported
25	Paediatric Asthma Quality of Life Questionnaire (PAQOL) (!)	specific	multi-dimensional	child, parent ($1)	subjective	Ph, Psy	HRQoL: not reported
26	Central Middlesex Hospital Children's Health Diary (CMHCHD) (!)	generic	multi-dimensional	child (11-16) child partly with help of parent (6-10) (-)	subjective	Ph,Psy,S	QOL: not reported (health status)
27	A:Paediatric Asthma QOL questionnaire (PAQLQ) (!), B: Health Utilites Index (HUI) 2 and 3 (!), C: the Feeling Thermometer (!), D: Standard Gamble (!)	A: specific, B: generic, C:specific, D:specific	A: multi-dimensional, B: utility, C:global, D: utility	child	At/mD: subjective	A&B: Ph, Psy C&D: O	HRQoL: A. impact of asthma condition on children's day-to-day life, B:health status, C. health state and the value children place upon it D: value that patients place on their own health state.
28	Assessment of quality of life in adolescents and children with allergic rhinoconjunctivitis (!)	specific	multi-dimensional	child	subjective	Ph,Psy,S	QOL: relevance of rhinitis to activities and moods.
29	A: spina bifida HRQOL instrument (!), B:global question of well-being (?)	A: specific, B: generic	A: multi-dimensional, B: global	parent (5-12y) and child (13-20y) (-)	A&B: objective	A: Ph, Psy, S, B: O	HRQoL: construct encompasses physical and occupational function, psychological state, social interaction and somatic sensation.
30	Sickness Impact Profile (SIP)(!)	generic	multi-dimensional	child-or-parent (-)	subjective	Ph,Psy,S	HRQoL: evaluates the impact of a disorder on the patient's HRQoL as perceived through its effect on patient daily activities, feelings and attitudes.
31	A. The 6-item health-related QOL survey (OM-6) for chronic and recurrent otitis media (!), B. Global measure of ear-	A. specific B. specific	A. multi-dimensional, B.global	caregiver	A. subjective, B. subjective	A: Ph,Psy, B: O	HRQoL is a subjective outcome that reflects the patient's perception of his or her health status.
32	self-made questionnaire (?)	specific	multi-dimensional	parents	subjective	Ph	QOL: not reported
33	Pediatric Oncology Quality of Life Scale (POQOLS) (!)	specific	multi-dimensional	mother, father ($1)	subjective	Ph,Psy,S	QOL: frequency of pediatric oncology patients' daily activity.
34	The RAHC Measure of Functioning (MOF) (!)	generic	global, utility?	clinicians and parents ($2)	subjective	O	HRQoL: a broad concept of child health covering 'physical, mental and social well-being, not merely the absence of disease and infirmity.
35	self-made questionnaire (?)	specific	multi-dimensional	child, parent ($1)	subjective	Ph,Psy,S	QOL: impact of the disease on child an family.
36	The Hague seizure severity (SS) and The Hague side-effects(SE) scales (!)	specific	multi-dimensional	parent	subjective	Ph	QOL is a multidimensional concept with physical and psychosocial issues.

Refr. nr.	Name QOL instrument (1)	Generic or specific intrument	Type of instrument	Informant (2)	Objective or subjective evaluations	QoL domains (3)	QoL definition
37	A. Paediatric Asthma Quality of Life Questionnaire (PAQOL) (!); B: The Feeling Thermometer (!), C: Global rating of change questionnaire (!)	A. specific, B. generic, C:specific	A. multi-dimensional, B: global, C: multi-dimensional	child	subjective	A: Ph,Psy, B: O, C: Ph,Psy	QOL: A: the impact of asthma on their lives. B: how the patient feels about his or her own health state, C: not reported.
38	one domain of the The Warewick Child Health and Morbidity Profile (WCHMP) (interview or postal questionnaire) (!)	generic	global	parent	subjective	O	HRQoL: overall picture of perception of the child's health and illness experience.
39	not reported (?)	not reported	not reported	not reported	not reported	Ph, Psy	QOL: demonstrated by (increases in) activity, appetite, conversation and good humor.
40	Diabetes QOL for youth (DQOLY) measure	specific	multi-dimensional	not reported	subjective	O, Ph, Psych, S	QOL: not reported
41	"observations" "could not be assessed in a quanititative fashion" (?)	not reported	not reported	parents, carers, physicians (-)	subjective	Ph, Psy	QOL manifest by improved alertness, cognition and general functioning.
42	multiattribute health status classification (MAHSC) (!)	generic	utility	parent	subjective	Ph, Psy	HRQoL: functional abilities from a patient's perspective, health status.
43	Perception of illness experience (PIE) (!)	specific	multi-dimensional	child, parent ($2)	subjective	Ph,Psy,S	QOL: child's perception of the illness experience; multi-dimensional: subjective.
44	Childhood Asthma Questionnaires: CAQA (4-7y), CAQB (8-11y.), CAQC (12-16y.) (!)	specific and generic part	multi-dimensional	child	subjective	Ph,Psy,S	QOL = HS related to how a person feels, and how he/she functions in daily activity". QOL=multidimensional.
45	Holmfrid Quality of Life Inventory (?!)	generic	multi-dimensional	parents and children (#1)	subjective	Ph, Psy	QOL: defined as a multidimensional construct that covers physical, emotional, mental, social and behavioral components of well-being and function as perceived by patients and observers.
46	HRQOL instrument self-made based on a synthesis of existing measures (?!)	generic and specific	multi-dimensional	parent	subjective	Ph	HRQoL: not reported
47	self-made questionnaire (?)	generic	multi-dimensional	parent (<8y) and child (>7y) (-)	subjective	Ph,Psy	QOL: not reported
48	self-made instrument (?)	generic	multi-dimensional	parent	subjective	Ph,Psy,S	QOL: various aspects of the child's well being and behaviour.
49	various instruments (!) and self-made instrument (?)	generic	battery	parents, children, teachers, test-results (psychologist) (-)	objective and subjective	Psy, O	QOL is being used as an indicator of the success of a particular intervention with increasing frequency, although as a concept it is subjective and therefore unscientific.
50	various instruments (?)	generic	battery	test-results	objective	Ph, Psy	QOL: not reported
51	self-made questionnaire (?)	not reported	global	parent	subjective	not reported	QOL: not reported
52	self-made questionnaire (?)	specific	not reported	parent	subjective	not reported	QOL: not reported (leading a normal life).

(1) (!)=good measurement properties, (?)=no measurement properties provided, (?!)=suggested good properties
(2) ($1)=cross-informant comparison, not tested; $2=cross-informant statistics; (-)=no test or comparison
(3) Ph=physcial, Psy=psychological, S=social, O=overall or global

All but three publications[29,49,50] assessed subjective measurements of QoL. One publication stated, remarkably, that the concept of QoL is subjective and therefore unscientific.[41] Only four publications satisfied all requirements.[28,29,43,44] They covered physical, psychological as well as social well-being in one instrument and reported good measurement properties, assessed QoL by the children themselves and asked for their subjective opinion. Some publications almost met the requirements but asked the parents to 'help if necessary'.[24,26,30] In this way, however, it is not clear how much the parent provided the answers instead of the child. All other studies lacked at least one of the requirements.

Finally, the column 'QoL definition' gives a nice illustration of the lack of consensus about a definition of QoL. Twelve publications used the term Health Related Quality of Life (HRQoL),[21,22,24,25,27,29,30,31,34,38,42,46] but the definitions of these studies did not substantially differ from the ones that used the term QoL. Eleven publications did not report a definition at all,[23-26,32,40,46,47,50-52] although three of them had QoL as their main study objective.[25,26,46] One paper administered an open question to explore the parental understanding of the concept QoL[38]: Eighty-seven percent of the parents thought that having a loving, caring family was most necessary in order for a child to have a good QoL. Good food, activity, health, and happiness were considered of lesser importance.

7.6.4 Longitudinal QoL research design

Table 3 starts with a schematic diagram of the research design focussing on quantity and timing of observations and interventions. As can be seen, in two publications the length of the period between assessments is not reported.[34(phase3 of 38)] The length of the total assessment period varies between 1 week (=0.25 months).[31,44,51] and 10 year (=120 months).[50] In Figure 1 the total assessment period from Table 3 is combined with the aim of the study from Table 1. On average, publications that describe a particular group (aim a) had the longest assessment periods, publications that aimed at reproducability (aim e), responsiveness to change (aim g) or both (aim e+g) had the shortest assessment periods, and the ones that aimed at treatment evaluation (aim f) or at identifying determinants (aim c) had assessment periods that fell between the two extremes.

Thirteen publications reported the instrument's recall period,[21-23,26,27,29,31,33,34,37,42,43,46] that varied between the 'previous three months'[42] and 'at that point of time'.[22,34,43] In one publication the recall period of the instrument coincides with the

period between assessments.[31] Ten publications did not report the sample size at the end of the study.[23,28,30,33,34,40,41,44,48,52] Eight publications did not report a longitudinal statistic to test the change in QoL,[23,25,28,39,41,45,50,51] although some had QoL as their main objective.[25,45] Of the four publications that satisfied all requirements for QoL assessment in Table 2, [28,29,43,44] all had clear assessment descriptions,[28,29,44,43] some reported a recall period,[29,43] some had good description of the sample size at the end of the study,[29,43] and some used longitudinal statistics,[29,43,44] leaving two studies that met QoL assessment as well as longitudinal requirements.[29,43] Both had a rather large age range : between 8-20[43] and between 5-20,[29] but one presented separate results for the 5-12 and 13 to 20 year olds.[29] This is a rather small number for drawing general conclusions about QoL changes in children.

Figure 1. The total assessment period in months by study aim

Legend:
A=describing a group,
B=describing QOL development,
C=identifying determinants,
D=predict morbitiy/mortality from baseline QoL,
E=testing instrument's reproducability,
F=treatment evaluation,
G=testing instrument's

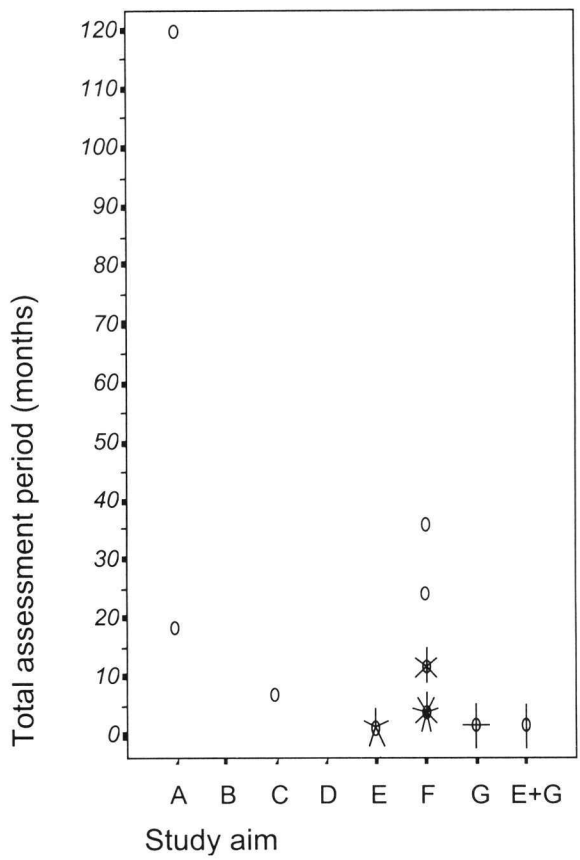


Table 3. Longitudinal QOL research design in the reviewed studies

Refr. nr.	Research type (1)	Assessment diagram (2)	Total period (in months)	Instrument(s) recall period	Sample size at start (3)	Sample size at end (3)	Longitudinal statistics
21	obs.	I: O1 - 1w-> O2 -2w-> O3	0.75	previous 7 days	75	74 (Istable=13, Ichange=61)	paired t-test, intraclass correlation coefficient
22	quasi.	It1: StartMI -1w-> O1 - 1w-> O2 -1w-> O3 It2: StartMI -1w-> O1 - 1w-> O2 -1w-> O3	0.75	A+B not reported, C+D: at that point of time	18 (It1=9, It2=9)	18 (I1=9, I2=9)	ANOVA, paired t-test, intra-class correlations
23	quasi.	It: start MI -1year(daily dairy)->O1 -1year(daily dairy)-> O2-1year(daily dairy)->O3 Ic: start -1year(daily dairy)->O1 -1year(daily dairy)-> O2-1year(daily dairy)->O3	36	1 day	300 (It=151, Ic=149)	not reported	not reported
24	quasi.	It: O1 ->PI -18/32mm-> O2 Ic: O1 -18/32m->O2	24	not reported	199 (It=104, Ic=95)	184 (It=95, Ic=89)	ANCOVA between groups with the pretest as covariates
25	obs.	I: O1 -1w-> O2 -4w-> O3 -4w-> O4	2.25	not reported	52	52	not reported
26	obs.	I1: O1-every day->O28 I2: O1-every day->O28 R: O1-every day->O28	1	1 day (missing data retrospectively completed within 1 week)	50 (I1=14, I2=11, R=25)	50 (I1=14, I2=11, R=25)	Mann-Whitney U test
27	obs.	enrolment -1w-> O1 -4w-> O2 -4w ->O3	2.25	A: previous week, B:"time frame is not specified", C:previous week, D:previous week.	52	52 (Istable=37, Ichange=15)	Intra-class correlation coefficient (ICC) and paired t-test
28	exp.	It: O1 -1w-> start Pla -1w >-start MI -4->end MI-1->O2 Ic: O1 -1w-> start Pla -1w >-start Pla -4->end Pla-1->O2	1.75	not reported	204 (It=102, Ip=102)	not reported	percentage
29	obs.	O1 -2w-> O2	0.5	A: not reported, B: at present	35	28	intra-class corelation coefficient
30	quasi.	O1,startMI - 1m-> O2 -3/12 m-> O3,endMI -1m-> O4 -2m-> O5	7 to 16	not reported	94	not reported	student t-test
31	quasi.	phase I responsiveness: O1 -±2w-> MI -±2w-> O2 phase II test-retest: O1 -1w-> O2	1 (phase I) ; 0.25 (phaseII)	A+B: during the past 4 weeks	186	110 (phase I=50 phase II=60)	A: standardised response mean SRM (mean change score divided by its SD and 95% confidence interval. B: correlations
32	quasi.	O1 -15d-> O2 start MI -1m->O3-1m->O4-1m->O5-1m-> O6-1m->O7-1m->O8	6.5	not reported	19	17	analysis-of-variance model

Refr. nr.	Research type (1)	Assessment diagram (2)	Total period (in months)	Instrument(s) recall period	Sample size at start (3)	Sample size at end (3)	Longitudinal statistics
33	quasi.	It1: DstartMI -1m-> O1 -1m-> O2 -4m-> O3 It2: DstartMI&PI -1m-> O1 -1m-> O2 -4m-> O3 Ic: D -6m-> O3	6	previous 2 weeks	162 (It1=45, It2=47, Ic=70)	not reported	repeated measure analyses of covariance (ANCOVA)
34	obs.	O1admission -?->O2discharge	not reported	current level of functioning	28	not reported	Wilcoxon's signed rank test
35	quasi.	O1startMI-6w->O2endMI -2w->O3	2	not reported	27	20	Wilcoxon matched-pairs, signed ranks test (two-tailed)
36	obs.	O1 -14d-> O2	0.5	not reported	22	18	test-retest reliability: Pearson's R
37	obs.	I: O1 -1w-> O2 -4w-> O3 -4w-> O4	2.25	A: previous week, B: not reported, C: since previous visit(1 to 4 weeks)	52	100 observations of 4 weeks (Istable = 46 obs., Ichanged =54 obs.)	paired t-test, Pearson correlations, within-subject standard deviation of change
38	obs.	phase 1: O1 -2w to 3 m-> O2 phase 2: not longitudinal phase 3: VO1 -?w->VO2	phase 1: 0.5 to 3. phase 3: not reported	not reported	phase 1: 128, phase 3: 40	phase 1: 88, phase 3: not reported	weighted kappas
39	quasi.	MI -1to3m->O1 -3mto11m->O2	6 to 12	not reported	5	5	not reported
40	exp.	It: O1 start SI -12?m->O2 Ic O1 -12?m-> O2	12	not reported	106 (It=52, Ic=54)	not reported	repeated measurement statistic
41	quasi.	startMI -1 to 2 m.-> O1 -1 to 2 m.->O2 -1 to 2 m.-> O3 -1 to 2 m.-> O4)	6	not reported	15	not reported	percentage of children with improvement
42	obs.	O1 -±3m-> A -4.4d-> Di -1y-> O2	18	previous 3 months	468	254	described (in %) by comparing the number of affected domains before admission with that one year after discharge
43	obs.	VO1 -(≤2m)->VO2	2 or less	as they felt 'now'	35	28	test-retest reliability

To be continued at the next page.

(1) obs=observational, quasi= quasi-experimental, exp: experimental; [Decision rule: I: are the subjects randomly assigned to conditions? II: has the experimenter functional control over independent variable(s): I yes+ II yes = exp.; I no+II yes = quasi.; I no+ II no = obs.]

(2) O=observation, D= newly diagnosed, MI=medical intervention, PI=psychological intervention, SI=sociological intervention, V=clinical visit, A=Admission to clinic, Di=discharge from clinic, Pla=placebo

(3) I=Index, It=Indexed treatment, Ic=Indexed control, Ip= Indexed placebo, R=healthy reference group

Continuation of Table 3

Refr. nr.	Research type (1)	Assessment diagram (2)	Total period (in months)	Instrument(s) recall period	Sample size at start (3)	Sample size at end (3)	Longitudinal statistics
44	obs.	I4-7y: O1 -1w-> O2 R4-7y: O1 -1w-> O2 I8-16y: O1 --3w-> O2 R8-16y: O1 --3w-> O2	4-7y: 0.25 ; 8-16y 0.75	not reported	535 (I4-7y=80 ; R4-7y=103; I8-11y=103; R8-11y=153; I12-16y=98)	not reported	test-retest: Pearson corr., median scores, intraclass correlations
45	quasi.	I: O1 -3m withdrawal of medication-> end medication -4m -> O2 R: O1 -7m -> O2	7	not reported	200 (I=100, R=100)	166 (I=83, R=83)	difference scores
46	quasi.	O1 -2w-> O2 ->startMI -1m->O3 -1m-> O4 -1m-> O5 -1m-> O6 -1m->O7 -1m->O8	6.5	"previous two weeks" or not reported	19	14	repeated measures analysis of variance
47	quasi.	O1 start MI-6m-> O2 -6m-> O3	12	not reported	107	44	Friedman's test
48	exp.	It1: O1 -2w->start MI -12w-> O2 -12w->O3 start Placebo-12w-> O4 -12w->O5 endMI It2: O1 -2w->start Placebo -12w-> O2 -12w->O3 start MI -12w-> O4 -12w->O5endMI	12	not reported	11 (It1=6, It2=5)	not reported	paired t-test
49	obs.	I: O1 -(M=8m,range 1-18m)->MI -3m->O2 R: O1	4 to 21	not reported	28+? (I=28, R=not reported)	28 + ? (I=28, R=not reported)	t-tests and Mann-Whitney U-test
50	obs.	MI -5y-> O1 -5y->O2	120	not reported	120	64	percentages (not tested)
51	quasi.	V-1d-> O1-1w->O2	0.25	not reported	51	51	not reported
52	exp.	It1: O1 MI-3->O2 Pla -3w->O3 MI -3w->O4 Pla-3w-> O5MI -3w->O6 Pla-3w->O7 It2: O1 Pla-3->O2 MI -3w->O3 Pla -3w->O4 MI-3w-> O5 Pla -3w->O6 MI-3w->O7	4.5	not reported	7 (3 versus 4 since randomization table is used)	not reported	binomial test for paired observations

(1) obs=observational, quasi= quasi-experimental, exp: experimental; [Decision rule: I: are the subjects randomly assigned to conditions? II: has the experimenter functional control over independent variable(s): I yes+ II yes = exp.; I no+II yes = quasi.; I no+ II no = obs.]

(2) O=observation, D= newly diagnosed, MI=medical intervention, PI=psychological intervention, SI=sociological intervention, V=clinical visit, A=Admission to clinic, Di=discharge from clinic, Pla=placebo

(3) I=Index, It=Indexed treatment, Ic=Indexed control, Ip= Indexed placebo, R=healthy reference group

7.6.5 Approaching change: prediction or plasticity

When considering the approach to change by the aim of the study, eight publications covered predictability (aim a:[42,50], aim c:[45], aim e:[29,36,38,43,44]), 22 covered plasticity (aim f:[22-24,28,30,32,33,35,39,40,41,46-49,51,52]; aim g:[21,25,26,34,37]) and two studies covered both (aim e+g:[27,31]). At first sight, the plasticity approach to change in QoL appeared far more popular. The picture changes when additional information is considered, collected by means of the two questions that illustrate the approach to change: 1) Was the QoL presumed to be stable (or continuous), and was this supported by the results of the study? (predictability); 2) What was presumed to elicit changes in QoL, and was this supported by the results of the study? (plasticity). Half of the publications appeared to use a mixed approach which might be summarised as: Predictability has to be tested in children whose physical condition has not changed, and plasticity in children whose physical condition did change.[21,24,26-29,31,36-38,43,44,46,48-50,52] Two publications presumed that changes in physical and psychological status would alter QoL.[33,43] One of these found that regardless of physical status the QoL improved during the six months period, which points to continuity or consistency in relative ranks. In other words, they presumed plasticity but found predictability.[33] The publication with study aim c (identifying a determinant)[45] measured a combination of predictability and plasticity rather than predictability alone: these investigators studied if changes in physical status could predict changes in QoL. The presumption about changing QoL by changing physical status was tested in 14 publications, either by testing between groups with or without changing physical status,[21,23,27,31,33,35,37,40,48] between disorder and healthy groups,[26,44,45,49] or by testing the influence of other variables on changes in QoL.[21,25-27,37]

7.7 Discussion

The discussion is organised around the approaches of change in QoL as found in the reviewed publications, supplemented with the methodological requirements that are needed to consider these approaches in full.

7.7.1 Predictability

About 50% of the publications presumed the QoL to be stable or continuous, at least when the physical status is stable. Evidence obtained from studies on adults, suggests that QoL is quite stable anyway, and often does not reflect changes in life

circumstances.[53] *Stability* in QoL is mainly influenced by personality traits.[53] Temperament or dispositional mood influences the QoL judgement of the individual.[53,54] and determines individual differences in the tendency to give socially desirable responses. Temperament shapes the pattern of experiences that individuals are exposed to, leading to a stable set of life circumstances,[53] guides the individual in how the experiences are interpreted, or which circumstances are noticed.[55] Another stabilising factor is adaptation and adjusting to changes in life circumstances. As a result the impact of changes in circumstances can be detected in the short term, but in the medium term QoL perceptions return to a stable baseline.[53] Adaptation can be influenced by social comparisons: seriously handicapped patients evaluate their QoL as high because they compare themselves to patients with similar problems rather than to healthy individuals.[53,56]

The effects of the child's age and level of development probably interfere with stability but can still imply *continuity*, the consistency in relative rank over time. One of the factors that influences predictions of QoL is the cognitive development of the child. The level of cognitive development influences the child's concepts of health and illness.[57] It influences the ability of the child to read and understand the QoL questions, to recall the relevant information and to formulate the answer.[5,6,58] A publication from our selection reported that if the statistics between scores of younger children and adults are worse than the statistics between scores of older children and adults, then this would be a strong indication that the younger children did not fully understand the ratings.[21] However, another publication from our selection studied the minimum skills required by children to complete various QoL instruments.[27] The Standard Gamble method, for instance, required better than grade 6 reading skills. Children are not able to consider whether they would prefer to remain in a certain health state or take a chance with a new (imaginary) treatment. The children did not have problems with the other three questionnaires in that publication. Another selected publication stated that from 4 or 5 years upwards children are able to introspect and report upon their QoL.[44] Furthermore, in a previous publication we showed that children, like their parents, can give valid information about the child's QoL, although the information can be somewhat different.[59]

It should be noted, that individual continuity does not imply a certain shape of the developmental functioning, nor does it imply that all children necessarily exhibit the same pattern of development.[10] Therefore, variance between children can increase when they become older, but a temporary increase in variance will occur too if individual

differences in timing of universal developmental events are important.[60] Some of the publications in our selection provided a comparison between two or more age groups,[25,29,37,41,44] but they did not provide variances that could support this assumption. It may be that the levels of function in various dimensions change with age like the relative weightings of QoL domains do.[3,61] The specific impact of a medical situation also varies with age. For example, hair loss associated with chemotherapy of childhood cancer, may be especially disturbing during adolescence and less during childhood.[6] A publication from our selection[37] accounted for this by individualising items from the activity domain, which was regarded most likely to show heterogeneity across age. Furthermore, they reported that in younger children fewer domains of QoL could be distinguished than in older children. As a result it is better to use a limited age range: changes in children's QoL are probably different from changes in adolescents or adults, and most instruments are age-related.

7.7.2 Plasticity

All studies presumed QoL to be changeable, at least when the physical status has changed. Since this presumption was tested in only 14 publications, we conclude that this presumption is so strong that most researchers do not feel the need to prove it. Nowadays, in defining QoL, a strictly "biomedical model of health" is replaced by a broader model in which QoL contains the three factors, physical, psychological and social functioning.[8] In approaching change, however, a biomedical model is still in use, suggesting that the only change of importance is a change in physical functioning. This approach is supported by the publications in our selection. The use of a biomedical model on QoL change is considered a *restricted view* because it ignores the possible influence of psychological or social factors on QoL. Probably the physician's interest in the influence of physical changes on QoL is related to the fact that these changes can be modified by medical treatment.[62] This means that the restriction to physical variables represents a choice and a wilful omission, which is sometimes stressed by using the term 'Health Related' QoL. In studies using the restricted view, only clinical variables were collected, and assessment of QoL was planned in relation to the medical intervention process only.[3,7,8] In our selection, 28 papers included extra physical variables in addition to the QoL data, but only six papers included psychological,[24,27,29,33,40,43] and two studies included social variables.[27,40] The restricted view furthermore implies a linear model in that it assumes that treatment A leads to physical change B which leads to QoL outcome C.[8]

Plasticity could alternatively be viewed according to a more comprehensive biopsychosocial model. This model recognises that health and QoL are determined by psychological and social as well as physical factors, all of which interact to produce the current QoL.[8] In studies using this *broad view*,[3] psychological and social variables were collected in addition to clinical variables, as was done by two papers in our selection.[29,40] Contextual (e.g. setting, sociocultural variations), interpersonal (e.g. social support, exposure to stress, major life events) and intrapersonal (e.g. coping skills, personality traits, physical health) variables are informative along with QoL measurements.[7]

Furthermore, in a broad view, medical treatments as well as psychological interventions could be beneficial in changing QoL. We found this in four publications.[24,33,40,45] One of these papers[45] studied a 'time-delay' model, which predicts that psychosocial complications may follow the development of disease symptoms, however with a considerable time delay. Consequently psychosocial reactions may persist after normalisation of physical symptoms. The parents in this study continued to report disease-related problems and remained worried although the children were considered cured. The children however showed an abrupt adaptation to news of being cured and even showed a super-positive QoL evaluation. According to a broad view, assessment of QoL is planned in relation to the physical effects of a medical intervention, as well as in relation to the psychosocial effect of the medical intervention. For instance, in addition to side effects from medication, school absenteeism, not being able to see friends, being homesick etc. are also considered important factors that could elicit changes in QoL.

The QoL changes could be the result of changing priorities and goals of the individual. Calman[63] defined QoL as the gap between the patient's expectations and achievements. The priorities and goals of an individual must be realistic and would therefore be expected to change with time and be modified by age and experience.[63] Changing priorities and goals could result in changing internal standards about what is important in QoL. This change of internal standards is also called 'response shift'. Response shift is the result of a psychological process that includes adaptation to the current physical situation. However, response shift is approached as a purely methodological problem. It is interpreted as a systematic bias, because the 'real' QoL changes as a result of a physical condition will be overshadowed.[64] Probably in order to meet with this bias, in some publications from our selection, the QoL informants were allowed to see their previous assessments.[37,48] This reasoning seems to express a restricted view on plasticity, in which psychosocial processes are interpreted as

confounders. In a broad view, response shift is the result of a combination of physical and psychological changes and not merely a measurement bias.

Measuring QoL plasticity is furthermore complicated by *situational variables* at the time of assessment. A certain mood may increase access to memories with congruent information.[53,65] Therefore, mood, diet, sleep, current level of stress, the setting (clinic, home or laboratory), all may influence the judgement of QoL.[7,8] Under time pressure or threat, adults as well as children will not consider all domains of one's life when giving an overall judgement of their QoL. Instead they will choose a simpler strategy in which the emotional state at the time of assessment will be used to base their QoL judgement upon.[1,53,58] This was even promoted, though unintentionally, by one of the selected publications, in which the investigators explained the recall period of 'last week' to the children by referring to something that happened a week ago.[27] Therefore, retrospective estimates of former QoL are highly correlated with the present state.[31] The information from a questionnaire with a very long recall period may therefore hardly differ from that assessed with a short recall period.

7.7.3 Interaction between predictability and plasticity

Although predictability (invariability) and plasticity (changeability) intuitively represent opposing characteristics, half of the publications used a mixed approach. As stated before this might be summarised as: predictability has to be tested in children whose physical condition has not changed, and plasticity in children whose physical condition did change. In some studies this was supplemented with the presumption that along with changes in physical status, changes in psychological status would alter QoL. Even the publications that tested the stability of QoL presumed that changes in physical status would alter QoL. This is understandable when taking into account the idea that factors which influence predictability in turn influence plasticity. Predictability factors like personality and cognitive development are potential moderators of QoL plasticity, both directly or through experience. Experiences influence the way the child judges his or her QoL.[5,8] This interrelation between predictability and plasticity implies that longitudinal research should contain both approaches, which has implications for the planning of the QoL assessments.[1,3,54,66]

7.7.4 Methodological considerations

Several publications discuss the basic necessities for conducting QoL studies in general[7-9,54,67,68] and longitudinal QoL studies in particular,[3,10,69,70] or give good suggestions for presentation.[71] Instead of repeating all requirements that are needed for a good longitudinal QoL design, we refer to these publications and to the results and discussion in this paper. In addition, the following aspects are considered important:

Firstly, the *sample size* should be big enough to avoid type 2 errors (concluding that there is no difference between groups when there really is a difference) although repeated measures designs have an increased power to detect between group differences.

Secondly, it would be advisable to obtain multiple *assessments* of baseline, because the baseline measurement plays a critical role in the end point.[69] In addition, QoL assessed prospectively (e.g. 'How did you feel today ') differs from QoL obtained in retrospect (e.g. 'How did you feel on a certain day one year ago'). Therefore, QoL is not recoverable once lost. Careful attention to the timing of measurement and consistency of measurement across treatment arms is important.[3,7]

Thirdly, QoL is essentially subjective and *patient self-administered*. Preferably the child him- or herself should be the informant.[44] However, children cannot always be used as informants, because they could lack the cognitive skills to understand the questions. Alternatively, a proxy respondent can be used. The closer the relationship between child and proxy, the higher the agreement is between them. Therefore, the parent is the most preferable proxy informant about the child's QoL. As the choice of informant influences the QoL judgements,[8,59,72,73] the same informant should be used at all points of measurement.

Fourthly, some of the selected publications used an instrument without limited measurement properties, and most instruments did not meet the basic requirements for measuring QoL. At present many QoL instruments for children are being developed and it must be possible to choose one that meets all requirements.[4,6,7,58,67,68,74-77]

Preferably, a multi-dimensional questionnaire with a limited number of scales should be used, because the repeated measurement of scales enlarges the volume of statistical tests needed, which enlarges the number of measurement errors. For that matter, finally, the choice of longitudinal statistics could direct possibilities in studying the changeability of QoL. For instance, correlations test the strength of a relation between time points, but not differences in height. This will imply that changes in QoL cannot be studied.

In conclusion, since many publications from the selection used large age ranges, various disorder groups, different QoL assessments and assessment periods, it is difficult to draw conclusions about children's QoL. Four publications defined and measured QoL according to the current consensus (multi-factorial, self-administered, subjective). Only two of these papers met the longitudinal requirements as well (clear assessment period, recall period, sample size at end of study, longitudinal statistics). Thus, more studies are needed that meet QoL as well as longitudinal requirements. Despite the growing consensus that QoL is variable over time, information about the underlying approach to change is scarce: Can QoL be stable over time (predictability), or if it changes, by what did it change (plasticity)? In our selection a mixed model of predictability and plasticity is used but seldom explicitly tested: stable physical health gives stable QoL and changes in physical health change QoL. It is rarely acknowledged that psychological, social and situational variables can change QoL as well. In future, more discussion is needed about how variable with time QoL really is, as this influences the planning of the assessments and guides the interpretation of changes in children as well as in adults.

7.8 Reference List

1. Leventhal H, Colman S. Quality of life: A process view. *Psychology and Health* 1997; 12: 753-767.
2. Spilker B. Introduction. In: Spilker B, ed. *Quality of life and pharmacoeconomics in clinical trials.* Philadelphia: Lippincott-Raven Publishers, 1996: 1-10.
3. Schipper H, Clinch JJ, Olweny CLM. Quality of life studies: Definitions and conceptual issues. In: Spilker B, ed. *Quality of life and pharmacoeconomics in clinical trials.* Philadelphia: Lippincott-Raven Publishers, 1996: 11-23.
4. Maillé AR, Kaptein AA, deHaes JCJM, Everaerd WTAM. Assessing quality of life in chronic non-specific lung disease - A review of empirical studies published between 1980 and 1994. *Qual Life Res* 1996; 5: 287-301.
5. Wiklund I, Erling A, Albertsson-Wikland K. Critical review of measurement issues in quality of life assessment for children with growth problems. In: Drotar D, ed. *Measuring health-related quality of life in children and adolescents: implications for research and practice.* Mahwah, New Jersey: Lawrence Erlbaum Associates, 1998: 255-271.
6. Levi R, Drotar D. Critical issues and needs in health-related quality of life assessment of children and adolescents with chronic health conditions. In: Drotar D, ed. *Measuring health-related quality of life in children and adolescents: implications for research and practice.* Mahwah, New Jersey: Lawrence Erlbaum Associates, 1998: 3-20.
7. Naughton MJ, Shumaker SA, Anderson RT, Czajkowski SM. Psychological aspects of health-related quality of life measurement: Tests and scales. In: Spilker B, ed. *Quality of life and pharmacoeconomics in clinical trials.* Philadelphia: Lippincott-Raven Publishers, 1996: 117-131.
8. Finkelstein JW. Methods, models, and measures of health-related quality of life for children and adolescents. In: Drotar D, ed. *Measuring health-related quality of life in children and adolescents: implications for research and practice.* Mahwah, New Jersey: Lawrence Erlbaum Associates, 1998: 39-52.
9. Wadsworth ME, Rodgers B, Wadsworth MEJ. Longterm follow-up studies; a critical overview. *Rev Epidemiol Sante Publique* 1989; 37: 533-547.

10. Ramey CT, Lee MW, Burchinal MR. Developmental plasticity and predictability: Consequences of ecological change. In: Bornstein MH, Krasnegor NA, eds. *Stability and continuity in mental development.* Hillsdale: Lawrence Erlbaum Associates, 1989: 217-233.
11. Donadieu J, Boutard P, Tchernia G, et al. [French: A phase II study of recombinant human granulocyte-colony stimulating factor (rHuG-CSF, lenograstim) in the treatment of agranulocytosis in children]. *Nouv Rev Fr Hematol* 1994; **36:** 441-448.
12. Loras DI, De-Potters S, Pharaon I, Olives JP, Hermier M. Qualité de vie des enfants en nutrition parenterale a domicile et de leurs parents [French: Quality of life of children with home parenteral nutrition and of their parents]. *Pediatrie* 1993; **48:** 555-560.
13. Paty E, De-Blic J, Brunet D, et al. Desensibilisation accelerée aux dermatophagoides pteronyssinus chez l'asthmatique severe. Bilan au bout d'un an d'immunotherapie [French: Accelerated desensitization with dermatophagoides pteronyssinus in severe asthmatic children. Evaluation after one year of immunotherapy]. *Arch Fr Pediatr* 1990; **47:** 173-179.
14. Pichlmayr R. Transplantations hépatiques. L'experience de Hanovre [French: Liver transplantation. The Hannover experience]. *Chirurgie* 1991; **117:** 225-229.
15. Gordeev VI, Ul'rikh EV. Kachestvo zhizni detei, perenesshikh neotlozhnye khirurgicheskie operatsii v periode novorozhdennosti [Russian: The quality of life of children who have undergone emergency surgical operations in the neonatal period]. *Vestn Khir Im I I Grek* 1997; **156:** 73-76.
16. Hyllienmark L, Ludvigsson J. Insulinpump--ett realistiskt alternative i behandlingen av diabetes hos barn och ungdom [Swedish: Insulin pump--a realistic alternative in the treatment of diabetes in children and adolescents]. *Läkartidningen* 1992; **89:** 1057-1062.
17. Micheli R, Giordano L, Balestrini MR. Tumori cerebrali in neurofibromatosi tipo 1 in eta pediatrica [Italian: Cerebral tumors in children with neurofibromatosis type 1]. *Minerva Pediatr* 1996; **48:** 89-97.
18. Rendo P, Biscigniano L, Turconi A, Sanchez AJ, Greco J. Tratamiento de la anemia en ninos sometidos a hemodialisis cronica con eritropoyetina humana recombinante [Spanish: Treatment of aenaemia in children undergoing chronic hemodialysis with recombinant human erythropoietin]. *Medicina B Aires* 1994; **54:** 392-398.
19. Scholtz W, Haubrock M, Lob CT, Gebert N, Wahn U, Szczepanski R. Kosten-Nutzen-Untersuchungen bei ambulanten Schulungsmassnahmen für asthmakranke Kinder und ihre Familien [German: Cost-effectiveness studies of ambulatory educational programs for children with asthma and their families]. *Pneumol* 1996; **50:** 538-543.
20. Willner AEE, Rodewald GE. *Impact of cardiac surgery on the quality of life: Neurological and psychological aspects.* (xiii edition) New York, NY, US: Plenum Press, 1990: 522.
21. Juniper EF, Howland WC, Roberts NB, Thompson AK, King DR. Measuring quality of life in children with rhinoconjunctivitis. *J Allergy Clin Immunol* 1998; **101:** 163-170.
22. Barr RD, Petrie C, Furlong W, Rothney M, Feeny D. Health-related quality of life during post-induction chemotherapy in children with acute lymphoblastic leukemia in remission: An influence of corticosteroid therapy - Clinical Study. *Int J Oncol* 1997; **11:** 333-339.
23. Cantani A, Arcese G, Lucenti P, Gagliesi D, Bartolucci M. A three-year prospective study of specific immunotherapy to inhalant allergens: evidence of safety and efficacy in 300 children with allergic asthma. *J Investig Allergol Clin Immunol* 1997; **7:** 90-97.
24. Bartholomew LK, Czyzewski DI, Parcel GS, et al. Self-management of cystic fibrosis: short-term outcomes of the Cystic Fibrosis Family Education Program. *Health Educ Behav* 1997; **24:** 652-666.
25. Guyatt GH, Juniper EF, Griffith LE, Feeny DH, Ferrie PJ. Children and adult perceptions of childhood asthma. *Pediatr* 1997; **99:** 165-168.
26. Gill LJ, Shand PAX, Fuggle P, Dugan B, Davies SC. Pain assessment for children with sickle cell disease: Improved validity of diary keeping versus interview ratings. *British Journal Of Health Psychology* 1997; **2 Part 2:** 131-140.

27. Juniper EF, Guyatt GH, Feeny DH, Griffith LE, Ferrie PJ. Minimum skills required by children to complete health-related quality of life instruments for asthma: comparison of measurement properties. *European Respiratory Journal* 1997; **10**: 2285-2294.
28. Meltzer EO, Orgel HA, Biondi R, et al. Ipratropium nasal spray in children with perennial rhinitis. *Ann Allergy Asthma Immunol* 1997; **78**: 485-491.
29. Parkin PC, Kirpalani HM, Rosenbaum PL, et al. Development of a health-related quality of life instrument for use in children with spina bifida. *Qual Life Res* 1997; **6**: 123-132.
30. Iorio R, Pensati P, Botta S, et al. Side effects of alpha-interferon therapy and impact on health-related quality of life in children with chronic viral hepatitis. *Pediatr Infect Dis J* 1997; **16**: 984-990.
31. Rosenfeld RM, Goldsmith AJ, Tetlus L, Balzano A. Quality of life for children with otitis media. *Arch Otolaryngol Head Neck Surg* 1997; **123**: 1049-1054.
32. Donadieu J, Boutard P, Bernatowska E, et al. A European phase II study of recombinant human granulocyte colony-stimulating factor (lenograstim) in the treatment of severe chronic neutropenia in children. Lenograstim Study Group. *Eur J Pediatr* 1997; **156**: 693-700.
33. Kazak AE, Penati B, Boyer BA, et al. A randomized controlled prospective outcome study of a psychological and pharmacological intervention protocol for procedural distress in pediatric leukemia. *J Pediatr Psychol* 1996; **21**: 615-631.
34. Dossetor DR, Liddle JLM, Mellis CM. Measuring health outcome in paediatrics: Development of the RAHC measure of function. *J Paediatr Child H* 1996; **32**: 519-524.
35. Berth-Jones J, Finlay AY, Zaki I, et al. Cyclosporine in severe childhood atopic dermatitis: a multicenter study. *J Am Acad Dermatol* 1996; **34**: 1016-1021.
36. Carpay HA, Arts WF, Vermeulen J, et al. Parent-completed scales for measuring seizure severity and severity of side-effects of antiepileptic drugs in childhood epilepsy: development and psychometric analysis. *Epilepsy Res* 1996; **24**: 173-181.
37. Juniper EF, Guyatt GH, Feeny DH, Ferrie PJ, Griffith LE, Townsend M. Measuring quality of life in children with asthma. *Qual Life Res* 1996; **5**: 35-46.
38. Spencer NJ, Coe C. The development and validation of a measure of parent-reported child health and morbidity: the Warwick Child Health and Morbidity Profile. *Child Care Health Dev* 1996; **22**: 367-379.
39. Konishi T, Masuko K, Naganuma Y, Hongou K, Yagi S. Flunitrazepam for sleep disturbance in children with intractable epilepsy. *Brain Dev* 1995; **17**: 69-72.
40. Marrero DG, Vandagriff JL, Kronz K, et al. Using telecommunication technology to manage children with diabetes: the Computer-Linked Outpatient Clinic (CLOC) Study. *Diabetes Educ* 1995; **21**: 313-319.
41. Buchanan N. The efficacy of lamotrigine on seizure control in 34 children, adolescents and young adults with intellectual and physical disability. *Seizure* 1995; **4**: 233-236.
42. Gemke RJ, Bonsel GJ, van-Vught AJ. Long-term survival and state of health after paediatric intensive care. *Arch Dis Child* 1995; **73**: 196-201.
43. Eiser C, Havermans T, Craft A, Kernahan J. Development of a measure to assess the perceived illness experience after treatment for cancer. *Arch Dis Child* 1995; **72**: 302-307.
44. French DJ, Christie MJ, Sowden AJ. The reproducibility of the Childhood Asthma Questionnaires: Measures of quality of life for children with asthma aged 4-16 years. *Qual Life Res* 1994; **3**: 215-224.
45. Aldenkamp AP. Psychosocial adaptation after normalisation of physical symptomatology: Cognitive complaints of children and their parents after seizure remission: In search of the time-delay model. Special Issue: Youth and disability in Holland. *Int J Adolesc Med H* 1994; **7**: 303-315.
46. Cleary PD, Morrissey G, Yver A, Oster G. The effects of rG-CSF on health-related quality of life in children with congenital agranulocytosis. *Qual Life Res* 1994; **3**: 307-315.
47. Van-Damme-Lombaerts R, Broyer M, Businger J, Baldauf C, Stocker H. A study of recombinant human erythropoietin in the treatment of anaemia of chronic renal failure in children on haemodialysis. *Pediatr Nephrol* 1994; **8**: 338-342.

48. Morris KP, Sharp J, Watson S, Coulthard MG. Non-cardiac benefits of human recombinant erythropoietin in end stage renal failure and anaemia. *Arch Dis Child* 1993; **69:** 580-586.
49. Wray J, Radley SR, Yacoub M. Effect of cardiac or heart-lung transplantation on the quality of life of the paediatric patient. *Qual Life Res* 1992; **1:** 41-46.
50. Hoppe-Hirsch E, Renier D, Lellouch TA, Sainte RC, Pierre KA, Hirsch JF. Medulloblastoma in childhood: progressive intellectual deterioration. *Childs Nerv Syst* 1990; **6:** 60-65.
51. Becker N, Lombardi P, Sidoti E, Katkin LS. Mylicon drops in the treatment of infant colic. *Clin Ther* 1988; **10:** 401-405.
52. Kekomäki M, Vilkki P, Gordin A, Salo H. Loperamide as a symptomatic treatment in pediatric surgery: A double-blind cross-over study. *Z Kinderchir Grenzgeb* 1981; **32:** 237-243.
53. Browne JP, McGee HM, O'Boyle CA. Conceptual approaches to the assessment of quality of life. *Psychology and Health* 1997; **12:** 737-751.
54. Edgerton RB. Quality of life from a longitudinal research perspective. In: Robert LS, Michael JB, eds. *Quality of life: Perspectives and issues.* Washington,DC: American Association on Mental Retardation, 1990: 149-160.
55. Scarr S, McCartney K. How people make their own environments: A theory of genotype environment effects. *Child Dev* 1983; **54:** 424-435.
56. Croyle RT, Ditto PH. Illness cognition and behavior: an experimental approach. *J Behav Med* 1990; **13:** 31-52.
57. Schmidt LR, Dlugosch GE. Entwicklungspsychologische Aspekte der Gesundheitspsychologie [German: Developmental aspects of health psychology]. *Z Klin Psychol* 1992; **21:** 36-47.
58. Fayers PM, Jones DR. Measuring and analysing quality of life in cancer clinical trials: a review. *Stat Med* 1983; **2:** 429-446.
59. Theunissen NCM, Vogels AGC, Koopman HM, et al. The proxy problem: child report versus parent report in health-related quality of life research. *Qual Life Res* 1998; **7:** 387-397.
60. Plomin R. Developmental behavioral genetics: stability and instability. In: Bornstein MH, Krasnegor NA, eds. *Stability and continuity in mental development.* Hillsdale: Lawrence Erlbaum Associates, 1989: 273-291.
61. Rosenbaum P, Cadman D, Kirpalani H. Pediatrics: assessing quality of life. In: Spilker B, ed. *Quality of life assessments in clinical trials.* New York: Raven Press, Ltd, 1990: 205-215.
62. Cleary PD. Future directions of quality of life research. In: Spilker B, ed. *Quality of life and pharmacoeconomics in clinical trails.* Philadelphia: Lippincott-Raven Publishers, 1996: 73-78.
63. Calman KC. Quality of life in cancer patients--an hypothesis. *J Med Ethics* 1984; **10:** 124-127.
64. Sprangers MA. Response-shift bias: a challenge to the assessment of patients' quality of life in cancer clinical trials. *Cancer Treat Rev* 1996; **22 Suppl A:** 55-62.
65. Croyle RT, Uretsky MB. Effects of mood on self-appraisal of health status. *Health Psychol* 1987; **6:** 239-253.
66. Orenstein DM, Kaplan RM. Measuring the quality of well-being in cystic fibrosis and lung transplantation. The importance of the area under the curve. *Chest* 1991; **100:** 1016-1018.
67. Guyatt GH, Jaeschke R, Feeny DH, Patrick DL. Measurements in clinical trials: Choosing the right approach. In: Spilker B, ed. *Quality of life and pharmacoeconomics in clinical trials.* Philadelphia: Lippincott-Raven Publishers, 1996: 41-48.
68. Landgraf JM, Abetz L. Measuring health outcomes in pediatric populations: Issues in psychometrics and application. In: Spilker B, ed. *Quality of life and pharmacoeconomics in clinical trails.* Philadelphia: Lippincott-Raven Publishers, 1996: 793-802.
69. Fairclough DL, Gelber RD. Quality of life: Statistical issues and analysis. In: Spilker B, ed. *Quality of life and pharmacoeconomics in clinical trials.* Philadelphia: Lippincott-Raven Publishers, 1996: 427-435.
70. Appelbaum MI, Burchinal MR, Terry RA. Quantitative methods and the search for continuity. In: Bornstein MH, Krasnegor NA, eds. *Stability and continuity in mental development.* Hillsdale: Lawrence Erlbaum Associates, 1989: 251-272.

71. Schoenfelder JR, Spilker B. Presentation of quality of life data. In: Spilker B, ed. *Quality of life and pharmacoeconomics in clinical trials.* Philadelphia: Lippincott-Raven Publishers, 1996: 477-487.
72. Achenbach TM, McConaughy SH, Howell CT. Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychol Bull* 1987; **101:** 213-232.
73. Koopman HM, Theunissen NCM, Vogels AGC, Zwinderman AH, Verloove-Vanhorick SP, Wit JM. Health related quality of life of children with a chronic illness: Parent versus child report. *Gedrag & Gezondheid* 1999; **27:** 118-125.
74. Tennant A. Quality of life - a measure too far? *Ann Rheum Dis* 1995; **54:** 439-440.
75. Spieth LE, Harris CV. Assessment of health-related quality of life in children and adolescents: An integrative review. *J Pediatr Psychol* 1996; **21:** 175-193.
76. Pal DK. Quality of life assessment in children: A review of conceptual and methodological issues in multidimensional health status measures. *Journal of Epidemiology and Community Health* 1996; **50:** 391-396.
77. Bullinger M, Ravens-Sieberer U. Health related QOL assessment in children: A review of the literature. *European Review of Applied Psychology/Revue Européenne de Psychologie Appliquée* 1995; **45:** 245-254.

Chapter 7

8 Changes in Quality of Life can be studied when the measurement instrument is different at different time points

N.C.M. Theunissen, J.J. Meulman, A.L. den-Ouden, H.M. Koopman, G.H.W. Verrips, S.P. Verloove-Vanhorick, J.M. Wit

8.1 Abstract

This chapter presents a strategy for the analysis of longitudinal Quality of Life (QoL) data that suffer from differences in measurement instruments over time. The strategy was applied to a set of longitudinal data from a cohort of 688 preterm born children. Health status data, differently defined at 5, 9 and 10 years of age, were prepared for longitudinal analyses with qualitative and quantitative item selection. Expert ratings fitted the data into physical, psychological and social health status domains. Principal component analyses (PCA) was used to match the data between measurements. Longitudinal PCAs were performed using the matched health status data. The impact of background variables such as gender and birth weight on health status changes was studied. This strategy aimed at reconstructing and combining an imperfect data set, provided valuable information about the development of health status in preterm born children. *Key words: health status, quality of life, preterm birth, data reconstruction, principal component analysis, categorical data, ordinal data, longitudinal studies, follow-up.*

8.2 Introduction

A longitudinal design is a useful tool for studying the late effects of perinatal factors such as, for instance, preterm birth. Unfortunately, this design has some drawbacks of which drop outs,[1,2] selection bias and the lengthy and expensive commitment are the most commonly known. Less emphasised is the necessity to use a consecutive series of age specific instruments, which may be difficult to compare between ages.[3] Another drawback of longitudinal studies are the changes in scientific ideas during the research period. For instance, outcome studies in children until the 80s primarily emphasised the physical functioning approach. More recently, a broader concept of quality of life has been commonly used in which physical, as well as

psychological and social functioning are considered essential. Furthermore, standardised quality of life instruments for children are relatively new and even a standardised instrument can change over time. Consequently, the measurement instruments may differ at different time points, which makes it difficult to analyse longitudinal changes. In this paper an analysis strategy is presented that can be used to overcome these obstacles. The strategy was applied to a set of longitudinal data collected from 688 children, preterm born in 1983 and enrolled in the follow-up program "Project On Preterm and Small for gestational age infants (POPS)".[4-6] Originally, the project was designed to investigate the relation of prenatal and perinatal factors with mortality and morbidity. The survival rate of these children has improved and therefore scientific interest has shifted from mortality and morbidity towards health status in general, described as a combination of physical-, psychological as well as social functioning. The health status domain physical functioning is in the literature divided into sub-domains like physical complaints, motor functioning and autonomy. In psychological functioning cognitive and emotional sub-domains are commonly used. Social functioning could be directed towards peers or adults.[7-9] Generally, health status should be the assessment by a person of his or her own health functioning. If it is not possible to obtain information from the person him/herself, a proxy can be used. Next to the child itself, the parent is the most preferable informant of the child's health status.[10] Although much health status information obtained from the parents was available, the early part of the POPS project was designed to fit longitudinal analyses using the physician and not the parent as main informant. As a result the content and the form of the assessment differed between points of measurement.

8.3 Method and results

8.3.1 Qualitative selection of items

The administration, the amount and the content of health status questions obtained from the parents differed between measurement points. At 5 years of age, a home visit by a specially trained paediatrician was preceded by a questionnaire for the parent. If necessary, the questionnaire was completed by the parent during the paediatrician's visit. At age 9, a postal questionnaire was used assessing mainly school achievement. At age 10 a new postal questionnaire was used, assessing all health status domains.

Assignment of the outcome variables at age 5, 9 and 10 to health status sub-domains has been performed as follows. First, the items were pre-selected if they applied to health status sub-domains according to independent rating by five experts: two paediatricians (A.L.d.O, S.P.V.V.) and three developmental psychologists (N.C.M.T, H.M.K, G.H.W.V.). The rating resulted in five independent opinions about every item. Second, an item was selected if at least three out of five experts agreed the item belonged to a health status sub-domain. If fewer than three out of five experts agreed about an item, additional selection was performed by a developmental psychologist and a paediatrician (N.C.M.T, A.L.d.O.). The final result was again presented to the forum of experts. Final decisions regarding item selection were made by group consensus. As a result, a set of 34 variables at 5 years of age, a set of 7 variables at 9 years of age and a set of 47 variables at 10 years of age became available (see Appendix at page 121).

8.3.2 Quantitative selection of items

A quantitative selection was necessary to match the qualitatively selected health status variables at different points of measurement. The data consist of ordinal categorical variables, and it could not be assumed a priori that the distances between the different categories were equal. Therefore a generalisation of classical principal component analysis (PCA) was used, called PRINCALS,[11-13] which is a PCA for categorical data. PRINCALS simultaneously fits the principal components model and finds an optimal scaling for the categories. Optimal scaling means that the average proportion of total variance accounted for is as large as possible, given the PCA model. The analyses require a data matrix with subjects in the rows and variables at age 5, 9 and at 10 in the columns.

First, all health status variables were included in a categorical PCA analysis to see if the variables would group together by health status sub-domains, regardless of time of measurement. This analysis accounted for a proportion of total variance of 20% (dimension 1 = 14%, dimension 2 = 6%), which is equivalent to a Cronbach's alpha of .93 for dimension 1 and an alpha of .82 for dimension 2 (the number of variables is taken into account in the formula).[14] The relationships between the variables are displayed graphically in Figure 1 and 2 at the next page.

Figure 1. Relationships between HS variables at all different times of measurement, resulting from the quantitative selection

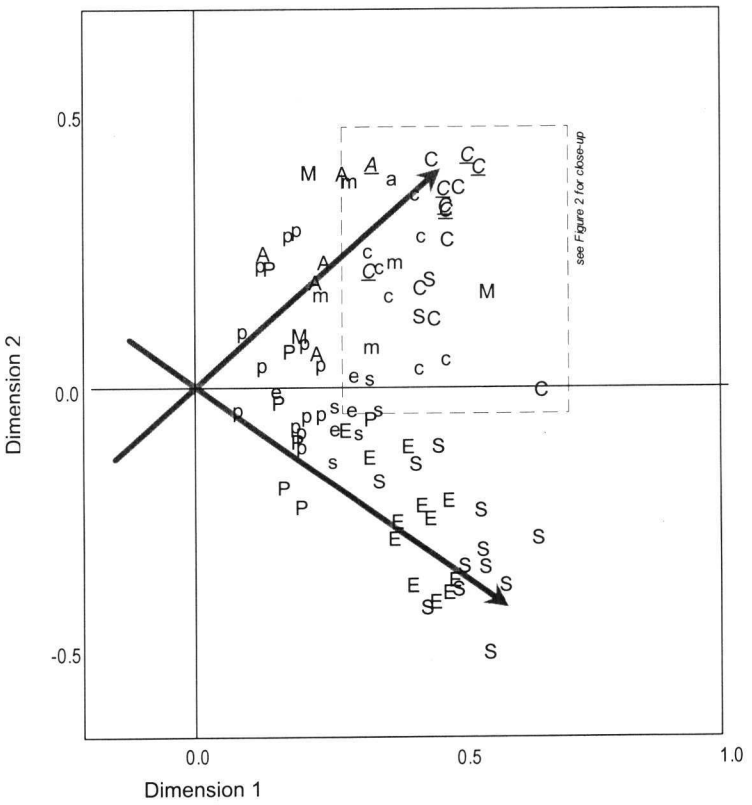
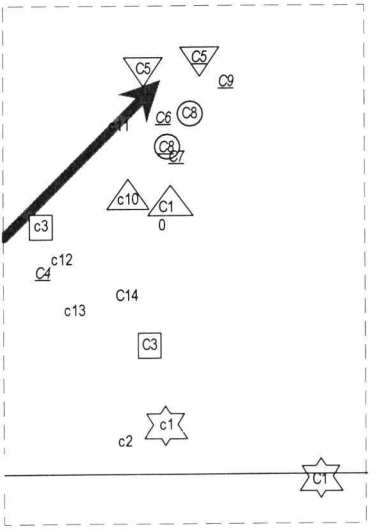


Figure 2. A close up from Figure 1, displaying the cognitive functioning items



Legend Figure 1:
no capitals = items at 5 years of age ; Italic *CAPITALS* underlined = items at 9 years of age; CAPITALS = items at 10 years of age ; C = cognitive functioning, S = social functioning, P = physical complaints, M = motor functioning, A = autonomy, E = emotional functioning (See text for more clarification).

Legend Figure 2:
The numbers in the figure correspond with the numbers in the Appendix. The circles, triangles, stars or squares mark the partner variables.

In Figure 1, the positions of the variables and their distance to the centre of the graph (zero point) are determined by their correlations with the principal components, the so-called component loadings. The distances between the variables represent the correlations between them. As can be seen in, there is a very strong first component (the horizontal axis): all variables point in the positive direction. In the second dimension, the analysis partitioned the variables into two clusters: motor (M), autonomy (A) and cognitive (C) variables on the top, emotional (E) and social (S) variables at the bottom, and physical (P) in between. The clustering of motor functioning with autonomy, and emotional with social functioning is considered an obvious one. Apparently, the children with motor-autonomy problems experience cognitive problems as well.

A more important result of this analysis is that the variables of each health status sub-domain are grouped together, irrespective of time of assessment. Therefore, the variables in these sub-domains could be matched at different points in time. Eighteen combinations of variables were inspected that a priori were considered to be good matches according to their meaning. For instance, the variable "cannot concentrate" at age 5 (c1) was almost identical in formulation to a similar variable at age 10 (C1) Their position in relation to the centre of the graph is alike and the distance between these variables is small. This is shown in Figure 2, a close-up of Figure 1. At age 5 another cognitive variable was assessed as being analogous in meaning to the "cannot concentrate" mentioned previously, that is "problems with concentration" (c2). In Figure 2 it can be seen that the distance between these variables at age 5 is very small. As we already found a match on this subject, this last variable could be removed from further analyses. Other variables, like for instance "received extra assistance at school" (*C4*) , were not repeated in time which made matching impossible. For some variables, matching at different time points was debatable and needed further analyses to clarify their usefulness. Therefore, a second PCA was performed including the 18 combinations mentioned above, excluding variables analogous in meaning to these 18 combinations, and excluding health status aspects that were not repeated in time, but including doubtful cases.

The second analysis accounted for a proportion of total variance of 22% (Cronbach's alphas: dimension 1 = .89, dimension 2 = .72). The same two main clusters were found as in the first analysis, confirming that the selected variables still covered the same health status aspects. In addition to the 18 combinations of variables already established, one new combination was found. *A priori* the two variables did not

have the same content, but it could be argued that they belonged together: "stiff movements" at age 5 and "walking" at age 10 both represents moving problems. A third PCA was performed with the 18 old and the one new combination, which accounted for a proportion of total variance of 24% (Cronbach's alphas: dimension 1 = .86, dimension 2 = .65), with the same two bundles found as in the first analysis.

Given that the categories of two variables to be combined were not comparable, the categories had to be merged in a rational way. Therefore the recoding was done by using the optimal quantifications for the separate variables. Table 1 shows how categories were merged. As only three variables at age 9 could be combined with partner variables at age 10, the variables at age 9 were excluded from further analyses leaving 16 variables.

Table 1. Matching categories by using optimal scaling to obtain a new longitudinal variable: "moving problems" and "speech problems"

Stiff movements (at 5 y.)		*Walking (at 10 y.)*		*Moving problems*
original ordinal categories	quantification	original ordinal categories	quantification	new ordinal categories combined by optimal scaling
1= rarely	-0.48	1= very good	-0.35	1
2= occasionally	0.33			2
3= rather often	1.44			3
		2= dubious or barely	2.77	4
4= almost always	3.28			5

Others understand this child (at 5 y.)		*Other people understand what child says (at 10 y.)*		*Speech problems*
original ordinal categories	quantification	original ordinal categories	quantification	new ordinal categories combined by optimal scaling
1= almost always	-0.43	1= very good	-0.44	1
2= most of time	0.65			2
		2= dubious or barely	2.24	3
3= rarely or not	3.71			4

8.3.3 Longitudinal analysis: relationships between health status variables over time

To study the health status development, the data were reshaped as follows. The columns of the data matrix contained the 16 selected variable combinations, and the rows the observations on the 688 children at time point 5 and 10, placed below each other. Thus, the data matrix contained 1376 rows, each subject appearing two times. The marginal frequencies of the categories were unevenly distributed and the majority of the cases were in the lower category, which indicates few problems.

In addition to this data matrix, categorical interaction variables were created combining the information about time point and each of the background variables gender, birth weight and multiple birth. Next, four categorical PCAs were performed. The first analysis was based on the original 16 health status variables only, in each of the subsequent analyses one of the three supplementary time × background variables was included. Because there were two points in time and, for instance, two sexes, the new time × gender variable contains four possible categories. The 16 health status variables were given an ordinal optimal scaling level, while the time × background variables were treated as nominal variables because their categories are unordered. The average proportion of variance accounted for by the categorical PCAs excluding and including time × background variables was 31%. Dimension 1 accounts for 21% (alpha=.76) and dimension 2 for 10 % (alpha=.46). To determine the resemblance between the PCAs excluding and including the time × background variables, correlations were computed between the subjects' scores (component scores) that were obtained from the four analyses. In Table 2, correlations between the subject scores of the PCA models are given, excluding and including time × background variables. Correlations are extremely high, indicating that adding a time × background variable to the health status variables does not change the PCA results substantially. Less then perfect correlations were found on Dimension 2 indicating differentiation of the time × background variable on this dimension.

Table 2. Correlations between the subject scores of the PCA models including and excluding time × birth-factor variables

	excluding time	time × gender	time × birth	time × multiple
excluding time × birth-factors		.871	.864	.865
time × gender	.996		.990	.985
time × birth weight	.997	.998		.995
time × multiple birth	.997	.998	.999	

Note: under the diagonal the correlations on Dimension 1 are given and above the diagonal correlations on Dimension 2.

The health status variables included in the analyses are ordinal categorical ones. Optimal scaling of a variable in the PRINCALS procedure amounts to assigning interval values to ordinal categories (=category quantifications). The category quantifications of a variable depend on the time × background variables that were included in the PCA models, therefore the quantification may differ from analysis to analysis. However, the correlations between the PCA models are extremely high (see Table 2) and therefore the differences between the quantifications are negligible. In this paper we use consequently the PCA including the time × gender variable as an example.

At the y-axis in Figure 3, category quantifications are displayed for ordinal variables having more than two categories. The zero score represents the mean score of the total sample over two time points. The x-axis gives the original category numbers. As can be seen in the figure, the categories of the presented variables are ordinal and in four out of six almost at interval level. In all health status variables the lowest category represents the absence of a problem (see the Appendix at the end of this chapter for the category contents).

Figure 3. Category optimal scaling scores of variables with more than two categories, resulting from the longitudinal analysis

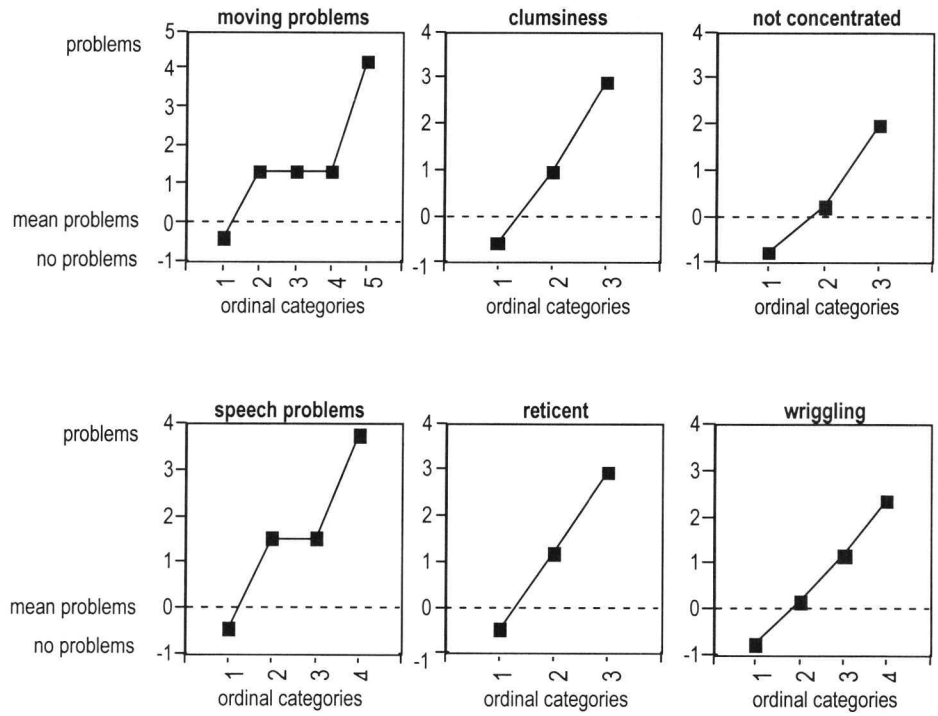
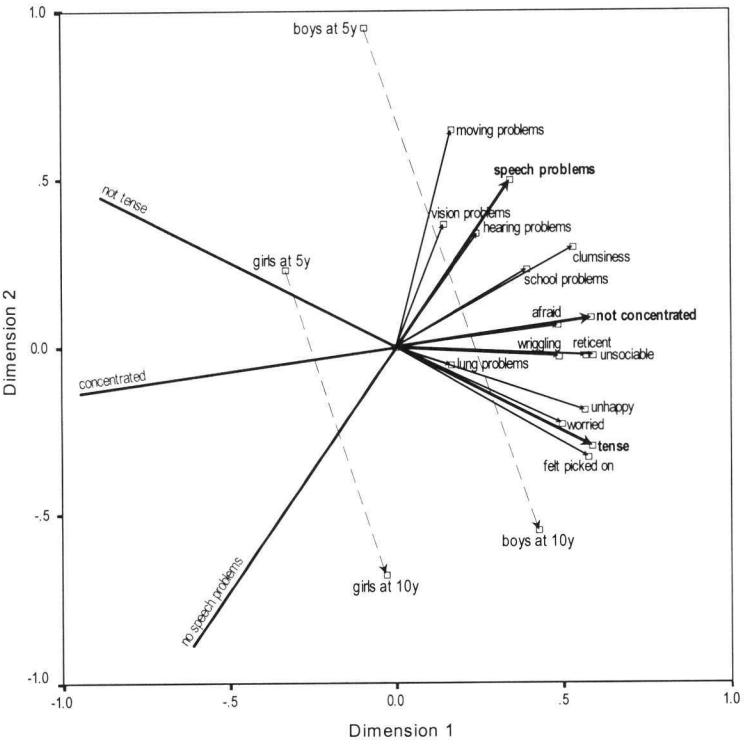


Figure 4. Relationships between HS variables in combination with gender displayed over time



Only the lowest categories have scores below zero, and the majority of the other categories, representing health status problems, obtained scores higher than zero. It can be concluded that a score of zero represents minor problems, which in turn corresponds with the fact that the study group on the whole had minor health status problems on average.

The health status variables are depicted graphically as vectors (arrows) in Figure 4. Again, the zero-point in the figure represents the mean component score of the sample. All variables have a positive correlation with the first dimension (x-axis) resulting in a Cronbach's alpha of .76. Since in these variables, high category scores mean more health status problems, the underlying factor of Dimension 1 is health status problems. "Respiratory problems" does not fit very well (short vector) and seems to be not important as a health status indicator. The lower score ends of the vectors (at the left-hand side of Figure 4) indicate the absence of problems such as "not tense" and "no speech problems". Not shown in this figure is that seventeen percent of the subjects have

their component scores (=object scores) in the area spanned by the no-problem ends of the vectors at age 5 but not at age 10. This indicates no health status problems at 5 yet health status problems at 10 years of age. Sixteen percent had no health status problems at age 10 although they had health status problems at age 5. Thirty-five percent of the subjects had health status problems at both ages and 32% of the sample had no health status problems at both ages. To investigate the nature of the problems at 5 or 10 years of age, the second dimension was inspected.

The second dimension (vertical axis), with a total Cronbach's alpha of only .46, separates the health status variables. It distinguishes between physical and motor problems in the upper part of the figure, and emotional en social problems in the lower part. Cognitive problems are positioned in between. Smaller angles between the vectors represent higher correlations between the variables. The bundle of "speech problems", "moving problems", "vision problems" and "hearing problems" has hardly any correlation with the bundle "tense", "unhappy", "worried" and "felt-picked-on". The first bundle can be interpreted as problems in basic functioning and the last as negative moods. Between these bundles is a bundle close to the horizontal axis with "school problems", "not concentrated", "unsociable", "wriggling", "closed", "afraid", and "clumsiness". Together they indicate concentration problems with both cognitive and behavioural components.

Usually, Cronbach's alpha between variables is computed on the basis of the original variables. In our case, the variables were optimally transformed. In the PCA model, a transformed variable is fitted by the sum of the component scores weighted by the component loading in each dimension. These transformed scores are calculated by multiplying the subject's component score by the variable's component loading for dimension 1 and dimension 2 consecutively, and summing the resulting figures. As we are looking at the representation of the variables in two dimensions, Cronbach's alphas of the three variable bundles were computed on the basis of the transformed variables. Basic functioning problems had a Cronbach's alpha of .97, concentration problems had an alpha of .99, and negative moods had an alpha of .996 .

From each bundle of variables a key variable was selected, to represent the other variables in the bundle in further analyses. Selection was not based on the highest loading, but on the representativeness for the other variables. This representativeness was based both on content and on a central position of the key variable in the bundle. However, choosing another variable from the bundle as the key variable will give comparable results. From the basic functioning bundle we selected 'speech problems', from the negative moods bundle we selected 'tense', and from the concentration bundle we selected 'not concentrated'.

8.3.4 Relationships between background variables and health status development

In Figure 4 the development of children over time is displayed by putting the time × gender variable next to all health status variables. The dotted lines in the figure shows a large time effect along the second dimension: Both boys and girls had more basic functioning problems at age 5 and more negative moods at age 10. This time effect was found for all background variables. In other words, scores on dimension 2 were dominated by time.

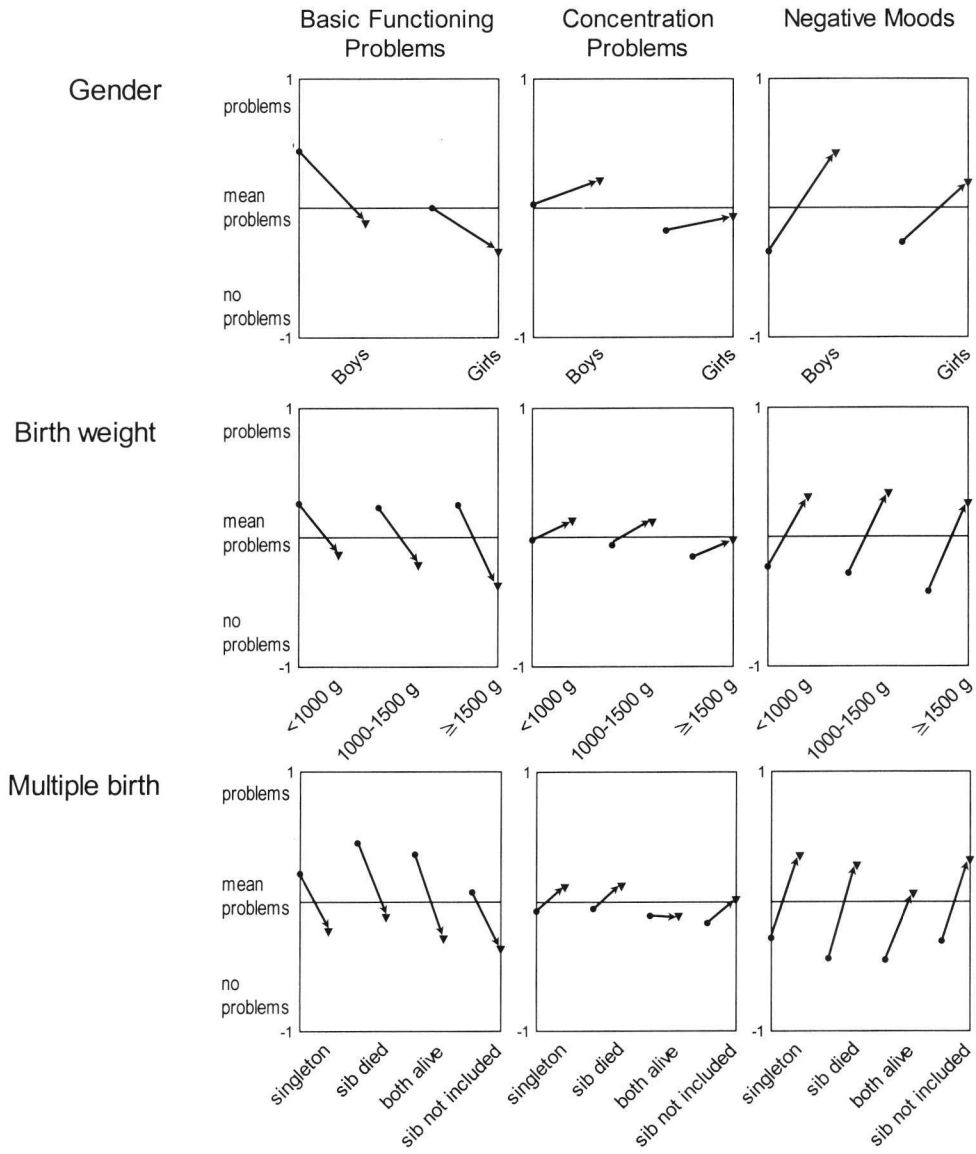
The relations between gender, time, and the key variables can be derived from Figure 4, and are more clearly presented in the Figures 5a to 5c (see next page). In these figures, for each background variable group the average of the key variables' optimal scaling scores are given. Differences between groups in arrow length shows group effects, total length shows change over time and parallel lines indicate no interaction between group and time. Differences with respect to key variables were tested with Repeated Measures ANOVA, testing the sources of variation between groups and within groups between age 5 and 10 (see Table 3).

Table 3. Repeated measures analysis of variance between birth-factors and age of measurement with respect to the three key variables

		Basic functioning problems *'speech problems'* F	Concentration problems *'not concentrated'* F	Negative moods *'tensed'* F
Gender	between subjects (groups) [a]	88.23 ***	40.14 ***	4.61 *
	within subjects (age)	450.01 ***	37.73 ***	608.41 ***
	interaction	25.11 ***	2.82	37.77 ***
Birth weight, g.	between subjects (groups)	1.94	3.35 *	3.11 *
	within subjects (age)	338.09 ***	28.28 ***	406.25 ***
	interaction	6.33 **	0.43	1.50
Multiple birth	between subjects (groups)	2.40	1.90	5.28 ***
	within subjects (age)	171.72 ***	9.50 **	183.53 ***
	interaction	3.18 *	2.39	1.00

[a] differences between measurements at 5 and 10 years of age; * p≤.05 ** p≤.01 *** p≤.001

Figure 5. Background variables at 5 and 10 years of age with respect to basic functioning problems, concentration problems and negative moods
5a = Gender, 5b = Birth weight, 5c = Multiple birth
Legend: ● = 5 years of age, ▼ = 10 years of age



Overall a large time effect is shown in all background variable groups. The amount of basic functioning problems tended to decrease in time, negative moods increased and concentration problems increased slightly in all background variable groups.

Boys have more problems in basic functioning and more concentration problems than girls (see Figure 5a). At 5 years of age, boys had fewer negative moods than girls. At 10 years of age, boys had slightly more negative moods than girls. The group differences between boys and girls as well as time differences between 5 and 10 years of age are significant. In addition, group × time interactions are significant in basic functioning problems and negative moods. It can be concluded that boys had more problems than girls did at both points of measurement, although the nature of the problems were the same for both sexes. Differences between birth weight groups were significant for negative mood and concentration problems only (see Figure 5b). Lower birth weight related to more HS problems. Problems in basic functioning decreased and negative mood increased in time, for each single or multiple birth group (see Figure 5c). Concentration problems gave a mixed picture with small but significant differences (see Table 3). The group differences were significant with respect to negative mood only: Children with a twin included in the study had the least problems.

8.4 Discussion

Changes in health status could be reconstructed in spite of the initial differences in measurement instruments at different time points. The qualitative item selection we used to start with, was inspired by Saigal et al.[15] in re-using data originally not organised to measure health status. They applied the Multi-attribute Health Status (MAHS) classification system retrospectively, to prospectively collected clinical and psychometric data. This application was performed to compare health status of children with extremely low birth weight with a control group. However, Saigal et al. recoded the original items into MAHS items by using computer algorithms entirely based on expert ratings. In the present study expert ratings were used to select but not to recode items. Interpretation of which original score represents which problem level is entirely empirically based. Consequently, it was possible to stay close to the original data. After matching the health status variables between points of measurement, longitudinal changes could be made visible. Although there are many longitudinal studies of preterm children,[16-28] only some of them include a physical, psychological as well as a social

health status domain.[17,24,26,28] Instead of studying the changes in health status sub-domains one by one as previously done in these studies, the present method allows for new combinations of domains. Consequently, it is possible to pinpoint the problems that are relevant for the development of preterm born children in particular.

In addition, half of the longitudinal studies on preterm born children we found, reported no longitudinal health status analyses,[26] or restricted analyses to correlations between assessments and did not report changes in health status.[17,22-25] This was probably caused by the use of standard instruments that were age specific. For instance, Caputo et al.[17] used age specific instruments to assess development in children with low birth weight. The Cattell Infant Intelligence Scale and the Gesell Developmental Schedules were used at 1 year of age. The Wechsler Intelligence Scale for Children-Revised and the Visual Motor Gestalt Test were used at 7 to 9 years of age. Although from both points of measurement a developmental quotient (DQ) was obtained, the content of these DQs was probably considered to be too age specific. By using the present reconstruction technique, it might have been possible to study changes in development and not merely correlations. In the other half of the longitudinal studies we inspected, changes were investigated,[16,18-21,27,28] but only one of them included all three health status domains[28] and reported changes similar to ours.

The longitudinal results of the present study may be an artefact of the reconstruction method we used. Some variables that we considered to measure the same property perhaps did not. This may have caused both underestimation and overestimation of longitudinal changes in all health status variables. Also, under- or overestimation can be influenced by the health status assessment itself. At age 5 the parents completed the questionnaire together with the paediatrician "if necessary". At age 10 the paediatrician was not present. The parents may have been biased at age 5 by the stress put on basic functioning problems by a paediatrician, since clinicians emphasise other HS problems than parents do. Nevertheless, the main problem areas we found and the order in which they appeared -- basic functioning, concentration and negative moods -- are considered to be clinically and psychologically relevant. More paediatric details about the health status changes of the POPS cohort are described elsewhere.[29] It is of interest to report here that the groups which appear at risk according to the present method, are known risk groups; this can be interpreted as a validation of the study's method.

In conclusion, longitudinal studies have to cope with the consequences of using age specific instruments and with changes in scientific ideas during the research period. In this paper we described a strategy to overcome the problem of using different measurement instruments at different time points. At the same time we stayed as close to the original data as possible. The "common knowledge" results we found in preterm born children gave confidence in the method used. Although prospective research will be needed to judge the validity of the findings, the method presented allows an imperfect data set to provide valuable information about changes in health status or quality of life.

Acknowledgements
We would like to thank all the previous participants from the Collaborative Project on Preterm and Small for Gestational Age (POPS) Infant in the Netherlands. We are indebted to the parents for their participation.

8.5 Reference List

1. Raboud JM, Singer J, Thorne A, Schechter MT, Shafran SD. Estimating the effect of treatment on quality of life in the presence of missing data due to drop-out and death. *Qual Life Res* 1998; **7**: 487-494.
2. Zwinderman AH. Statistical analysis of longitudinal quality of life data with missing measurements. *Qual Life Res* 1992; **1**: 219-224.
3. Wadsworth ME, Rodgers B, Wadsworth MEJ. Longterm follow-up studies; a critical overview. *Rev Epidemiol Sante Publique* 1989; **37**: 533-547.
4. Hille ETM, den-Ouden L, Bauer L, Van-den-Oudenrijn C, Brand R, Verloove-Vanhorick SP. School performance at nine years of age in very premature and very low birth weight infants: Perinatal risk factors and predictors at five years of age. *J Pediatr* 1994; **125**: 426-434.
5. Veen S, Ens-Dokkum MH, Schreuder AM, Verloove-Vanhorick SP, Brand R, Ruys JH. Impairments, disabilities, and handicaps of very preterm and very-low-birthweight infants at five years of age. *Lancet* 1991; **338**: 33-36.
6. Verloove-Vanhorick SP, Verwey RA, Ebeling MCA, Brand R, Ruys JH. Mortality in very preterm and very low birth weight infants according to place of birth and level of care: Results of a national collaborative survey of preterm and very low birth weight Infants in the Netherlands. *Pediatr* 1988; **81**: 404-411.
7. Bullinger M, Ravens-Sieberer U. Stand der Forschung zur gesundheitsbezogenen Lebensqualität von Kindern: eine Literaturanalyse [German: State of the art of health-related quality of life research in children: a literature review]. *Präv Rehab* 1995; **7**: 106-121.
8. Grunberg SM, Groshen S, Steingass S, Zaretsky S, Meyerowitz B. Comparison of conditional quality of life terminology and visual analogue scale measurements. *Qual Life Res* 1996; **5**: 65-72.
9. Vogels T, Verrips GH, Verloove-Vanhorick SP, et al. Measuring health-related quality of life in children: the development of the TACQOL-Parent Form. *Qual Life Res* 1998; **7**: 457-465.
10. Theunissen NCM, Vogels AGC, Koopman HM, et al. The proxy problem: child report versus parent report in health-related quality of life research. *Qual Life Res* 1998; **7**: 387-397.

11. van-der-Ham T, Meulman JJ, van-Strien DC, van-Engeland H. Empirically based subgrouping of eating disorders in adolescents: A longitudinal perspective. *Brit J Psychiat* 1997; **170:** 363-368.
12. Kerkhof AJFM, van-der-Wal J, Hengeveld MW. Typology of persons who attempted suicide with predictive value for repetition: a prospective cohort study. In: Möller HJ, Schmidtke A, Welz R, eds. *Current issues of suicidology.* Berlin: Springer-Verlag, 1988: 194-203.
13. *SPSS Categories 8.0.* Chicago: SPSS Inc., 1998.
14. Heiser WJ, Meulman JJ. Homogeneity analysis: Exploring the distribution of variables and their nonlinear relationships. In: Greenacre M, Blasius J, eds. *Correspondence analysis in the social sciences: Recent developments and applications.* London, San Diego, New York, Boston, Sydney, Tokyo, Toronto: Academic Press, Harcourt Brace & Company, 1994: 179-209.
15. Saigal S, Rosenbaum P, Stoskopf B, et al. Comprehensive assessment of the health status of extremely low birth weight children at eight years of age: Comparison with a reference group. *J Pediatr* 1994; **125:** 411-417.
16. Bosch JD, Veerman JW, Samson JF, Smith F. Ex-neonatale couveusepatieentjes vijf en tien jaar later. Een follow-up-onderzoek naar mogelijke gedragsproblemen [Dutch: Ex-neonatal intensive-care-unit (NICU) patients five and ten years later: A follow-up study of risk for behavioral problems]. *Kind en Adolescent* 1995; **16:** 1-15.
17. Caputo DV, Goldstein KM, Taub HB. Neonatal compromise and later psychological development: A 10-year longitudinal study. In: Friedman SL, Sigman M, eds. *Preterm birth and psychological development.* New York, etc: Academic Press, 1981: 353-385.
18. Fedrizzi E, Inverno M, Botteon G, Anderloni A, Filippini G, Farinotti M. The cognitive development of children born preterm and affected by spastic diplegia. *Brain Dev* 1993; **15:** 428-432.
19. Gross SJ, Slagle TA, D'Eugenio DB, Mettelman BB. Impact of a matched term control group on interpretation of developmental performance in preterm infants. *Pediatr* 1992; **90:** 681-687.
20. Kalmar M. The course of intellectual development in preterm and fullterm children: An 8-year longitudinal study. *International Journal of Behavioral Development* 1996; **19:** 491-516.
21. Largo RH, Molinari L, Kundu S, Lipp A, Duc G. Intellectual outcome, speech and school performance in high risk preterm children with birth weight appropriate for gestational age. *Eur J Pediatr* 1990; **149:** 845-850.
22. Largo RH, Graf S, Kundu S, Hunziker U, Molinari L. Predicting developmental outcome at school age from infant tests of normal, at-risk and retarded infants. *Dev Med Child Neurol* 1990; **32:** 30-45.
23. Menyuk P, Chesnick M, Liebergott JW, Korngold B. Predicting reading problems in at-risk children. *Journal of Speech and Hearing Research* 1991; **34:** 893-903.
24. Portnoy S, Callias M, Wolke D, Gamsu H. Five-year follow-up study of extremely low-birthweight infants. *Dev Med Child Neurol* 1988; **30:** 590-598.
25. Saigal S, Szatmari P, Rosenbaum PL. Can learning disabilities in children who were extremely low birth weight be identified at school entry? *J Dev Behav Pediatr* 1992; **13:** 356-362.
26. Saylor CF, Casto G, Huntington L. Predictors of developmental outcomes for medically fragile early intervention participants. *J Pediatr Psychol* 1996; **21:** 869-887.
27. Schothorst PF, van-Engeland H. Long-term behavioral sequelae of prematurity. *J Am Acad Child Adolesc Psychiatry* 1996; **35:** 175-183.
28. Silva PA, McGee R, Williams S. A longitudinal study of the intelligence and behavior of preterm and small for gestational age children. *J Dev Behav Pediatr* 1984; **5:** 1-5.
29. Theunissen NCM, den-Ouden AL, Meulman JJ, Koopman HM, Verloove-Vanhorick SP, Wit JM. Health status development in a cohort of preterm children. *submitted* 1999.

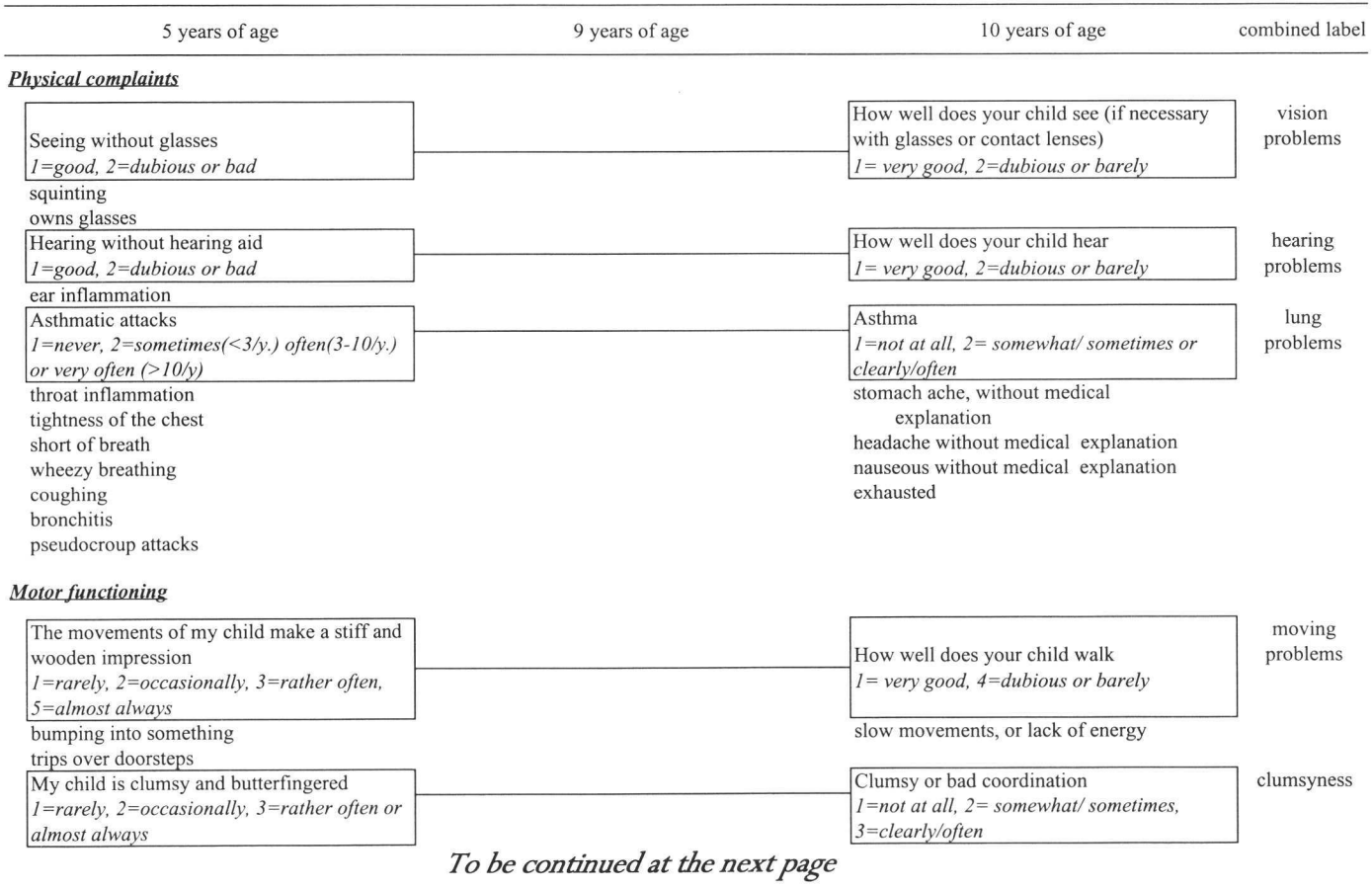
Appendix. The qualitative selection result of health status variables at 5, 9 and 10 years of age (translated from the Dutch original)

5 years of age	9 years of age	10 years of age	combined label
Cognitive functioning			
c1: My child cannot concentrate, cannot focus on something for a long time *1=not applicable, 2= somewhat, 3=clearly/often*		**C1:** Cannot concentrate, cannot focus on something for a long time *1=not at all, 2= somewhat/ sometimes, 3=clearly/often*	not concentrated
c2: problems with concentration			
c3: My child gets low grades at school *1=no, 2= somewhat or clearly/often*		**C3:** Low grades at school *1=not at all, 2= somewhat/ sometimes or clearly/often*	school problems
	C4: received extra assistance at school		
	C5: School results: language *1=good, 2=satisfactory, 3=weak or unsatisfactory*	**C5:** School results: language or Dutch *1=good, 2=satisfactory, 3=weak or unsatisfactory*	language problems
	C6: school:reading		
	C7: school:writing		
	C8: School results: math *1=good, 2=satisfactory, 3=weak or unsatisfactory*	**C8:** School results: math or rythmetics *1=good, 2=satisfactory, 3=weak or unsatisfactory*	math problems
	C9: school: other subjects		
c10: Children are often better understood by their parents than by other people. Others understand my child *1=almost always 2=most of time, 4=rarely or not*		**C10:** How well does other people understand what your child says *1= very good, 3=dubious or barely*	speech problems
c11: talking		**C14:** how well does the child understand others?	
c12: finding words			
c13: intelligibility/audibility			

To be continued at the next page

Continuation of the Appendix

5 years of age	9 years of age	10 years of age	combined label
Social functioning			
My child feels that others pick on him/her *1=not applicable, 2= somewhat or clearly/often*		Feels that others pick on him/her *1=not at all, 2= somewhat/ sometimes or clearly/often* be ragged a lot	felt picked on
My child can't get on well with other children *1=not applicable, 2= somewhat or clearly/often*		Can't get on well with other children *1=not at all, 2= somewhat/ sometimes or clearly/often* disrupts other children's games activities with friends getting along with other children others don't like him/her how many really good friends?	unsociable
My child is reticent, others do not know what is on the child's mind *1=not applicable, 2= some what, 3=clearly/often*		Reticent, others do not know what is on the child's mind *1=not at all, 2= somewhat/ sometimes, 3=clearly/often*	reticent
My child is wriggling constantly *1=rarely, 2=occasionally, 3=rather often, 4=almost always* fiddling all the time doesn't do what it has to do		Makes often restless movements with hands or feet or wiggles in chair *1=not at all, 2= somewhat/ sometimes, 3=rather often 4= very often* does not respond disobedient at home behavior towards parents objects or arguments a lot	wriggling



To be continued at the next page

Continuation of the Appendix

5 years of age	9 years of age	10 years of age	combined label
Autonomy			
	School results: gymnastics *1=good, 2=satisfactory, 3=weak or unsatisfactory*	Sports achievement compared with peers *1=above average, 2=average, 3=beneath average*	sports problems
problems with laces and buttons		medical reasons for not joining sports or gymnastics at school	
		how well is the child performing hobbies in comparison with peers	
		wet one's pants	
		wet one's bed	
Emotional functioning			
My child is unhappy, sad, depressed *1=not applicable, 2= somewhat or clearly/often*		Unhappy, sad, depressed *1=not at all, 2= somewhat/ sometimes or clearly/often*	unhappy
My child is worried *1=not applicable, 2= somewhat or clearly/often*		Is worried *1=not at all, 2= somewhat/ sometimes or clearly/often*	worried
My child is fearful or too afraid *1=not applicable, 2= some what or clearly/often*		Is fearful or too afraid *1=not at all, 2= somewhat/ sometimes or clearly/often*	afraid
		shy	
My child is nervous, tense *1=not applicable, 2= some what or clearly/often*		Nervous or tensed *1=not at all, 2= somewhat/ sometimes or clearly/often*	tense
		sulks and pouts	
		fights a lot	
		destroys its own belongings	
		cruel or mean to others	
		physically attacks others	
		outbursts of anger	
		quickly jealous	

9 Health status development in a cohort of preterm children

N.C.M. Theunissen, A.L. den-Ouden, J.J. Meulman, H.M. Koopman,
S.P. Verloove-Vanhorick, J.M. Wit

9.1 Abstract

Objective: The complications that accompany a preterm birth can have long-term effects. The Health Status (HS) at the age of 5 and 10 years in a cohort of children born preterm is studied to determine the impact of preterm birth on HS development.

Sample: 688 children, born in 1983 with a gestational age of less than 32 weeks and a birth weight of less than 1500 grams. Design: Prospectively collected HS variables, obtained from the parents, were analysed in a longitudinal perspective, using principal component analyses. Results: One third of the sample had minor to severe HS problems at both ages of measurement. One third had problems on one assessment only. The remainder of the sample had no HS problems at either age. The analyses grouped the HS variables into three combinations. Problems in Basic Functioning, such as mobility or speech, decrease with age. Negative Moods substantially increase, and Concentration Problems increase slightly. Specifically at risk were preterm born children with handicaps, boys, and Small for Gestational Age (SGA) children.

Conclusion: According to the parents, one third of the cohort had no HS problems at either age. The pattern of HS problems of the preterm born children changed between 5 and 10 years of age. *Key words: health status, preterm birth, premature birth, very low birth weight, longitudinal studies, follow-up.*

9.2 Introduction

Since neonatal intensive care was introduced in the 1960s, survival rate of Very Preterm (VPT: < 32 weeks) and Very Low Birth Weight (VLBW: < 1500 g) infants has substantially improved. The prevalence of major disabilities and handicaps, however, has remained stable.[1-4] It is well documented that VPT or VLBW birth can result in psychological, social, and physical problems.[3,5-14] The World Health Organization in 1948 defined health as a combination of physical-, psychological as well as social functioning, and Health Status (HS) is seen as a combination of these three domains.[15-17] The HS domains are most often divided into sub-domains: physical

functioning into physical complaints and motor functioning or daily functioning, psychological functioning into cognitive functioning and emotional functioning, and social functioning is functioning towards peers or adults.[15-17] However, in studies of HS in preterm born children the use of a combination of these domains is rare. Most studies have focussed primarily on the relation between perinatal factors on one hand, and a HS outcome in a specific domain at a certain age in childhood on the other. It is of interest to clinicians and parents to know whether these problems in childhood are stable or in which HS domain problems will decrease or increase.

The aim of this article is to study in a large national cohort of VPT and VLBW children the relations between the HS domains, the changes with age in HS, and the impact of factors such as gestational age on changes in HS.

Table 1. Distribution of clinical and demographic variables

Perinatal factor	Definition	N	(%)
Gestational age (in weeks)	< 28	62	(9)
	28-30	159	(23)
	30-32	275	(40)
	> 32	190	(28)
Percentile of birth weight against gestational age[a]	≤ 2.3 (VSGA)	128	(19)
	< 10 (SGA)	122	(18)
	> 10 (AGA)	437	(64)
Gender	female	353	(51)
Handicaps	no handicap	550	(80)
	motor handicap	22	(3)
	sensory handicap	8	(1)
	cognitive handicap	91	(13)
	combination of handicaps	17	(3)
SES [20]	low	205	(37)
	medium	217	(39)
	high	137	(24)
Maternal age at enrollment	< 25	200	(29)
	25-29	253	(37)
	> 29	235	(34)

[a] controlled for sex and parity[21]; VSGA= very small for gestational age, SGA= small for gestational age, AGA= appropriate for gestational age

9.3 Method

9.3.1 Sample

In a nation-wide cohort of VPT and VLBW infants, 94% (n=1338) of all infants, born alive in 1983 in the Netherlands, with a gestational age of less than 32 completed weeks and/or a birth weight less than 1500 grams were enrolled. Prospective data collection was performed during infancy, and at 5 and 10 years of age. Children with data until 10 years of age were included, which is 74% (n=688) of the survivors at 5 years of age. The non-response group included more children with low Social Economical Status (SES), handicaps and disabilities at 5 years of age. This resulted in a positive selection bias.[18] The history of the cohort is described in more detail in previous publications.[1,18,19] Table 1 shows the distribution of gestational age, birth weight, gender, handicaps, SES and maternal age that could influence HS development.

Table 2. The grouping of the 16 HS variables before and after analyses

Theoretical HS sub-domains	*Result of principal component analyses* [a]
Physical complaints	**Basic functioning problems**
vision problems	speech problems (key variable)
hearing problems	vision problems
respiratory problems	hearing problems
	moving problems
Motor functioning	
moving problems	**Concentration problems**
clumsiness	not concentrated (key variable)
	school problems
Cognitive functioning	clumsiness
not concentrated	unsociable
school problems	wriggling
speech problems	reticent
	afraid
Emotional functioning	
unhappy	**Negative moods**
worried	tense (key variable)
tense	unhappy
afraid	worried
	felt picked on
Social functioning	
felt picked on	
unsociable	
reticent	
wriggling	

[a] Respiratory problems appeared not necessarily as HS indicator.

9.3.2 *Measures and analysis strategy*

Sixteen variables were used that were prospectively collected from the parents and that covered all HS domains according to experts. To study HS development between 5 and 10 years of age the association between these variables were computed using principal component analysis.[22-24] Variables with a strong correlation form a cluster in the solution of the principal component analysis. These clusters of variables were labelled according to their content. In Table 2 (see previous page) the 16 variables are given grouped according to theoretical HS sub-domains and according to the results of the principal component analysis. From each cluster a key variable was selected, to represent the other variables in the cluster in further analyses. The influence of clinical and demographic factors (Table 1) on HS changes with age was assessed by relating them to the key variables. Statistical significance was tested with repeated measures ANOVA. Full details about the selection of the 16 variables and the analysis strategy are described elsewhere.[25]

9.4 Results

9.4.1 *Relationships between HS variables over age*

The average proportion of variance accounted for by the principal component analysis in two dimensions is 31%. Dimension 1 accounts for 21% which is equivalent to a Cronbach's alpha of .76 and dimension 2 accounts for 10% which is an alpha of .46.[26] According to the principal component analysis, 35% of the children had HS problems at both ages and 32% had no HS problems at either age. The remaining third either had problems at age 5 (16%) or at age 10 years (17%). The variables were clustered by the analyses in three main problem areas: basic functioning, concentration and negative moods (see Table 2). The variable "respiratory problems" did not belong to one of these areas and was therefore excluded. From each cluster one variable could be selected to represent the cluster in further analyses. The key variable "speech problems" represents basic functioning problems, "tensed" represents negative moods, and "not concentrated" represents concentration.

Table 3. Repeated measures analysis of variance between perinatal factors and age of measurement with respect to the three key variables

		Basic functioning problems *'speech problems'* F	Concentration problems *'not concentrated'* F	Negative moods *'tense'* F
Gender	between subjects (groups)	88.23 ***	40.14 ***	4.61 *
	within subjects (age)	450.01 ***	37.73 ***	608.41 ***
	interaction	25.11 ***	2.82	37.77 ***
Gestational age, in weeks	between subjects (groups)	4.99 **	10.3 ***	20.45 ***
	within subjects (age) [a]	286.28 ***	37.12 ***	431.32 ***
	interaction	5.61 ***	0.38	1.09
Birth weight percentile score	between subjects (groups)	4.3 ***	7.38 ***	8.58 ***
	within subjects (age)	353.87 ***	35.98 ***	463.59 ***
	interaction	1.52	2.92 **	4.17 ***
Handicaps	between subjects (groups)	222.98 ***	43.97 ***	15.32 ***
	within subjects (age)	222.58 ***	1.64	112.25 ***
	interaction	24.99 ***	2.63 *	4.59 ***
Maternal age by SES at enrollment	between subjects (groups)	7.69 ***	6.18 ***	4.6 ***
	within subjects (age)	354.12 ***	28.35 ***	360.61 ***
	interaction	8.25 ***	1.46	1.78

[a] differences between measurements at 5 and 10 years of age; * p≤.05 ** p≤.01 *** p≤.001

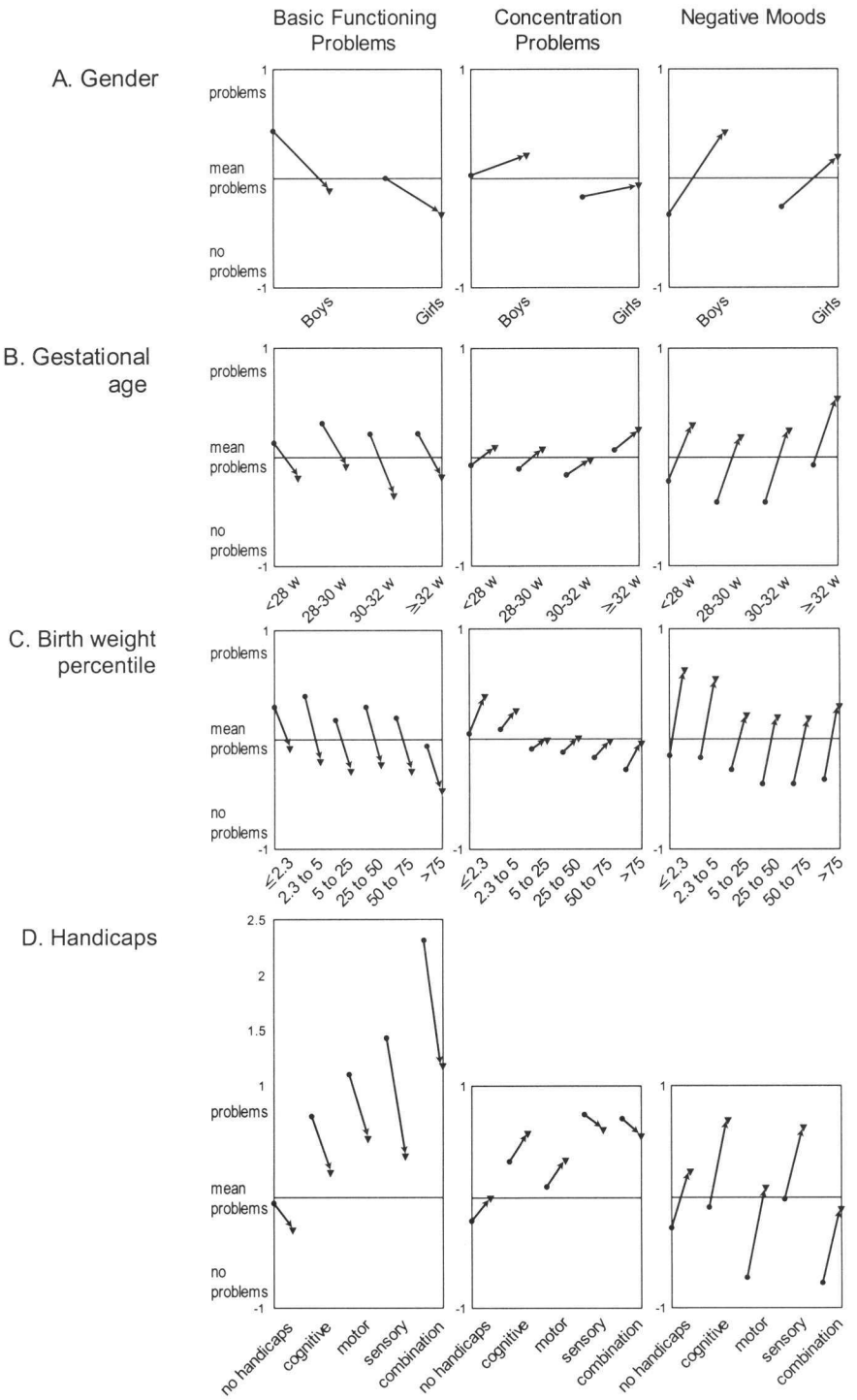
9.4.2 Relationships between clinical and demographic factors and HS development

The relations between the key variables and the age factor are inspected for different groups according to clinical and demographic factors; they are depicted in Figure 1a to 1d. A score of zero in the figures equals the mean HS problem score of the sample. The mean HS score indicates a minor problem level at the two time points according to means of the original variables. A score below zero indicates no HS problems and a score above zero indicates minor to severe problems. Differences between groups in arrow length show group effects, total length shows change over

time and parallel lines indicate no interaction between group and time. The magnitude of the differences between groups and with age with respect to the key variables are given in Table 3 at the previous page.

HS appeared to change with age in all groups: problems in basic functioning decrease with age, negative moods increase, and concentration problems increase slightly. Figure 1a shows that boys had more problems than girls with one exception: at 5 years of age, boys had fewer negative moods than girls. As can be seen in Table 3, the differences between boys and girls as well as differences between 5 and 10 years of age are statistically significant. In addition, gender × age interactions are significant in basic functioning problems and negative moods. It can be concluded that boys overall have more problems than girls and the change, either positive or negative is larger for boys than for girls. More concentration problems and negative moods were found at lower gestational ages but problems in basic functioning were not linearly related to gestational age (Figure 1b). Children born ≥ 32 weeks had the most negative moods and concentration problems. An explanation could be that, due to the inclusion criteria, these children were all small for gestational age. This is confirmed by the relation found between birth weight percentile and HS problems. Figure 1c shows that HS problems generally were more prevalent in children with lower birth weight percentile scores. Both group and age differences were significant, and group × age interactions were significant for concentration problems and negative moods. The most important for HS, however, were not the perinatal data but the handicap status, especially at age 5 (see Figure 1d). Children with handicaps had far more problems in basic functioning and concentration problems than the children without handicaps, at both ages. Children with a combination of handicaps had many concentration problems and the most problems in basic functioning. Negative moods were primarily found in children with cognitive and sensory problems and in the children without handicaps. Surprisingly, even in children with handicaps problems in basic functioning clearly decreased with age. Although negative moods increased in all groups, these problems were significantly less prevalent in children with motor handicaps and in children with a combination of handicaps at 10 years of age.

Figure 1. Clinical and demographic factors at 5 and 10 years of age with respect to basic functioning problems, concentration problems and negative moods (A higher score represents more HS problems). Legend: ● = 5 years of age, ▼ = 10 years of age


Basic Functioning Problems
Concentration Problems
Negative Moods
A. Gender
B. Gestational age
C. Birth weight percentile
D. Handicaps
problems
mean problems
no problems
Boys
Girls
<28 w
28-30 w
30-32 w
≥32 w
≤2.3
2.3 to 5
5 to 25
25 to 50
50 to 75
>75
no handicaps
cognitive
motor
sensory
combination

The overall pattern of a decrease in basic functioning problems and an increase in concentration problems and negative moods also applied to each maternal age and SES group (data not shown). Children of younger mothers have more HS problems. SES did not have a linear relation with HS. Group × age interaction was significant in basic functioning problems only.

9.5 Discussion

Two thirds of the study group had HS problems either at one or at both assessments, which were mostly minor. Not surprisingly, children without handicaps have far fewer basic functioning and concentration problems than children with handicaps do, and children with a combination of handicaps show the most problems. Negative moods are relevant at 10 years of age in children with a cognitive or a sensory handicap, but also in preterm children without a handicap. A similar result about this non-handicap group was reported by Schothorst et.al.[27] Furthermore, boys had more problems than girls at both ages. This gender difference is often found: not only is mortality in VPT and VLBW boys higher than in girls, when boys survive they remain more at risk.[28,29] This is, however, found in children born full-term as well.[27]

Although there was a relation between gestational age and HS problems, the birth weight percentile appeared a more important risk factor. Probably the substantial number of VLBW children of relatively long gestational age in the present sample has masked the influence of gestational age. In other studies small for gestational age children had more HS problems than normal for gestational age children of similar gestational ages.[30,31] Silva et al. formulated it as: "it is better to be born too early than too small".[31]

The non-linear relationship between HS problems and SES that we found is influenced by the age of the mother. Children of young mothers with low SES have most HS problems. Older mothers may have obtained the maturity and the responsiveness which is important for an adequate raising of preterm children, in spite of low SES.[30,32,33] Another explanation for the ambiguous impact of SES we found may be the relatively small class differences with regard to exposure to health service created by the health security system in the Netherlands.[29]

The relative importance of HS problems in the cohort of preterm born children appeared to change with age. Problems in basic functioning decrease while concentration problems and negative moods increase. Although there are many longitudinal HS studies in preterm born children,[27-39] for several reasons comparison with the results of these studies is complicated. Firstly, these previous studies had small samples of preterm children (range 30 to 145), with diverse selection criteria (with or without handicaps and various cut of points for gestational age). They used a variety of instruments and informants of HS (clinicians, parents, psychologist, teachers). As a result, generalisations are difficult to draw from these studies. Secondly, studies including a physical, psychological as well as a social HS domain are still limited, and these studies explored the HS domains one by one.[30,31,34,39] We explored new combinations of HS domains in order to pinpoint the problems that are relevant for preterm born children in particular. Thirdly, in half of the longitudinal studies we found, analyses were restricted to correlations between different time points. Correlations tested the strength of a relation between time points, but not differences in height. As a result, changes in HS could not be reported. In the other half of the longitudinal studies we found, some reported changes similar to ours,[28,31,32,38] other studies deviated.[27,29,33] From this selection, only Silva et al.[31] combined all HS domains. They measured changes in intelligence and problem behaviour (social and physical problems) in 31 children born preterm (< 37 w.), 71 SGA children and 748 children born full-term with normal birth weight. Intelligence appeared stable and problem behaviour increased not only in the preterm groups but also in the control group. Therefore, if the scores of the preterm groups had been standardised by the scores of the control group, no increases would have been reported.[33] In our study we cannot prove that the absolute changes we found are relative changes as well when compared with a norm or control group. A limitation of this study is that a norm or control sample was not available.

The most remarkable HS change we found is the decrease in basic functioning problems, even in the group of children with a combination of handicaps. There is however a plausible explanation for these results. Parents might develop a tolerance for some problems as their children grow older.[27,29] On the other hand, not all kinds of problems are detected at the same time. The normal development of children at a certain age, is reflected in the type of problems the parents reported in the present study. Basic functioning (like vision and speech) is important from early childhood onwards. At "kindergarten age", concentration problems become more important to the

parents. When children grow older they start to communicate more about their moods. Therefore, the parents could successively detect problems in basic functioning, concentration problems and negative moods. As a result, at the age of 10 years the parents could have grown quite accustomed to the first, less to the second, but not to the last kind of HS problems. For instance, parents and children can get used to the burden of a child's handicap. They become physically adjusted with the aid of the available medical treatments and devices, and they become psychologically adjusted because 'one can get used to almost anything'. The basic functioning problems that appeared to be so overwhelming at five years of age, perhaps were less intrusive in daily life at ten years of age. As a result, parents may report fewer basic functioning problems when their handicapped child is 10 years of age, whereas the clinician might report as much problems as before, because the child remained handicapped. Since parents are the main decision-makers in respect of the rearing and medical treatment of their children, their perception of their child's HS problems is most relevant.

In conclusion, a VPT or VLBW birth can have long-term effects. Only one third of the cohort born in 1983 had no HS problems at age 5 and 10 years. The rest of the cohort had minor to severe HS problems. Problems in basic functioning decreased and negative moods and concentration problems increased. Groups at the highest risk were preterm born children with handicaps, with a birth weight below the 5th percentile and boys. In the future, new HS follow-up studies, preferably with a control group, are needed to judge the validity of this study's results in current cohorts of preterm born children.

Acknowledgement
We would like to thank all participants from the Collaborative Project on Preterm and Small for Gestational Age (POPS) Infant in the Netherlands. We also like to thank Dr. G. H.W. Verrips for his review of the manuscript. We are indebted to the parents for their participation.

9.6 Reference List

1. Veen S, Ens-Dokkum MH, Schreuder AM, Verloove-Vanhorick SP, Brand R, Ruys JH. Impairments, disabilities, and handicaps of very preterm and very-low-birthweight infants at five years of age. *Lancet* 1991; **338:** 33-36.
2. Fawer CL, Besnier S, Forcada M, Buclin T, Calame A. Influence of perinatal, developmental and environmental factors on cognitive abilities of preterm children without major impairments at 5 years. *Early Hum Dev* 1995; **43:** 151-164.
3. Robertson CM, Etches PC, Kyle JM. Eight-year school performance and growth of preterm, small for gestational age infants: a comparative study with subjects matched for birth weight or for gestational age. *J Pediatr* 1990; **116:** 19-26.
4. Tin W, Wariyar U, Hey E. Changing prognosis for babies of less than 28 weeks' gestation in the north of England between 1983 and 1994. Northern Neonatal Network. *Brit Med J* 1997; **314:** 107-111.
5. Hall A, McLeod A, Counsell C, Thomson L, Mutch L. School attainment, cognitive ability and motor function in a total Scottish very-low-birthweight population at eight years: a controlled study. *Dev Med Child Neurol* 1995; **37:** 1037-1050.
6. Middle C, Johnson A, Alderdice F, Petty T. Birthweight and health and development at the age of 7 years. *Child Care Health Dev* 1996; **22:** 55-71.
7. Halsey CL, Collin MF, Anderson CL. Extremely low birth weight children and their peers: a comparison of preschool performance. *Pediatr* 1993; **91:** 807-811.
8. Saigal S, Rosenbaum P, Stoskopf B, et al. Comprehensive assessment of the health status of extremely low birth weight children at eight years of age: Comparison with a reference group. *J Pediatr* 1994; **125:** 411-417.
9. Lane SJ, Attanasio CS, Huselid RF. Prediction of preschool sensory and motor performance by 18- month neurologic scores among children born prematurely. *American Journal of Occupational Therapy* 1994; **48:** 391-396.
10. Levy-Shiff R, Einat G, Har Even D, et al. Emotional and behavioral adjustment in children born prematurely. *J Clin Child Psychol* 1994; **23:** 323-333.
11. Ross G, Lipper EG, Auld PA. Social competence and behavior problems in premature children at school age. *Pediatr* 1990; **86:** 391-397.
12. McCormick MC, Workman Daniels K, Brooks Gunn J. The behavioral and emotional well-being of school-age children with different birth weights. *Pediatr* 1996; **97:** 18-25.
13. Stanton WR, McGee R, Silva PA. Indices of perinatal complications, family background, child rearing, and health as predictors of early cognitive and motor development. *Pediatr* 1991; **88:** 954-959.
14. Wolke D, Ratschinski G, Ohrt B, Riegel K. The cognitive outcome of very preterm infants may be poorer than often reported: an empirical investigation of how methodological issues make a big difference. *Eur J Pediatr* 1994; **153:** 906-915.
15. Vogels T, Verrips GH, Verloove-Vanhorick SP, et al. Measuring health-related quality of life in children: the development of the TACQOL-Parent Form. *Qual Life Res* 1998; **7:** 457-465.
16. Bullinger M, Ravens-Sieberer U. Stand der Forschung zur gesundheitsbezogenen Lebensqualität von Kindern: eine Literaturanalyse [German: State of the art of health-related quality of life research in children: a literature review]. *Präv Rehab* 1995; **7:** 106-121.
17. Grunberg SM, Groshen S, Steingass S, Zaretsky S, Meyerowitz B. Comparison of conditional quality of life terminology and visual analogue scale measurements. *Qual Life Res* 1996; **5:** 65-72.
18. Hille ETM, den-Ouden L, Bauer L, Van-den-Oudenrijn C, Brand R, Verloove-Vanhorick SP. School performance at nine years of age in very premature and very low birth weight infants: Perinatal risk factors and predictors at five years of age. *J Pediatr* 1994; **125:** 426-434.
19. Verloove-Vanhorick SP, Verwey RA, Ebeling MCA, Brand R, Ruys JH. Mortality in very preterm and very low birth weight infants according to place of birth and level of care: Results of a national collaborative survey of preterm and very low birth weight Infants in the Netherlands. *Pediatr* 1988; **81:** 404-411.

20. van-Westerlaak JM, Kropman JA, Collaris JWM. *Beroepenklapper [Dutch: Professional index].* Nijmegen: Instituut voor Toegepaste Sociologie [Institution for Applied Sociology], 1975.
21. Kloosterman GJ. Over intra-uteriene groei en de intrauteriene groeicurve [Dutch: Intrauterine growth and intrauterine growth curves]. *Maandschr Kindergeneeskd* 1969; **37:** 209-225.
22. *SPSS Categories 8.0.* Chicago: SPSS Inc., 1998.
23. van-der-Ham T, Meulman JJ, van-Strien DC, van-Engeland H. Empirically based subgrouping of eating disorders in adolescents: A longitudinal perspective. *Brit J Psychiat* 1997; **170:** 363-368.
24. Kerkhof AJFM, van-der-Wal J, Hengeveld MW. Typology of persons who attempted suicide with predictive value for repetition: a prospective cohort study. In: Möller HJ, Schmidtke A, Welz R, eds. *Current issues of suicidology.* Berlin: Springer-Verlag, 1988: 194-203.
25. Theunissen NCM, Meulman JJ, den-Ouden AL, et al. Changes in quality of life can be studied when the measurement instrument is different at different time points. *submitted* 1999.
26. Heiser WJ, Meulman JJ. Homogeneity analysis: Exploring the distribution of variables and their nonlinear relationships. In: Greenacre M, Blasius J, eds. *Correspondence analysis in the social sciences: Recent developments and applications.* London, San Diego, New York, Boston, Sydney, Tokyo, Toronto: Academic Press, Harcourt Brace & Company, 1994: 179-209.
27. Schothorst PF, van-Engeland H. Long-term behavioral sequelae of prematurity. *J Am Acad Child Adolesc Psychiatry* 1996; **35:** 175-183.
28. Largo RH, Molinari L, Kundu S, Lipp A, Duc G. Intellectual outcome, speech and school performance in high risk preterm children with birth weight appropriate for gestational age. *Eur J Pediatr* 1990; **149:** 845-850.
29. Bosch JD, Veerman JW, Samson JF, Smith F. Ex-neonatale couveusepatieentjes vijf en tien jaar later. Een follow-up-onderzoek naar mogelijke gedragsproblemen [Dutch: Ex-neonatal intensive-care-unit (NICU) patients five and ten years later: A follow-up study of risk for behavioral problems]. *Kind en Adolescent* 1995; **16:** 1-15.
30. Caputo DV, Goldstein KM, Taub HB. Neonatal compromise and later psychological development: A 10-year longitudinal study. In: Friedman SL, Sigman M, eds. *Preterm birth and psychological development.* New York, etc: Academic Press, 1981: 353-385.
31. Silva PA, McGee R, Williams S. A longitudinal study of the intelligence and behavior of preterm and small for gestational age children. *J Dev Behav Pediatr* 1984; **5:** 1-5.
32. Kalmar M. The course of intellectual development in preterm and fullterm children: An 8-year longitudinal study. *International Journal of Behavioral Development* 1996; **19:** 491-516.
33. Gross SJ, Slagle TA, D'Eugenio DB, Mettelman BB. Impact of a matched term control group on interpretation of developmental performance in preterm infants. *Pediatr* 1992; **90:** 681-687.
34. Portnoy S, Callias M, Wolke D, Gamsu H. Five-year follow-up study of extremely low-birthweight infants. *Dev Med Child Neurol* 1988; **30:** 590-598.
35. Largo RH, Graf S, Kundu S, Hunziker U, Molinari L. Predicting developmental outcome at school age from infant tests of normal, at-risk and retarded infants. *Dev Med Child Neurol* 1990; **32:** 30-45.
36. Saigal S, Szatmari P, Rosenbaum PL. Can learning disabilities in children who were extremely low birth weight be identified at school entry? *J Dev Behav Pediatr* 1992; **13:** 356-362.
37. Menyuk P, Chesnick M, Liebergott JW, Korngold B. Predicting reading problems in at-risk children. *Journal of Speech and Hearing Research* 1991; **34:** 893-903.
38. Fedrizzi E, Inverno M, Botteon G, Anderloni A, Filippini G, Farinotti M. The cognitive development of children born preterm and affected by spastic diplegia. *Brain Dev* 1993; **15:** 428-432.
39. Saylor CF, Casto G, Huntington L. Predictors of developmental outcomes for medically fragile early intervention participants. *J Pediatr Psychol* 1996; **21:** 869-887.

10 Quality of life and self-esteem in children treated for idiopathic short stature

N.C.M. Theunissen, G.A. Kamp, H.M. Koopman,
K.A.H. Zwinderman, T.G.C. Vogels, J.M. Wit

10.1 Abstract

Changes in Health Related Quality of Life (HRQoL) and self-esteem were studied in children with idiopathic short stature (ISS) participating in a prospective randomised controlled study on the effect of Growth Hormone (GH) treatment. The sample consisted of forty prepubertal children (age 4 to 10 years old at start) with ISS (height < -2 SD). The children were randomly assigned to a treatment or control group. HRQoL and self-esteem were assessed three times: shortly after randomisation (T1), after one (T2) and two years (T3). Children with ISS, their parents and the paediatrician completed questionnaires. At T1, children with ISS did not have a lower HRQoL and self-esteem than the norm population, except for social functioning as reported by child and parents. Children, parents and paediatrician assessed HRQoL differently: At T3, children of the treatment group reported in some scales lower HRQoL and self-esteem than the control group did. The parent reports did not differ between groups, but the paediatrician reported improved HRQoL in the treatment group. Changes in HRQoL and self esteem between T2 and T3 hardly related to growth (objectively measured or as perceived by the child). Instead, changes in several HRQoL and self-esteem scales related to the height appreciation by the child her/himself. The assumption that GH treatment improves HRQoL in children with ISS could not be supported in this study. *Key words: health related quality of life, health status, self-esteem, idiopathic short stature, growth hormone treatment.*

10.2 Introduction

If a child has a short stature without an underlying disease or deficiency, the diagnosis of idiopathic short stature (ISS) is made. More precisely, a child with ISS has: a height of more than 2SD below the mean of an age and sex specific population reference, no GH deficiency, a normal size for gestational age at birth, normal body proportions, no evidence of chronic organic disease, no psychiatric disease or severe emotional disturbance, and normal food intake.[1,2]

It is widely assumed that ISS can result in psychological, social, and physical problems.[3] Short children appear younger than they are, and are often treated as such by adults and peers.[4-6] The physical limits they meet in sports and play with peers could make them the subject of ragging and baiting.[5,6] Nevertheless, other studies showed that short children have unimpaired self-esteem and normal patterns of behaviour.[7,8] With regard to adults, short stature is generally considered a social disadvantage in Western society, although it was recently found that quality of life was little affected by shortness.[9] For more than a decade, the answer to the potential problems due to short stature has been to try to accelerate growth through growth hormone (GH) treatment.[10,11] Although GH deficiency is still the main indication for treatment, there is also evidence that GH treatment stimulates growth in short children without GH deficiency, but the effect on final height appears limited.[6,12,13] However, most studies into the GH treatment of ISS, had methodological flaws, such as inadequate controls and the use of mixed diagnostic groups, and no accurate predictors of the growth response.[3,13] Therefore, we conducted a prospective randomised controlled GH-dose-response study, preceded by an extensive biochemical characterisation of GH sensitivity.[14] Furthermore we investigated whether participating in the study, as part of the treatment or control group, would influence the well-being of the children.

In our study we used the construct Health related Quality of Life (HRQoL) to study the physical, psychological and social consequences of being short . HRQoL is defined by the individual's perception of problems in Health Status (HS), combined with the affective responses to such problems.[15-17] An aspect of psychological functioning often referred to in the field of short stature is self-esteem.[3,10,18-20] Self-esteem is defined by the way someone perceives and appreciates her or himself.[18] Short stature is supposed to reduce the HRQoL and self-esteem in childhood as well as in adulthood.[3,9]

All participants intended to undergo treatment if they were allocated a place by lot, but half of the children were randomly assigned to the control group. We hypothesised that this might result in lower HRQoL or self-esteem, although this effect might vanish during the study. Treatment with GH imposes a heavy burden on the child and his or her parents as it requires daily injections for several years.[21] Nevertheless it was expected that if short stature is accompanied by low HRQoL or self-esteem, both should be improved by increasing height. As a result, in the treatment group the HRQoL and self-esteem might become better than in the control group. On the other hand, since the effect of GH on ISS is not certain, disappointment (in the treatment group) or surprise (in the control group) about achieved height, might prevent the treatment group from having better HRQoL than the control group.

The aim of the present paper is to describe the effect of GH treatment on the HRQoL and self-esteem in children with ISS. In addition, the impact of growth expectations, growth achievements and the appreciation of the current height on the changes in HRQoL were investigated.

Table 1. Distribution of baseline characteristics

		Treatment N (%)	Control N (%)	*p* *
Child's sex	male	15 (75)	14 (70)	*.72*
Age (in years)	4-7	9 (45)	11 (55)	
	8-12	11 (55)	9 (45)	*.53*
Responding parent's sex	female	19 (95)	17 (85)	*.35*
Hospital	Eindhoven	8 (40)	13 (65)	
	Rotterdam	9 (45)	5 (25)	
	Amsterdam	3 (15)	2 (10)	*.25*
Maternal educational level	low	8 (42)	7 (39)	
	medium	8 (42)	11 (61)	
	high	3 (16)	0 (0)	*.17*
Paternal educational level	low	9 (47)	11 (61)	
	medium	6 (32)	3 (17)	
	high	4 (21)	4 (22)	*.56*

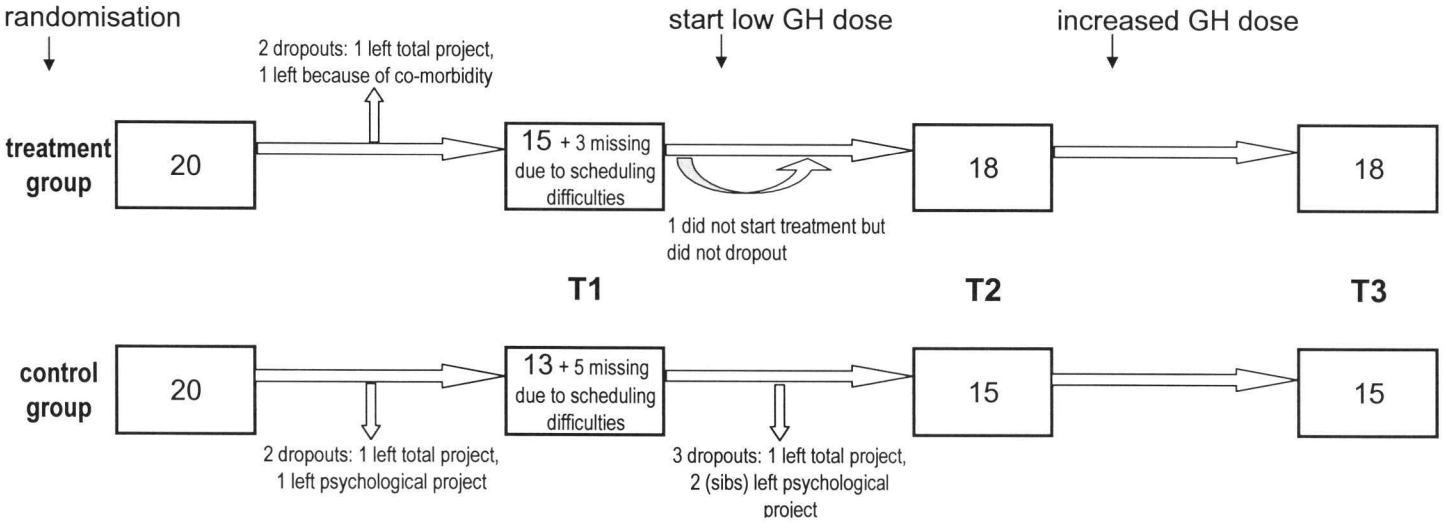
* chi-square test

10.3 Method

10.3.1 Subjects and data collection procedures

Data were collected from 40 prepubertal children (age 4 to 10 years at start) with ISS and their parents. They were enrolled between 1994 and 1997 (born between 1983-1992) in the Dutch multicenter study on the responsiveness to short and long-term GH therapy. Most children were boys (73%) and most children were accompanied by their mother (90%). Two children in the treatment group and one in the control group had parents who originated from Turkey. Table 1 shows the demographic data, which were distributed evenly over the two groups according to chi-square analyses.

Figure 1. Composition of the study group over time



The children were randomly assigned to a GH treatment group (n=20) and a control group (n=20). The children of the GH treatment group underwent an extensive biochemical assessment to verify their GH responsiveness during the first year. High dose GH therapy (Pharmacia & Upjohn, Stockholm, Sweden) was started and given for at least two years. To obtain medical and biochemical parameters, the children in the treatment group visited the paediatrician (GAK) four times a year, the children in the control group one or two times a year. To obtain data on HRQoL and self-esteem, all children visited the hospital three times extra: after randomisation but two weeks before treatment or an equivalent period in the control group (T1), one year (T2) and two years (T3) after start of treatment. Children and their parents were scheduled at the same time but were interviewed independently in different rooms. The parents were seen by a developmental psychologist (NCMT) and the children by different psychology graduate students. The sessions took two hours at T1 and one-and-a-half hours at T2 and T3. The protocol was reviewed and approved by the medical ethics committees of the three participating centres in the Netherlands (Catharina Hospital Eindhoven, Sophia Children's Hospital Rotterdam, Free University Hospital Amsterdam). The parents of all children gave written consent to the study. When appropriate, the consent of the children was also obtained. Figure 1 presents a flowchart on the dropouts.

10.3.2 Measures

TACQOL (TNO-AZL Children's Quality Of Life) questionnaire

This is a 56-item generic instrument for assessing HRQoL of children, aged 6 to 12 year, in medical research and clinical trials.[15-17,22] Two parallel questionnaires for the child's HRQoL were available: a child form (CF) and a parent form (PF), both with good measurement properties. The TACQOL explicitly offers respondents the possibility of differentiating between their functioning and the way they feel about it. The reference period is formulated as 'the last few weeks'. The questionnaire contains seven scales of eight items each: physical complaints, motor functioning, autonomy, cognitive functioning, social functioning, positive emotions and negative emotions. The final version of the TACQOL appeared in 1995. As the first children were enrolled in 1994, an earlier version was used during the whole research period. This version used 117 items which also included the definitive ones. Besides some minor rephrasing of items which we considered negligible, the main difference was the number of categories used. In both versions a concretely and specifically formulated Health Status problem,

if reported, leads to a question about the emotional response. Figure 2 shows an example. Whereas the final version gives four options by which to express an emotional evaluation (very good, not so well, rather bad, bad), the previous version used in the present study gave five (good, rather good, ordinary, rather bad, bad). Based on categorical analyses (HOMALS, data not shown) we decided to combine 'rather good' and 'ordinary' and recoded them into one category. To explore the difference between the four and five category solutions we computed scales for both options. Spearman's rho correlations and paired t-tests revealed negligible differences between the four and five category solutions. As the previous version was used for both groups at all three points of measurement, recoding of the items from five to four categories could not have influenced the comparison between groups in this longitudinal study.

Items were scored 0-4 (added in Figure 2 in italics, between brackets). Scale scores were obtained by summing item scores within scales, and transforming crude scale scores linearly to a 0-100 scale, with higher scores indicating better HRQoL. Emotional responses were not assessed in items of the two emotion scales, because this would have led to illogical items. In these scales HS equalled HRQoL, item scores were on a 0-2 scale (0 = often, 1 = sometimes, 2 = never) and scale scores ranged from 0-100. Item scores of the other scales were also encoded excluding the emotional evaluation to obtain HS scores (0 = often, 1 = sometimes, 2 = never) and scale scores ranged from 0-100, with higher scores representing better HS. As a result, four series of scales could be used: a HRQoL parent form, HRQoL child form, HS parent form, and a HS child form.

Figure 2. An example of a TACQOL-PF question. The TACQOL-CF equivalent of the question is 'Have you had difficulty running?' and 'During this I felt:'

Has your child had difficulty running? ❑ never *[4]* ❑ sometimes ❑ often

During this my child felt:

❑(very)good *[3]* ❑ not so well *[2]* ❑ rather bad *[1]* ❑ bad *[0]*

ISSQOL

This is an eight item ISS specific scale covering vitality, which is used in addition to the generic TACQOL. A disorder-specific instrument has the potential for increased responsiveness, because only important aspects of HRQoL are included relevant to the patients being studied.[23-25] Vitality was chosen since GH treatment can

improve the energy level in children or adults with GH deficiency.[4,26] From the 117 items of the old version of the TACQOL, 8 items were selected that covered vitality and were not included in the 56 items from the definitive TACQOL. The ISSQOL parent form (PF) had Cronbach's alphas of .71 (HRQoL) and .75 (HS), the alphas of the ISSQOL child form (CF) were .66 (HRQoL) and .56 (HS). As in the TACQOL, the ISSQOL scale score was obtained by adding item scores within scales, and transforming crude scale scores linearly to a 0-100 scale, with higher scores indicating better HRQoL. In addition, HS scores were obtained, excluding the emotional evaluation.

DUCATQOL (Dutch Children's AZL/TNO Quality Of Life)

This is a 25-item generic self-report HRQoL for school-aged children (5 to 16 years).[27] HRQoL was defined as the children's affective evaluation of various aspects of their daily functioning. The DUCATQOL has good validity and reliability and covers the four domains, home, physical, emotional, social functioning, plus a total HRQoL score. The reference period is formulated as 'recently'. The items use abstract faces (smiley) as answer categories, with expressions from happy to sad, thus constructing a five-point Likert scale. If reading skills are not sufficient, questions can be read and children can answer how they feel about a subject by pointing to a face. This makes the instrument useful for longitudinal research starting at an early age. Scale scores were obtained by adding item scores within scales, and transforming crude scale scores linearly to a 0-100 scale, with higher scores indicating better HRQoL.

Self-Perception Profile

This is a 36-item questionnaire measuring self-esteem in children 8 to 12 years old.[18,28] The questionnaire stresses the multi-dimensional nature of the child's sense of competence. It contains six scales of six items each: scholastic competence, social acceptance, athletic competence, physical appearance, behavioural conduct and general self-worth. Statements are presented in a forced choice format and the child is asked to compare her or himself to peers. When children lacked sufficient reading skills the questions are allowed to be read aloud. Therefore we could use the instrument for children younger than 8 years old as well. Scale scores are obtained by adding item scores within scales, and transforming crude scale scores linearly to a 0-100 scale, with higher scores indicating better perceived self-esteem.

Therapy Evaluation Scale (TES)

The TES was constructed by the Dutch Working Group 'Psychologists and Growth Hormone.[29] It provides qualitative and quantitative information on the expectations and experiences about participating in a growth study. The TES has versions for child and parent, treatment and control group, and versions that are slightly adapted for assessment before versus during treatment. If reading skills are not sufficient, the questions are allowed to be read aloud. Perceived growth 'since the last year', and growth expectations 'in the next year' were obtained using a 1-6 categorical scale ("not" to "very much"). Other aspects were covered by open-ended questions. At T1 the TES was used to infer the motivations of children and parents to participate in the project. At T1, T2 and T3 subjects were asked why they wanted to continue participation, and if they could mention drawbacks of the project.

Judgement of the paediatrician: Global HRQoL of the child

This is a short global HRQoL measure, obtaining the judgement of the paediatrician. It contained three questions: seriousness of present medical situation, suffering of the child from current medical situation, and suffering of the child from participating in the research program. The questions used a 1 to 7 point Likert scale ("not serious at all" to "very serious").

Child's and parent's characteristics

In addition, questions were asked which assessed the child's and parent's characteristics. This included general information such as parental education. Changes in height were measured by the paediatrician and expressed in centimetres and as Standard Deviation Score (SDS).[30]

10.3.3 Data analysis

Dutch reference data were available for the TACQOL-CF, Self-Perception Profile, and DUCATQOL concerning children aged 8-12 years, and on the TACQOL-PF concerning children aged 6-12 years. To see if the reference data could also be used for children outside these age range, differences between children within and outside these age ranges were tested with MANOVA. If no differences were found, reference data of the age group nearest by were used to calculate standardised scores of children outside the age ranges. Raw scores were standardised with the mean and standard

deviation of a healthy norm group, to correct for age-related changes. Since standardisation of the TACQOL to the norm group may be influenced by the recoding of the items from five to four categories, raw scores were used as well in all analyses. The results of the treatment and control group were expressed as the mean scores and the standard deviations at all points of measurement. Differences between groups at each time point with respect to HRQoL, HS and self-esteem were tested by Mann-Witney U test for independent samples. The measurements over time were analysed using mixed-model ANOVA for repeated measurements (RM-ANOVA) with patient as random factor, and time and treatment and their interaction as fixed factors.[31] This model can account for most missing data mechanisms and allows for non-linear changes. The RM-ANOVA gave three indicators: time effect, group effect, and the interaction of time with group. A significant interaction effect indicates that HRQoL and self-esteem changed differently between groups during the project, the mean scores of the groups show the direction of change. Linear regression analyses were computed for every scale, to reveal possible relationships between the dependent variable changes in HRQoL or self-esteem, and the independent variables: growth (objective and subjective), growth expectations, height appreciation and group. In this analysis the difference between T2 and T3 was used for two reasons: at T1 some children had missing data, and most growth was expected after T2 because of increasing GH doses. To control for potential differences between the groups at that time, in addition T2 scores were used as a covariate. The results of the linear regression analyses will be reported as beta-weights which are standardised regression coefficients. A beta-weight means that one standard deviation change of an independent variable relates to beta units of change in HRQoL or self-esteem, given that other independent variables are held invariable. Agreement between children and parent reports on the TACQOL was tested and quantified with paired t-tests and intraclass correlations.[32]

Overall, results of statistical analyses were reported as significant with a p-value of .05 or less or as trends with a p-value between .05 and .10. Since only 20 children per group were available, we expected to have little power for clinically important differences. Therefore, Bonferroni correction for multiple testing was not performed. Statistical analyses were performed using SPSS 8.0.[33]

The open-ended questions from the TES were subjected to qualitative analyses according to 'grounded theory' principles.[34] The answers were categorised by content, and the distribution of the data over the categories in the two groups was studied. Although the frequency of certain answers can say something about the importance of a statement, results can be used to generate hypotheses but not to test them.

Table 2. Child's perception: Health Related Quality of Life and self-esteem at all measurements

	T1		T2		T3		Longitudinal changes[b]		
	Treatment M (SD)	Control M (SD)	Treatment M (SD)	Control M (SD)	Treatment M (SD)	Control M (SD)	*p* time	*p* group	*p* time X group
TACQOL-CF (in SDS)	n=12	n=9	n=17	n=10	n=15	n=14			
physical complaints	-1.11(1.42)	-1.41(1.67)	-1.50(1.06)	-1.38(1.29)	-1.67(0.74)	-0.74(1.13)*[a]	.74	.30	.05
motor functioning	-0.10(0.86)	-0.45(1.91)	-0.14(1.24)	0.43(0.50)	-0.37(0.99)	0.29(0.92)*	.74	.17	.24
autonomy	-0.54(1.63)	-0.69(1.85)	-0.80(1.84)	0.22(0.53)*	-0.41(1.13)	0.13(0.60)	.59	.14	.16
cognitive functioning	0.28(1.19)	0.20(0.90)	0.24(0.81)	0.65(0.89)**	-0.11(1.01)	0.44(0.98)*	.67	.19	.60
social functioning	-2.37(1.52)	-1.25(2.44)#	-1.68(1.73)	-0.69(1.54)	-2.64(1.51)	-0.56(1.03)**	.22	.00	.61
positive emotions	-0.51(1.05)	-0.41(1.20)	-0.78(1.20)	-0.22(1.29)	-0.75(1.28)	0.10(0.94)#	.87	.11	.95
negative emotions	0.05(0.95)	0.00(1.23)	-0.21(0.92)	0.00(1.06)	-0.17(0.96)	0.03(1.12)	.14	.66	.80
ISSQOL-CF (scale 0-100)	n=12	n=9	n=17	n=10	n=15	n=14			
vitality	77(18)	74(27)	76(16)	82(11)	70(13)	85(14)**	.59	.13	.04
DUCATQOL (in SDS)	n=15	n=13	n=18	n=15	n=18	n=15			
home	-0.15(1.12)	-0.40(1.24)	0.19(0.80)	0.01(0.92)	0.19(1.11)	0.24(0.97)	.12	.67	.81
physical	-0.12(1.41)	-0.29(1.51)	0.36(1.02)	0.07(1.22)	0.19(1.15)	0.10(1.37)	.26	.58	.66
emotional	0.10(1.36)	-0.03(1.28)	0.61(0.77)	0.42(0.80)	0.34(0.91)	0.31(1.40)	.11	.69	.89
social	-0.35(1.59)	-0.10(1.42)	0.51(0.84)	0.17(1.16)	0.09(1.08)	0.67(1.12)	.06	.60	.19
total QoL	-0.16(1.47)	-0.20(1.31)	0.55(0.72)	0.24(1.11)	0.26(1.01)	0.42(1.37)	.05	.82	.47
height appreciation: "about how tall I am I feel" (item from the physical scale)	-0.39(1.43)	-0.57(1.68)	-0.24(1.58)	-0.03(1.13)	-0.12(1.29)	-0.67(1.60)	.48	.61	.50
Self-Perception Profile (in SDS)	n=14	n=12	n=18	n=15	n=18	n=15			
scholastic competence	0.66(0.87)	0.26(1.05)	0.57(1.22)	0.51(0.94)	0.41(0.95)	0.84(1.12)	.22	.95	.05
social acceptance	-0.17(1.36)	0.19(0.87)	0.39(1.09)	0.36(0.87)	0.25(1.13)	0.85(0.86)	.21	.27	.27
athletic competence	-0.05(0.72)	0.20(0.61)	0.27(1.04)	0.22(0.73)	0.31(1.10)	0.43(0.93)	.11	.68	.84
physical appearance	-0.06(1.01)	-0.16(1.17)	-0.22(1.17)	0.17(1.00)	-0.09(1.20)	0.14(0.87)	.65	.52	.92
behavioural conduct	0.41(1.15)	0.55(1.57)	-0.02(1.05)	0.64(1.37)#	0.09(1.29)	1.00(1.22)*	.67	.10	.27
general self-worth	0.18(1.15)	0.32(0.95)	0.43(0.97)	0.22(1.22)	0.01(1.25)	0.36(0.87)	.68	.76	.26

Note: higher scores represent better HRQoL or self-esteem; [a] Between groups: Mann-Withney U test: ** p≤.01, * p≤.05, # p≤.10; [b] Results of the RM-ANOVAs, p-values of .10 and lower are underlined

Table 3. Parent's and paediatrician's perception: Health Related Quality of Life and height at all measurements

	T1		T2		T3		Longitudinal changes[b]		
	Treatment M (SD)	Control M (SD)	Treatment M (SD)	Control M (SD)	Treatment M (SD)	Control M (SD)	*p* time	*p* group	*p* time X group
TACQOL-PF (in SDS)	n=15	n=13	n=18	n=15	n=17	n=15			
physical complaints	-0.78(1.43)	-0.98(1.46)	-1.25(2.05)	-1.17(2.01)	-1.50(1.89)	-1.37(1.62)	.22	.98	.96
motor functioning	-0.26(0.93)	-0.22(0.98)	-0.29(2.20)	-0.25(1.25)	-0.15(0.95)	0.05(0.60)	.60	.73	.95
autonomy	-0.76(1.65)	-1.03(1.73)	-1.30(3.24)	-1.50(3.81)	-0.21(1.22)	0.01(0.63)	.07	.92	.83
cognitive functioning	0.34(0.80)	-0.77(1.75)#	-0.79(2.42)	-0.10(1.54)	-0.79(1.67)	-0.53(1.91)	.29	.94	.47
social functioning	-2.34(1.35)	-2.69(2.17)	-2.96(2.12)	-2.48(2.11)	-2.16(1.88)	-2.25(1.77)	.21	.95	.54
positive emotions	-0.05(1.04)	-2.10(2.36)**[a]	-0.86(1.92)	-0.15(1.04)	-0.21(1.19)	-0.07(0.90)	.29	.45	.03
negative emotions	-0.48(1.19)	-0.68(1.16)	-0.93(1.27)	-0.38(1.43)	-0.78(0.92)	-0.39(1.28)	.79	.43	.61
ISSQOL-PF (scale 0-100)	n=15	n=13	n=18	n=15	n=17	n=15			
vitality	75(14)	76(15)	75(23)	77(17)	76(19)	78(19)	.94	.65	.93
Judgement of the paediatrician (scale 0-100)	n=15	n=13	n=18	n=15	n=18	n=14			
seriousness of current medical situation	81(9)	77(20)	75(19)	68(20)	83(0)	81(15)	.05	.08	.73
suffering from current medical situation	56(25)	49(24)	63(23)	57(24)	79(11)	58(27)**	.02	.04	.27
suffering from participating in research program	63(28)	55(25)	64(26)	66(25)	79(17)	63(22)**	.05	.17	.27
Objective measurements	n=15	n=12	n=18	n=15	n-18	n=14			
height (in centimetres)	115(11)	114(9)	123(9)	120(9)	132(9)	124(8)**	.00	.24	.00
standard deviation score (SDS) for height at given age	-2.95(0.53)	-2.70(0.35)	-2.53(0.47)	-2.55(0.48)	-1.85(0.56)	-2.50(0.51)*	.00	.32	

Note: higher scores represent better HRQoL or self-esteem; [a] Between groups: Mann-Withney U test: ** p≤.01, * p≤.05, # p≤10;
[b] Results of the RM-ANOVAs , p-values of .10 and lower are underlined

10.4 Results

10.4.1 Children with ISS in comparison to a reference group

Reference data were available for limited age ranges (6-12 or 8-12 depending on the questionnaire). As differences between children within and outside these age ranges were not significant in this study group (data not shown), all available data were used in the analyses. Analyses with the raw data (not shown) were similar to the results with the data standardised according to age and sex, indicating a negligible effect of age and sex. The standardised scores of the children's HRQoL (TACQOL, ISSQOL and DUCATQOL) and self-esteem (Self-Perception Profile) questionnaires are given in Table 2 and the parent's and paediatrician's measurements are given in Table 3 (see previous pages). Most standard deviation scores (SDS) at T1 were around zero. The only scales showing scores at approximately -2 SDS , were the HRQoL social functioning scales (TACQOL-CF and PF). There was therefore little support for the hypothesis that children with ISS have lower HRQoL and self-esteem than a normal reference group.

10.4.2 Control group versus treatment group

Low HRQoL or self-esteem at T1 was expected in the control group, due to possible disappointment at start because of not being selected for treatment. This effect was found in the HRQoL positive emotions (significant) and cognitive functioning (trend) scales as reported by the parents only, and disappeared at T2 (see Table 3). An indication of primary disappointment in the control group was not found in the child reports. It was expected that the treatment group would gain more HRQoL and self-esteem than the control group. According to child reports (see Table 2) there were few significant differences between groups according to the Mann-Witney U test, and all in a direction opposite to the one expected. Most differences were found at T3 in the TACQOL-CF: the children from the treatment group had significantly worse HRQoL than the children in the control group according to the physical complaints, motor functioning, cognitive functioning, and social functioning scales. No differences between groups were reported by the parents except for the ones at T1 mentioned above (see Table 3). The paediatrician reported at T3 more suffering by the control group than by the treatment group (Table 3).

The RM-ANOVAs revealed some interaction effects, indicating that the scores changed differently between groups during the project. According to the children, the treatment group decreased and the control group increased in the TACQOL-CF physical complaints, Self-Perception Profile's scholastic competence, and ISSQOL-CF vitality scales. According to the parents, the TACQOL-PF positive emotions scale changed differently in both groups, probably because of the low scores in the control group at T1. Height expressed in centimetres improved more in the treatment group than in the control group, whereas the Height SDS improved significantly in the treatment group only (see also Figure 3 at the next page). The appreciation of height -- covered by an item from the DUCATQOL physical scale "about how tall I am I feel..." -- was not significantly different between groups (see also Figure 4 at the next page).

A group effect in the TACQOL-CF social functioning and Self-Perception Profile's behavioural conduct, indicates a systematic difference between groups according to the children, with lower scale scores for the treatment group. The group effect in the judgement of the paediatrician on the seriousness of, and suffering from the current medical situation, indicates that a difference between groups was experienced from the start.

10.4.3 Agreement between informants

According to paired t-tests between the TACQOL-CF and PF (not shown) the children reported significantly lower HRQoL than their parents on physical complaints at T2, higher HRQoL on cognitive and social functioning at T2 and T3 and higher HRQoL on negative emotions (=less negative emotions reported) at T3. The intraclass correlations between TACQOL-CF and PF were very low (-.29 to .54) and mostly not significant, except for negative emotions (T1: .52 and T2: .46), positive emotions (T2: .49 and T3: .54), and physical complaints (T2: .38). In the previous section it was shown that the pattern of longitudinal changes differed between child, parent and paediatrician. It can be concluded that the agreement between informants is low.

Figure 3. Height SDS over time, tested by mixed model ANOVA for repeated measurements

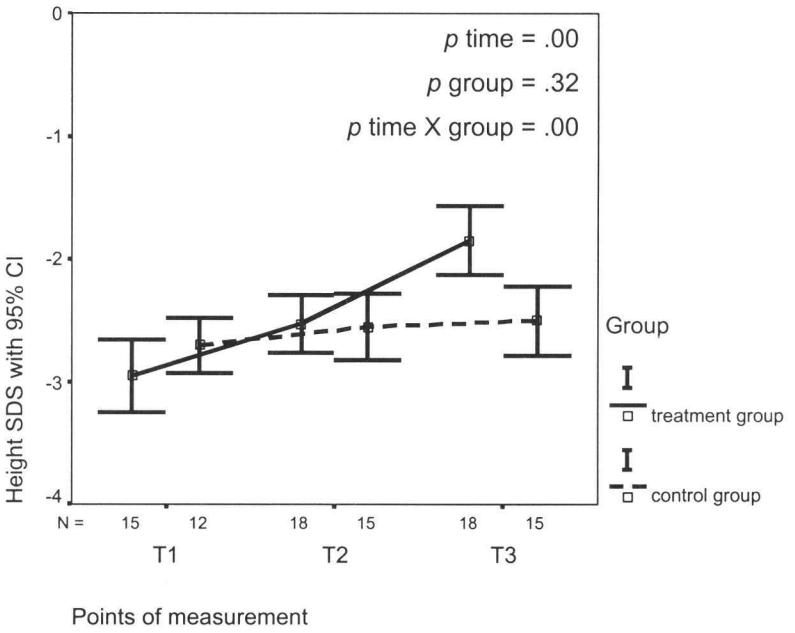
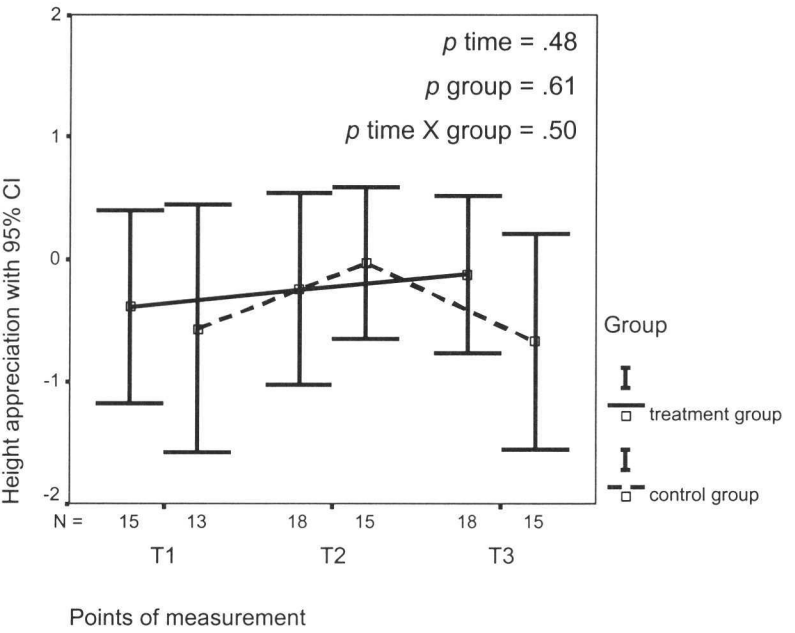


Figure 4. Height appreciation over time, tested by mixed model ANOVA for repeated measurements



10.4.4 HRQoL versus HS

The TACQOL and the ISSQOL presented HRQoL scores (with affective evaluation) as well as HS scores (without affective evaluation), which made comparison possible between these two constructs. In the longitudinal analysis a significant interaction effect was only found for HRQoL on the TACQOL-CF physical complaints scale, but not for HS. The significant group effect for HS in the TACQOL-CF motor functioning scale was not found in HRQoL. The most salient difference was, that the standard deviations of the mean HS scores were in general somewhat lower than those of the HRQoL scores.

10.4.5 Growth expectations, growth and height appreciation versus changes in HRQoL and self-esteem

Table 4 shows that the changes in HRQoL and self-esteem between T2 and T3, as expected, were generally predicted by the scores at T2, which did not differ between groups (see also Table 2). A significant beta-weight means that one standard deviation change of the T2 score of a scale relates to beta units of change in HRQoL or self-esteem. The lower the T2 scores were, the greater the improvement in HRQoL or self-esteem.

Group differences were discussed in a previous section, but Table 4 shows again that greater improvements in HRQoL or self-esteem were related to the control group. Height appreciation, which did not differ significantly between groups, had a stronger relationship with changes in HRQoL than both growth variables. More appreciation for the current height related to improvement in some HRQoL and self-esteem scales. The DUCATQOL physical functioning scale was the only scale that had a significantly relation to height SDS: growth related to improvements in HRQoL. Growth as perceived by the child him or herself related positively to the TACQOL-CF positive emotions scale and the athletic competence, physical appearance and general self-worth scales from the Self-Perception Profile. According to one-way ANOVA, the treatment group perceived their growth significantly higher than the control group did (data not shown). Growth expectations did not contribute to the HRQoL changes between T2 and T3.

Table 4. The independent variables growth expectations, growth (objective and subjective) and height appreciation versus the dependent variable changes in HRQoL and self-esteem between T2 and T3 (reported are the standardised regression weights from linear regression)

				Growth		
Difference in scale score between T2 and T3	Group [a]	Scale score at T2	The child's growth expectations at T2 ('in the next year')	Height SDS difference between T2-T3	Growth at T3 ('since last year') as perceived by the child	The child's height appreciation at T3
TACQOL-CF						
physical complaints	0.79*	-0.66*	--	--	--	--
motor functioning	0.77	-0.69*	--	--	--	--
autonomy	--	-0.74*	--	--	--	--
cognitive functioning	0.76	-0.73*	--	--	--	0.37*
social functioning	1.43*	-0.83*	--	--	--	0.44*
positive emotions	--	-0.72*	--	--	0.57*	--
negative emotions	--	-0.61*	--	--	--	--
ISSQOL-CF						
vitality	--	-0.52*	--	--	--	--
DUCATQOL						
home	--	-0.41*	--	--	--	0.47*
physical	0.69*	-0.76*	--	0.68*	--	0.73*
emotional	--	-0.67*	--	--	--	0.88*
social	1.02	--	--	--	--	--
total QoL	0.61*	-0.71*	--	--	--	0.82*
Self-Perception Profile						
scholastic competence	--	-0.60*	--	--	--	--
social acceptance	0.65	-0.59*	--	--	--	--
athletic competence	--	-0.57*	--	--	0.54*	--
physical appearance	--	-0.67*	--	--	0.55*	0.28
behavioural conduct	--	-0.56*	--	--	--	--
general self-worth	0.65	-0.71*	--	--	0.47*	0.37*

Note: Only beta's with a p-values of .10 and lower are presented; * $p \leq .05$; [a] *group* is the only nominal independent variable: the treatment group was coded as 1 and control group as 2. As a result, a positive relation between group and scale score means higher scores for the control group.

10.4.6 Pros and cons for participation: qualitative data

The Therapy Evaluation Scale (TES) included some open-ended questions, which were subjected to qualitative analysis. On the whole, children often responded that they did not know. With regard to the motivation for participation, it was found that most parents and children participated because the child was too short and needed to grow. During the project, however, the control group parents as well as children mentioned this less often as a reason for participating. From the start, parents from the control group stated more often that they appreciated that their child was closely monitored by professionals, and they mentioned that they started to participate for the benefit of the research project. The treatment group parents emphasised from the start that their child had problems with being short and that they wanted to use every opportunity to let the child grow. Some parents of controls hoped at the start and at T1 that their children could have treatment after all, but thereafter they stopped mentioning this. The children from the control group mentioned with increasing frequency that they did not know why they were participating, or that their parents decided to go on with the project. Parents in the treatment group were optimistic about the growth results and kept participating since the treatment had good results. Children in the treatment group said with increasing frequency that they wanted to grow, partly because they wanted to resolve psychosocial problems (e.g. being teased less). Both groups provided answers indicating that participating in the project was not a burden.

In response to the question about possible drawbacks, half of the total research group, children as well as parents, stated that there were no drawbacks. Nevertheless, some parents from the control group initially feared that participating would put more stress on being short, but this was mentioned less often with time. Children in the control group mentioned having blood samples taken as a drawback, for children in the treatment group it was the daily injections. Parents from the control group mostly mentioned blood sample taking and absence from school as a drawback, whereas the parents from the treatment group mentioned the daily injections and going to the hospital. Parents in the treatment group mentioned uncertainty about possible side-effects of GH and about the final height that could be realised with the treatment.

10.5 Discussion

Treating ISS children with growth hormone does not change the physical health but rather the physical appearance. It was hoped that increasing height would cause people to give more age appropriate reactions to the children, which in turn would improve their HRQoL. In other words, an intervention with GH is based on psychosocial grounds,[3,35] and therefore the psychological variable 'self-esteem' was measured in addition. However, the results of the present study provided only limited support for the hypothesis that short children are disadvantaged.[36] At start, the HRQoL of children with ISS in the treatment and control group was not lower than the HRQoL of a reference group, except for social functioning according to both parents and children. Self-esteem did not differ from the reference group, which accords with other recent work.[36]

It is unclear if these findings can be generalised to all children with ISS, since the study group represented a special group of children who were presented by their parents to a paediatrician because of ISS. Seventy-three percent of subjects were boys, which is typical for studies evaluating short stature.[3,10,12,37] Analyses on sex differences could not be performed because of the small numbers of girls. However, sex and age effects are unlikely because we did not find differences between the results with standardised and raw scores. It is said that especially children who are referred to growth clinics suffer from being short[3,9,21], and that parents who visit a paediatrician for this reason, may be more worried about ISS than parents who do not.[9] These worries could lower the parent's valuation of the child's HRQoL (referral bias).[3,36] Nevertheless, apart from social functioning, parents did not report lower HRQoL than parents from the norm group. Some proof was found for parental disappointment about their children being assigned to the control group, but this disappeared during the study. In another study it was found, that HRQoL was not affected in the adults with ISS sampled from the general population, in contrast to the ISS adults sampled from hospital files.[9] In adults from the hospital files, social functioning appeared to be negatively affected, which corresponds to the results in the present study. In another study, social problems were found in children with GH deficiency, but less in the physically more affected children with achondroplasia.[35] Lack of energy was seen as causing the social disadvantages in a group with GH deficiency.[9] In the present study, reference data for the vitality scale we used was not available. However, although in other studies a relationship was found between GH treatment and energy level,[23-25] the vitality scores

did not improve in the treatment group according to parents and children reports. In contrast, according to the children's reports, vitality improved more in the control group.

Although the paediatrician reported an improvement of HRQoL and self-esteem in the children treated for short stature, the parents reported no change, and the children in the treatment group reported the same or sometimes even worse HRQoL or self-esteem than the control group. This, and the low correlations between parents and children's report of the ISS child's HRQoL, illustrates the low agreement between informants. In other publications this was found as well.[10,35,37,38] Yet, the attitude of parents and significant others is of importance in determining how a child copes with short stature.[20] We found some agreement in what parents and children mentioned as pros and cons for participation, but this was not reflected in general agreement on HRQoL.

One limitation of the study was that the first point of measurement was planned after the random selection. The results at T1 were already influenced by the random selection itself. Parents of the control group indicated temporarily less positive emotions. In addition, results indicate that the paediatrician experienced a difference between groups from T1 onward. Furthermore, subjects in the two groups gave different answers on the retrospective open-ended question about why they started to participate in the study. From the start, parents from the control group emphasised that their child was being closely monitored by professionals, and that they participated for the benefit of the research, whereas the parents from the treatment group stressed that their child had psychosocial problems with his or her height and that they wanted to use every opportunity to promote growth. It is not likely that these differences between groups originated from before the randomisation, although the retrospective question pointed at that period. In the absence of a real baseline measure, we suppose that the differences between groups are the results of psychological changes due to the randomisation.

On the whole, the children in the treatment group did not have better HRQoL than the children in the control group. We hypothesised in advance about the possibility that disappointment or surprise about achieved height might influence HRQoL. Growth expectations of the children in the treatment group could be unrealistically high, since the effect of GH on ISS is not certain. On the other hand, children in the control group may have grown more than they expected at the start. For instance, in the Wessex growth study children who had either short or normal

height at age 5-6 years, showed some overlap in heights at age 12 years.[36] Although in the present study height SDS improved significantly in the treatment group only, height expressed in centimetres improved in both groups, which could be a surprise to children in the control group. Disappointment or surprise about the achieved height could influence HRQoL according to the theory known as 'Calman's Gap'.[39,40] Calman (1984)[40], defined HRQoL by the gap between the patient's expectations and achievements. The smaller the gap between expectations and achievements, the better the HRQoL would be. Lower expectations in the control group could have made the small growth achievements less frustrating. However, the key elements that form Calman's Gap, growth expectations and achieved growth, were either not related or only little to changes in HRQoL or self-esteem. Nevertheless, it is possible that other expectations and achievements than those about growth affect HRQoL. For instance, becoming more independent or participating in the project to help other children could be an achievement as well. Yet, children from the treatment group who expected to achieve psychosocial benefits from their growth, might be disappointed. Furthermore, uncertainty in the treatment group about the final height, even if they grew well at that time, could postpone the feeling that growth was really achieved.

In the present paper, we found that the child's appreciation of the current height but not height itself, related positively to improvements in some HRQoL and self-esteem scales. Another publication reported that shorter children had lower appreciation for their height, which correlated with worse self-esteem concerning body-image.[36] Our observation that there was no difference between groups in the appreciation of height is probably an indication of a psychological phenomenon called coping, the ability to adjust to difficult situations. In our definition of HRQoL we accounted for individual differences as a result of differences in coping. HRQoL was defined as the individual's perception of problems in health status (HS), combined with the affective responses to such problems.[15-17] The standard deviations of the HS group scores were in general somewhat lower than the group scores of the HRQoL. This means that the individual differences between the children with ISS were larger when affective evaluations were taken into account.

It is a matter of concern that some parents in the treatment group were afraid of unknown side-effects. Although the paediatrician stressed that the GH was a safe product, the fear for side-effects of medication is widespread and therefore difficult to eradicate.[41,42]

In conclusion, children with ISS did not have a lower HRQoL and self-esteem than the norm population, except for the HRQoL domain of social functioning as reported by child and parents. The paediatrician reported an improvement of HRQoL in the children treated for short stature, the parents reported no change, and the children in the treatment group reported a similar or sometimes even worse HRQoL than the control group. Differences in HRQoL and self esteem were scarcely related to growth (objectively measured or as perceived by the child). Instead, changes in some HRQoL and self-esteem scales were related to the appreciation of the current height by the child her/himself. Future research could investigate whether these results will hold for children after reaching their final height, and which achievements of these children are important next to growth achievements. The assumption that GH treatment improves HRQoL in children with ISS could not be confirmed in this study.

Acknowledgements
We would like to express our appreciation to Dr. J.J.J. Waelkens, Dr. S.M.P.F de Muick Keizer-Schrama, Prof dr. H.A. Delamarre-van de Waal, Dr. F.Slijper, Dr. J. Huisman, and Drs. J. den Hartog for sharing their knowledge and taking part in data collection. We are indebted to the children and parents for their participation.

10.6 Reference List

1. Ranke MB. Towards a consensus on the definition of idiopathic short stature. Summary. *Horm Res* 1996; **45:** 64-66.
2. Rekers-Mombarg L. Spontaneous growth in idiopathic short stature. *Arch Dis Child* 1996; **75:** 175-180.
3. Voss LD, Wiklund I. Short stature and psychosocial assessment. *Acta Paediatr Suppl* 1995; **411:** 69-74.
4. Stabler B, Clopper RR, Siegel PT, et al. Links between growth hormone deficiency, adaptation and social phobia. *Horm Res* 1996; **45:** 30-33.
5. Huisman J, Slijper FME, Sinnema G, et al. Psychosocial functioning of short prepubertal children and girls with Turner's syndrome. *Acta Paediatr Scand* 1991; **377:** 162-162.
6. Rekers-Mombarg LTM, Busschbach JJV, Massa GG, Dicke J, Wit JM. Quality of life of young adults with idiopathic short stature: effect of growth hormone treatment. *Acta Paediatr* 1998; **87:** 865-870.
7. Voss LD, Bailey BJR, Mulligan J, Wilkin TJ, Betts PR. Short Stature and School Performance: The Wessex Growth Study. *Acta Paediatr Scand* 1991; **Suppl 370:** 29-31.
8. Downie AB, Mulligan J, McCaughey ES, Stratford RJ, Betts PR, Voss LD. Psychological response to growth hormone treatment in short normal children. *Arch Dis Child* 1996; 32-35.
9. Busschbach JJV, Rikken B, Grobbee DE, DeCharro FT, Wit JM. Quality of life in short adults. *Horm Res* 1998; **49:** 32-38.

10. Sandberg DE, Brook AE, Campos SP. Short stature: a psychosocial burden requiring growth hormone therapy? *Pediatr* 1994; **94:** 832-840.
11. Sandberg DE, Michael P. Psychosocial stressses related to short stature: Does their presence imply psychological dysfunction? In: Drotar D, ed. *Measuring health-related quality of life in children and adolescents: implications for research and practice.* Mahwah, New Jersey: Lawrence Erlbaum Associates, 1998: 287-312.
12. Neely EK, Rosenfeld RG. Use and abuse of human growth hormone. *Annu Rev Med* 1994; **45:** 407-420.
13. Wit JM, Kamp GA, Rikken B. Spontaneous growth and response to growth hormone treatment in children with growth hormone deficiency and idiopathic short stature. *Pediatr Res* 1996; **39:** 295-302.
14. Kamp G, Waelkens JJ, Delemarre-Vandewaal HE, DeMuinck Keizer-Schrama SM, Wit JM. Biochemical characterisation of growth hormone sensitivity in idiopathic short stature: Increments of alkaline phosphatase and PIIINP predict the response to GH treatment. *Horm Res* 1999; **51:** 54-54.
15. Verrips GH, Vogels AGC, Koopman HM, et al. Measuring health-related quality of life in a child population. *Eur J Public Health* 1999; **9:** 188-193.
16. Verrips GH, Vogels AGC, Verloove-Vanhorick SP, et al. Health-related quality of life measure for children -- the TACQOL. *J Appl Therapeutics* 1997; **1:** 357-360.
17. Theunissen NCM, Vogels AGC, Koopman HM, et al. The proxy problem: child report versus parent report in health-related quality of life research. *Qual Life Res* 1998; **7:** 387-397.
18. Reijden-Lakeman van der IEA, Slijper FME, Dongen-Melman van JEWM, Waal de WJ, Verhulst FC. Self-concept before and after two years of growth hormone treatment in intrauterine growth-retarded children. *Horm Res* 1996; **46:** 88-94.
19. Pilpel D, Leiberman E, Zadik Z, Carel CA. Effect of growth hormone treatment on quality of life of short- stature children. *Horm Res* 1995; **44:** 1-5.
20. Wiklund I, Wiren L, Erling A, Karlberg J, Albertsson-Wikland K. A new self-assessment questionnaire to measure well-being in children, particularly those of short stature. *Qual Life Res* 1994; **3:** 449-455.
21. Leiberman E, Pilpel D, Carel CA, Levi E, Zadik Z. Coping and Satisfaction with Growth Hormone Treatment among Short-Stature Children. *Horm Res* 1993; **40:** 128-135.
22. Vogels T, Verrips GH, Verloove-Vanhorick SP, et al. Measuring health-related quality of life in children: the development of the TACQOL-Parent Form. *Qual Life Res* 1998; **7:** 457-465.
23. Guyatt GH, Jaeschke R, Feeny DH, Patrick DL. Measurements in clinical trials: Choosing the right approach. In: Spilker B, ed. *Quality of life and pharmacoeconomics in clinical trials.* Philadelphia: Lippincott-Raven Publishers, 1996: 41-48.
24. Wiklund I, Erling A, Albertsson-Wikland K. Critical review of measurement issues in quality of life assessment for children with growth problems. In: Drotar D, ed. *Measuring health-related quality of life in children and adolescents: implications for research and practice.* Mahwah, New Jersey: Lawrence Erlbaum Associates, 1998: 255-271.
25. Sprangers MAG, Cull A, Bjordal K, Groenvold M, Aaronson NK. The European Organization for Research and Treatment of Cancer approach to quality of life assessment: guidelines for developing questionnaire modules. *Qual Life Res* 1993; **2:** 287-295.
26. Wiklund I, Albertsson-Wikland K, Wiren L. How should the Psychosocial impact of Short Stature and Response to Growth Hormone Treatment be Assessed in Children? *Acta Paediatr Scand* 1991; **Suppl.377:** 174-175.
27. Roede MJ, van-Wieringen JC. Growth diagrams 1980. Netherlands third nationwide survey. *Tijdsch Soc Gezondheidszorg* 1985; **63:** 1-34.
28. Dongen-Melman JEWMv, Koot HM, Verhulst FC. Cross-cultural validation of Harter's self-perception profile for children in a Dutch sample. *Educational and Psychological Measurement* 1993; **53:** 739-753.
29. Reijden-Lakeman van der IEA. *Growing Pains? Psychological evaluation of children with short stature after intrauterine growth retardation, before and after two years of growth hormone treatment.* Rotterdam: Thesis Erasmus University Rotterdam, 1996: 1-140.
30. Roede MJ, van-Wieringen JC. Growth diagrams 1980. Netherlands third nationwide survey. Tijdsch Soc Gezondheidszorg 1985; 63: 1-34.

31. Zwinderman AH. Statistical analysis of longitudinal quality of life data with missing measurements. *Qual Life Res* 1992; **1**: 219-224.
32. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures: Statistics and strategies for evaluation. *Control Clin Trials* 1991; **12**: 142S-158S.
33. *SPSS Categories 8.0.* Chicago: SPSS Inc., 1998.
34. Strauss AL. *Qualitative analysis for social scientists.* Cambridge: Cambridge University Press, 1987: 1-319.
35. Haverkamp F, Noeker M. 'Short stature in children - a questionnaire for parents': a new instrument for growth disorder-specific psychosocial adaptation in children. *Qual Life Res* 1998; **7**: 447-455.
36. Downie AB, Mulligan J, Stratford RJ, Betts PR, Voss LD. Are short normal children at a disadvantage? The Wessex growth study. *Brit Med J* 1997; **314**: 97-100.
37. Erling A, Wiklund I, Albertsson-Wikland K. Prepubertal children with short stature have a different perception of their well-being and stature than their parents. *Qual Life Res* 1994; **3**: 425-429.
38. Koopman HM, Theunissen NCM, Vogels AGC, Zwinderman AH, Verloove-Vanhorick SP, Wit JM. Health related quality of life of children with a chronic illness: Parent versus child report. *Gedrag & Gezondheid* 1999; **27**: 118-125.
39. Schipper H, Clinch JJ, Olweny CLM. Quality of life studies: Definitions and conceptual issues. In: Spilker B, ed. *Quality of life and pharmacoeconomics in clinical trials.* Philadelphia: Lippincott-Raven Publishers, 1996: 11-23.
40. Calman KC. Quality of life in cancer patients--an hypothesis. *J Med Ethics* 1984; **10**: 124-127.
41. Horne R. Representations of medication and treatment: Advances in theory and measurement. In: Petrie KJ, Weinman J, eds. *Perceptions of health and illness: current research and applications.* Singapore: Harwood Academic Publishers, 1997: 155-188.
42. Fallsberg M. Reflections on medicines and medication: a qualitative analysis among people on long-term drug regimens. Linköping: Thesis Linköping University, 1991: 1-188.

11 General discussion

N. C. M. Theunissen

In this thesis, we investigated Health related Quality of Life (HRQoL) in children. As a generally accepted definition or theoretical framework is missing, we defined HRQoL as: the individual's perception of problems in health status, combined with the affective responses to such problems. In this definition Health Status (HS) is the assessment by a person of his or her own health. The 'health' component in HRQoL refers to quality of life as a result of a certain state of health, which involves physical, psychological as well as social functioning. The definition of HRQoL incorporates individual and culturally determined differences in coping with HS problems and reflects internal standards about HRQoL, factors which are emphasised by several authors.[1-5] Furthermore, we subscribe to the notion that HRQoL is multi-factorial (physical, psychological and social well-being), patient self-administered, subjective and variable over time.[6,7] Although the studies described in the previous chapters served various purposes, four themes arose repeatedly: the difference between HRQoL and HS, measuring HRQoL in children, the choice of an informant on the child's HRQoL, and the ideas about how much HRQoL is related to health. In the next paragraphs these themes are discussed.

11.1 HRQoL versus HS

Throughout the thesis it was stated that HRQoL differs from HS. In both constructs the person gives a subjective assessment, but in HS he or she describes the *quantity* of problems, and in HRQoL the person describes the *emotional impact* that the problems have on the person's life. When we started studying HRQoL in children in 1994, most available instruments measuring "quality of life" used a HS approach. This is illustrated by the review in Chapter 7, in which half of the longitudinal studies used a HS approach in defining HRQoL as was described in Table 2. This could be explained by the fact that HRQoL is a construct originating from the medical tradition. In that tradition it is unusual to incorporate subjective opinions, which are viewed as unscientific.[8] A construct in which affective evaluations of the patients themselves are included, is therefore even more unusual. As long as children with severe disorders were studied, for instance with cancer, it was considered obvious that the HRQoL of

these children was low. When the HRQoL concept became more popular, children with less severe disorders were studied as well. In these situations it was recognised that children with a minor disorder could live a happy life. This recognition probably created an opening for the study of subjective and individual differences in dealing with HS. Nowadays, scientists operating in the medical tradition have started to recognise that even children with severe disorders can be happy and can have a high HRQoL. According to the psychological tradition, subjective opinions can be studied as long as the measurement method itself is objective and reproducible. The co-operation of psychologists and paediatricians in our research group, resulted in instruments that allowed for HRQoL as well as HS scores. Therefore we could compare these two constructs in this thesis, which produced some interesting results.

Firstly, it was found that sometimes the HRQoL score was better than the HS score, and sometimes poorer. For instance, it was shown in samples from the general population, that only half of the reported HS problems were combined with negative affective responses to those problems (Chapters 2 and 5). As a result, HRQoL is often reported as much better than HS. Furthermore, according to the parents of pre-school children born very preterm, motor functioning including affective responses (=HRQoL) was better than motor functioning excluding affective responses (= HS). In contrast, lungs, stomach and sleeping scores including affective responses were poorer than lungs, stomach and sleeping scores excluding affective responses (Chapter 6). In a group of children with idiopathic short stature the standard deviations of the mean HS scores were in general somewhat lower than that of the HRQoL scores. This means that the differences between children became larger when the affective evaluations were taken into account. (Chapter 10).

Secondly, the relationship between HRQoL and physical health can differ from the relationship between HS and physical health, which was shown in a group of pre-school children with problems with respect to the neonatal period (Chapter 6). Correlations between perinatal factors and motor functioning including affective responses (=HRQoL) were lower and less often significant than between perinatal factors and motor functioning excluding affective responses (=HS).

Finally, agreement between informants was sometimes different for HRQoL and HS. For instance, the global judgement of a neonatologist correlated with HS, but not with HRQoL (both estimated by the parent). It was concluded that the neonatologists did not include the affective evaluation of HS problems in her judgement which the parents did (Chapter 6). Moreover, conceptualisation of HRQoL as a combination of HS and affective evaluation had consequences for agreement between

parent and child as well (Chapter 3). Children may wish to incorporate the fact that they do not feel bad about a certain HS problem by rating their HS problem as more or less severe than a proxy informant such as a parent, or a doctor would. According to previous studies, agreement seemed to be relatively good for observable measures.[9-11] Affective evaluations are probably less observable for parents than HS. Therefore, agreement between parent and child on HRQoL could be expected to be lower than agreement on HS. Indeed, HRQoL agreement on motor functioning and autonomy was significantly lower than HS agreement. However, HRQoL agreement matched HS agreement on the other scales, which indicates that adding affective evaluation does not influence observability. Perhaps these HS scales already have a strong subjective component: when children or parents report about the quantity of problems in, for instance, social functioning, there subjective evaluation might influence the counting of these problems. By adding affective evaluation to HRQoL, agreement on motor functioning and autonomy became poor, but the level of subjectiveness became probably distributed more even among the HRQoL scales. The possibility that patients have a health problem but do not feel bad about it may bias patient's self-reporting in typical HS questionnaires.

Conclusions and recommendations

In this thesis it was shown that the distinction between HRQoL and HS is justified. Consequently, the HS of a cohort of children born preterm (Chapter 9) cannot be equated to the HRQoL of these children. If it matters how children feel about their functioning rather than how they are functioning, measuring HS alone does not provide all relevant information. Children as well as parents distinguish between the HS problems they observe, and the emotional evaluation of these problems incorporated in HRQoL.

11.2 Measuring HRQoL in children

Instruments that measure HRQoL in children should be multi-factorial (physical, psychological and social well-being), self-administered and subjective. In this thesis four generic HRQoL questionnaires were used that met these requirements. Two instruments were exhaustively described, the TACQOL designed for children age 6 to 15 years (Chapters 2 and 3) and the TAPQOL designed for children 1 to 5 years (Chapter 5). Two other generic instruments were added for comparison, DUCATQOL

(Chapter 10) and KINDL (Chapter 2). The multi-factorial structure of the TACQOL and the TAPQOL was confirmed by statistical analyses. The instruments had good reliability and validity (Chapters 2, 3 and 5). Validity was extensively tested by Multi-Trait Multi-Method modelling on the TACQOL scores of 8-11 year olds and their parents (Chapter 3). The multi-factorial structure was not only confirmed by statistical analyses, it appeared useful as well. Children reported different HRQoL at different scales. For instance, children born very preterm had a low HRQoL in different scales to children with other problems in the neonatal period (Chapter 6). Furthermore, children with idiopathic short stature had the same HRQoL as children from a reference group, except for social functioning in which the HRQoL was lower (Chapter 10).

The life of 6 to 15 year old children differs from that of 1 to 5 year olds, which has to be reflected in the content of HRQoL instruments (Chapter 7). The TACQOL contains scales which are different to those of the TAPQOL, both in objective and in items. For instance, the TACQOL contained items like 'riding a bicycle ' and 'reading' (Chapter 2), which are irrelevant to pre-schoolers. Instead, the sleeping scale in the TAPQOL appeared relevant for children age 1 to 5 years (Chapter 2). The TACQOL's cognitive functioning scale contains schoolability items, the TAPQOL's communication scale measures cognitive skills that are considered relevant for pre-school children. Within the TAPQOL, some scales are suitable for children of at least one-year-and-a-half (Chapters 2 and 3). Within the TACQOL age range, there were minor differences between scores of various age groups (Chapters 3 and 10).

Although the items from the TACQOL and TAPQOL appeared relevant, the fact remains that the construction followed a top-down procedure. The items were selected by adult investigators. Children's knowledge about health and disease is age- and experience-related and therefore different from the knowledge of adults.[12,13] A bottom-up procedure in which the items are derived from the children themselves perhaps would give a different result.

The feasibility of instruments for self administering HRQoL by the child is limited by the cognitive skills of the child.[14] Therefore the TAPQOL uses the parent as informant. The TACQOL has a parent form to inquire about the HRQoL of children age 6-15 years old, and a child form which can be filled in by the child hem/herself at age 8-15 years. As child and parent have different views on the child's HRQoL (see next paragraph), switching between informants is advised against. Consequently, when conducting a longitudinal study from 1 to 12 years of age, the parent should be main informant. But even when the same informant is used throughout the study, the content

of the instruments changes with age, which may hamper longitudinal studies. However, in this thesis a strategy was presented that could overcome the problem of changing measurement instruments between time points (Chapter 8).

Conclusions and recommendations

In the thesis it was shown, that one-third of the publications that studied longitudinal HRQoL in children, used instruments without clear-cut measurement properties (Chapter 7). This should not be necessary in future research, because currently there are many good generic HRQoL instruments for children that can be used [15-19], and four are presented in this thesis. Most HRQoL instruments are paper and pencil questionnaires. As the use of these questionnaires is limited in young children, it would be interesting to develop other instruments like observation systems or interviews. Nevertheless, it can be concluded that the scientific tools are available to measure HRQoL in children. As stated before, the results have shown that 'subjectivity can be made scientific'.[20]

11.3 Informant of HRQoL

Although HRQoL aims for the individual's perception, children cannot always serve as informant. They may lack the cognitive skills or could be too ill to fill in questionnaires. In that case someone has to act as a proxy, for instance the parent or the physician.[9-11,21-25] Proxies may not have the same knowledge and internal standard as the child him or herself, which can influence their reports.[26] For instance, the HRQoL that parents reported about their pre-school children, related to the feelings that the parent had towards their child (Chapter 6). It is not clear if the reported HRQoL of the pre-school child was the result or the cause of the parent's feelings towards this child. Unfortunately we do not have instruments to obtain the child's HRQoL from the pre-school children themselves. We could, however, get an impression of the differences between child and parent report, in a sample of children 8 to 11 years of age and their parents.

Parent reports were only moderately correlated to child reports in the general population (Chapters 2 and 3). Yet, both child- and parent reports proved to be valid, as was described in the previous paragraph. The mean differences between the TACQOL scores of children and parents differed between studies. In the open population, children reported significantly lower HRQoL than their parents on the

physical complaints, motor functioning, autonomy, cognitive functioning and positive emotions scales (Chapter 3). In a group of children with a chronic disorder, children reported significantly lower HRQoL on physical complaints, motor functioning and positive emotions, but higher HRQoL than their parents on social functioning and negative emotions (Chapter 4). In a group of children with idiopathic short stature, the children reported significantly lower HRQoL than their parents on physical complaints at time 2, higher HRQoL on cognitive and social functioning at time 2 and time 3 and higher HRQoL on negative emotions (=less negative emotions reported) at time 3 (Chapter 10). Since the differences between parent and child at group level depend on the sample, it is not possible to give a formula for 'translating' parent scores into child scores. In two samples (Chapters 3 and 4) we found that agreement relates to height of the HRQoL scores. Child scores appear to be less extreme than parent scores. When parents are very pessimistic, children seem to say "it isn't so bad", and when parents are very optimistic, children seem to say "it isn't that good". One might consider parents failing as informants about the child and children as lacking a time perspective, but still, both child and parent reports proved to be valid.

As the parent is close to the child, the parent is the preferable proxy. In a clinical situation, however, the treatment program is established in concordance between physician and parent, if possible in consultation with the child. Therefore it is important to study the agreement between the physician and the parent, and between the physician and the child. This agreement cannot be estimated during a regular consultation. Developmental psychology has shown that children are used to treating adults as their tutors; when a question is asked they will refer to an adult to learn if their answer is correct.[27] As a result the opinion of the child is highly influenced by the presence of a parent or a physician. However, even parents can be highly influenced by the social status of the physician during the consultation.[28] This may give the impression to the physician that agreement with the parent is higher than it is in reality, which could have consequences for parental satisfaction with the consultation.

In this thesis HRQoL information is obtained from physicians, independently of the information obtained from parents and children. In a study about the HRQoL of pre-school children born preterm, it was found that parents noticed many motor problems(=HS). However, the parents considered these problems not to be of great emotional impact (=HRQoL). The judgement of the neonatologist related to the HS but not to the HRQoL as obtained by the parents. As a result, in a clinical situation the neonatologist is surprised that a parent does not want to have full treatment for the

child's motor functioning problems. The parent simply does not consider the motor problems to be as serious as the neonatologist does. In reverse, parents evaluated problems with lungs, stomach and sleeping as being more of an emotional burden than the neonatologist does. The parent therefore does not understand why these problems in their child receive less attention from the neonatologist.

Child, parent and physician's perception on the child's HRQoL was studied in a group of children with idiopathic short stature. Half of the group was treated with growth hormone, the other half acted as a control group (Chapter 10). It was found that the pattern of longitudinal changes differed between child, parent and physician. The physician reported an improvement of HRQoL in the children treated for short stature, the parents reported no change, whereas according to the children themselves the treatment group had the same or sometimes even poorer HRQoL than the control group.

Conclusions and recommendations

It appeared that in healthy as well as chronically ill children, the children and their proxies assessed HRQoL differently. It would be interesting to investigate what the effect is of feedback to the informants about their disagreement. Furthermore, if the observability is influencing agreement, studies are needed in future about what exactly children and their proxies observe. As no gold standard exists and both child's and parent's opinion were valid, it seems best to obtain both parent's and children's evaluations whenever possible. The judgement of the physician should be obtained in addition to help clinical communication and decision making.

11.4 Is HRQoL health related or not?

As is indicated by the term *Health Related* Quality of Life, a relation is implied between health and HRQoL. This relation can be studied in a cross-sectional design by testing the hypothesis that children with a health problem had poorer HRQoL than healthy children. In a longitudinal design the hypothesis could be formulated as: changes in the child's health can result in changes in HRQoL, or, improved health will give improved HRQoL. In the introduction, health is defined as 'a state of complete physical, mental, and social well being, and not merely the absence of disease or infirmity' (WHO,1948)[29]. This implies that children with a physical, mental or social health problems could be subject of study. In this thesis we limited ourselves to studies

in children with a physical health problem. Partly because the three kinds of health problems each require different knowledge and approaches, partly because HRQoL is a construct that is specifically popular in the medical world where people with a physical health condition are studied. In this thesis some support was found for the hypothesis that children with a physical health problem had poorer HRQoL than healthy children. In a group of 6 to 11 year old children from the open population, children with a chronic illness, children who had undergone medical treatment, and even children who had a common illness (a cold or influenza) had significantly lower HRQoL than healthy children (Chapter 2). In a group of 1 to 5 year old children from the open population again it was found that children with a chronic illness (mostly children with respiratory problems) had poorer HRQoL than healthy children (Chapter 5). Furthermore, children born preterm had lower HRQoL than a healthy reference group (Chapters 5 and 6).

Even if studies are limited to children with a physical health problem, as a rule, changes in HRQoL can be the result of physical, psychological as well as social changes (Chapter 7). Investigations that endorse the importance of this rule, use a biopsychosocial model of change in HRQoL.[30] This model recognises that health and HRQoL are determined by psychological and social as well as physical factors, all of which interact to produce the current HRQoL. To study these changes the research design must allow for collecting psychological and social variables along with physical parameters. Furthermore, medical treatments as well as psychological interventions could be beneficial in changing HRQoL. Even if a medical treatment is the objective, assessment of HRQoL has to be planned not only in relation to the physical effects of this medical intervention, but also in relation to the psychosocial effect. For instance, in addition to side effects of medications, also school absenteeism or not being able to see friends are considered important factors that could elicit changes in HRQoL. Five out of the 32 longitudinal studies reviewed in Chapter 7, used a biopsychosocial model of change.[31-35] One of the studies in this thesis described the HRQoL of children with idiopathic short stature (ISS), which is short stature without an underlying disease or deficiency (Chapter 10). By treating these children with growth hormone, it is not their physical health which is changed but rather their physical appearance. It was hoped that increasing height would give more age appropriate reactions to the children, which in turn would improve their HRQoL. In other words, not physical health but psychological or social health was object of study, and therefore the psychological variable 'self-esteem' was measured in addition. At start the HRQoL of children with

ISS was not lower than the HRQoL of a reference group, except for social functioning. Self-esteem did not differ from the reference group. This could indicate that ISS is more a social than a physical or psychological problem. Although the physician reported an improvement in HRQoL in the children treated for short stature, the parents reported no change, and the children in the treatment group reported the same or sometimes even poorer HRQoL than the control group. Furthermore, it was found that changes in HRQoL of these children did hardly relate to growth (objectively measured or as perceived by the child). Instead, changes in some HRQoL and self-esteem scales were related to the height appreciation by the child her/him self. The appreciation of height did not differ between groups, which is an indication of a psychological phenomenon called coping, the ability to adjust to difficult situations.

In the foregoing paragraphs, the ability of children to change in HRQoL over time is stressed, and changes in physical, psychological or social health are seen as the motor of change. We call this the plasticity approach of change in HRQoL. Another possible approach to change in HRQoL is the predictability approach. This is change defined as the maintenance of relative position on particular characteristics over time, which can denote both stability (absolute levels of HRQoL remain stable over time), as well as continuity (consistency in relative rank over time on HRQoL). Both approaches are extensively discussed in Chapter 7.

The thesis contains a study illustrating the predictability approach as well. It involved a longitudinal study in children born preterm between 5 and 10 years of age (Chapter 9). Problems in basic functioning decreased while negative moods increased, and concentration problems increased slightly. The changes in HS found in this study were mainly age-related instead of health-related. The most remarkable HS change was the decrease in basic functioning problems in children with more than one handicap. The results of this study support the idea of continuity in change, in which a consistency in relative rank over time could be found. The normal development of children at a certain age was reflected in the type of problems that the parents reported. Basic functioning is important from early childhood onwards. At 'kindergarten age', concentration problems become more important. When children grow older they start to communicate more about their moods. Therefore, parents could successively detect problems in basic functioning, concentration problems and negative moods. As a result, at the age of 10 years the parents could have grown accustomed to the first, less to the second, but not to the last kind of HS problems. In this study changing parental standards appeared to be the motor of change. Although these results are about HS, these probably would have been found in HRQoL as well. It appeared that the priorities

and goals of the parents changed with time and were modified by the age of the child and experience with the child. These changes narrow Calman's Gap[7,36], the gap that Calman positioned between the patient's expectations and achievements or possibilities. The better the gap is closed the higher the QoL should be.

A review in 32 longitudinal studies about HRQoL in children -- half of the time defined like we defined HS -- revealed the presumption that stable physical health gives stable HRQoL and that changes in physical health change the HRQoL (Chapter 7). This presumption appeared so strong that most investigators did not feel the need to prove it. However, as shown above, not only changes in physical health, but also changes in psychological and social health, and changing internal standards of the informant could influence changes in HRQoL.

Conclusions and recommendations

The two terms HRQoL and QoL are often intermingled in publications, even in a standard work such as Spilker's 'Quality of Life and pharmacoeconomics in clinical trials'.[37] To avoid the impression that the terms HRQoL and QoL describe different constructs, it might be better to choose between these terms in the future. It is important to realise that when the term *health related* QoL is used, psychological and social health should be considered as well as physical health. All three kinds of health should be considered when studying differences in HRQoL between groups, or seeking the cause of changes in HRQoL. The notion that not only physical health is relevant in HRQoL, requires more research about processes that influence psychosocial health like coping and adaptation. This implies that when changes in physical health cannot be realised in children with a chronic illness, changes in psychological and social health may be able to improve the HRQoL of children.

11.5 Reference List

1. Ebrahim S. Clinical and public health perspectives and applications of health-related quality of life measurement. *Soc Sci Med* 1995; 41: 1383-1394.
2. Bullinger M, Ravens-Sieberer U. Stand der Forschung zur gesundheitsbezogenen Lebensqualität von Kindern: eine Literaturanalyse [German: State of the art of health-related quality of life research in children: a literature review]. *Präv Rehab* 1995; 7: 106-121.
3. Lindström B. Measuring and improving quality of life for children. In: Lindström B, Spencer N, eds. *Social Paediatrics.* Oxford: Oxford University Press, 1995: 571-585.
4. Rosenberg R. Health-related quality of life between naturalism and hermeneutics. *Soc Sci Med* 1995; 41: 1411-1415.
5. Amir M, Bar-On DM, Penso R. Positive-Negative Evaluation (PNE) scale: A new dimension of the subjective domains of quality of life measure. *Qual Life Res* 1996; 5: 73-80.

6. Leventhal H, Colman S. Quality of life: A process view. *Psychology and Health* 1997; **12:** 753-767.
7. Schipper H, Clinch JJ, Olweny CLM. Quality of life studies: Definitions and conceptual issues. In: Spilker B, ed. *Quality of life and pharmacoeconomics in clinical trials.* Philadelphia: Lippincott-Raven Publishers, 1996: 11-23.
8. Buchanan N. The efficacy of lamotrigine on seizure control in 34 children, adolescents and young adults with intellectual and physical disability. *Seizure* 1995; **4:** 233-236.
9. Magaziner J, Simonsick EM, Kashner TM, Hebel JR. Patient-proxy response comparability on measures of patient health and functional status. *J Clin Epidemiol* 1988; **41:** 1065-1074.
10. Epstein AM, Hall JA, Tognetti J, Son LH, Conant L. Using proxies to evaluate quality of life. *Med Care* 1989; **27:** S91-S98.
11. Hays RD, Vickrey BG, Hermann BP, et al. Agreement between self reports and proxy reports of quality of life in epilepsy patients. *Qual Life Res* 1995; **4:** 159-168.
12. Koopman HM. *De kijk van kinderen: Een cognitief structurele visie op de ontwikkeling van het denken over ziekte [Dutch: the children's view: a cognitive structural approach to the development of thinking about illness].* Amsterdam: Thesis Publishers, 1993.
13. Hosli EJ. *Ziektebegrip bij kinderen: de ontwikkeling van het denken over ziekte [Dutch: Understanding illness: The development of children's thinking about illness].* Utrecht: Thesis Utrecht University, 1998: 1-263.
14. Juniper EF, Guyatt GH, Feeny DH, Griffith LE, Ferrie PJ. Minimum skills required by children to complete health-related quality of life instruments for asthma: comparison of measurement properties. *European Respiratory Journal* 1997; **10:** 2285-2294.
15. Levi R, Drotar D. Critical issues and needs in health-related quality of life assessment of children and adolescents with chronic health conditions. In: Drotar D, ed. *Measuring health-related quality of life in children and adolescents: implications for research and practice.* Mahwah, New Jersey: Lawrence Erlbaum Associates, 1998: 3-20.
16. Spieth LE, Harris CV. Assessment of health-related quality of life in children and adolescents: An integrative review. *J Pediatr Psychol* 1996; **21:** 175-193.
17. Pal DK. Quality of life assessment in children: A review of conceptual and methodological issues in multidimensional health status measures. *Journal of Epidemiology and Community Health* 1996; **50:** 391-396.
18. Bullinger M, Ravens-Sieberer U. Health related QOL assessment in children: A review of the literature. *European Review of Applied Psychology/Revue Européenne de Psychologie Appliquée* 1995; **45:** 245-254.
19. Bruil J. *Development of a quality of life instrument for children with a chronic illness.* Leiden: Doctoral dissertation Leiden University, 1999.
20. O'Boyle C. Making subjectivity scientific. *Lancet* 1995; **345:** 602-602.
21. Bellman MH, Paley CE. Pain control in children: Parents underestimate children's pain. *Brit Med J* 1993; **307:** 1563-1563.
22. Zivcic I. Emotional reactions of children to war stress in Croatia. *J Am Acad Child Adolesc Psychiatry* 1993; **32:** 709-713.
23. Achenbach TM, McConaughy SH, Howell CT. Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychol Bull* 1987; **101:** 213-232.
24. Sprangers MAG, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: A review. *J Clin Epidemiol* 1992; **45:** 743-760.
25. Ennett ST, DeVellis BM, Earp JA, Kredich D, Warren RW, Wilhelm CL. Disease experience and psychosocial adjustment in children with juvenile rheumatoid arthritis: Children's versus mothers' reports. *J Pediatr Psychol* 1991; **16:** 557-568.
26. Landgraf JM, Abetz LN. Functional status and well-being of children representing three cultural groups: Initial self-reports using the CHQ-CF87. *Psychology and Health* 1997; **12:** 839-854.

27. Elbers E. Interactie tussen volwassenen en kinderen: Theorie, onderzoek, toepassing [Dutch: Interaction between adults and children: Theorie, reseach, application]. In: Meeuwesen L, Houtkoop-Steenstra H, eds. *Studies naar mondelinge interacties [Dutch: Studies about verbal interactions].* Utrecht: ISOR, 1995: 13-30.
28. Bensing J. *Doctor-patient communication and the quality of care: An observation study into affective and instrumental behavior in general practice.* Utrecht: NIVEL (Thesis Erasmus University Rotterdam), 1991: 1-243.
29. World Health Organization. *Constitution of the World Health Organization.* Geneva: World Health Organization, 1948.
30. Finkelstein JW. Methods, models, and measures of health-related quality of life for children and adolescents. In: Drotar D, ed. *Measuring health-related quality of life in children and adolescents: implications for research and practice.* Mahwah, New Jersey: Lawrence Erlbaum Associates, 1998: 39-52.
31. Parkin PC, Kirpalani HM, Rosenbaum PL, et al. Development of a health-related quality of life instrument for use in children with spina bifida. *Qual Life Res* 1997; **6:** 123-132.
32. Marrero DG, Vandagriff JL, Kronz K, et al. Using telecommunication technology to manage children with diabetes: the Computer-Linked Outpatient Clinic (CLOC) Study. *Diabetes Educ* 1995; **21:** 313-319.
33. Bartholomew LK, Czyzewski DI, Parcel GS, et al. Self-management of cystic fibrosis: short-term outcomes of the Cystic Fibrosis Family Education Program. *Health Educ Behav* 1997; **24:** 652-666.
34. Kazak AE, Penati B, Boyer BA, et al. A randomized controlled prospective outcome study of a psychological and pharmacological intervention protocol for procedural distress in pediatric leukemia. *J Pediatr Psychol* 1996; **21:** 615-631.
35. Aldenkamp AP. Psychosocial adaptation after normalisation of physical symptomatology: Cognitive complaints of children and their parents after seizure remission: In search of the time-delay model. Special Issue: Youth and disability in Holland. *Int J Adolesc Med H* 1994; **7:** 303-315.
36. Calman KC. Quality of life in cancer patients--an hypothesis. *J Med Ethics* 1984; **10:** 124-127.
37. Spilker B. *Quality of life and pharmacoeconomics in clinical trials.* (2th edition) Philadelphia: Lippincott-Raven-Publishers, 1996: 1-1259.

References

Aaronson, N.K. (1988). Quality of life: what is it? How should it be measured? *Oncology Huntingt., 2*(5), 69-76,64.

Aaronson, N.K. (1989). Quality of life assessment in clinical trials: methodologic issues. *Controlled Clinical Trials, 10*(4 Suppl), 195S-208S.

Achenbach, T.M., McConaughy, S.H., & Howell, C.T. (1987). Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*(2), 213-232.

Aldenkamp, A.P. (1994). Psychosocial adaptation after normalisation of physical symptomatology: Cognitive complaints of children and their parents after seizure remission: In search of the time-delay model. Special Issue: Youth and disability in Holland. *International Journal of Adolescent Medicine and Health, 7*(4), 303-315.

Amir, M., Bar-On, D.M., & Penso, R. (1996). Positive-Negative Evaluation (PNE) scale: A new dimension of the subjective domains of quality of life measure. *Quality of Life Research, 5*, 73-80.

Apajasalo, M., Sintonen, H., Holmberg, C., Sinkkonen, J., Aalberg, V., Pihko, H., Siimes, M.A., Kaitila, I. , Makela, A., Rantakari, K., Anttila, R., & Rautonen, J. (1996). Quality of life in early adolescence: a sixteen-dimensional health-related measure (16D). *Qual.Life Res., 5*(2), 205-211.

Appelbaum, M.I., Burchinal, M.R., & Terry, R.A. (1989). Quantitative methods and the search for continuity. In M. H. Bornstein & N. A. Krasnegor (Eds.), *Stability and continuity in mental development.* (pp. 251-272). Hillsdale: Lawrence Erlbaum Associates.

Baker, G.A., Smith, D.F., Dewey, M., Jacoby, A. , & Chadwick, D.W. (1993). The initial development of a health-related quality of life model as an outcome measure in epilepsy. *Epilepsy Research, 16*(1), 65-81.

Bancalari, E., Abdenour, G.E., Feller, R., & Gannon, J. (1979). Bronchopulmonary dysplasia: clinical presentation. *Journal of Pediatrics, 95*(5 Pt 2), 819-823.

Barr, R.D., Petrie, C., Furlong, W., Rothney, M., & Feeny, D. (1997). Health-related quality of life during post-induction chemotherapy in children with acute lymphoblastic leukemia in remission: An influence of corticosteroid therapy - Clinical Study. *International Journal of Oncology, 11*(2), 333-339.

Bartholomew, L.K., Czyzewski, D.I., Parcel, G.S., Swank, P.R., Sockrider, M.M., Mariotto, M.J., Schidlow, D.V., Fink, R.J., & Seilheimer, D.K. (1997). Self-management of cystic fibrosis: short-term outcomes of the Cystic Fibrosis Family Education Program. *Health Educ.Behav , 24*(5), 652-666.

Becker, N., Lombardi, P., Sidoti, E., & Katkin, L.S. (1988). Mylicon drops in the treatment of infant colic. *Clin.Ther., 10* (4), 401-405.

Becker, P.T., Engelhardt, K.F., Steinmann, M.F., & Kane, J. (1997). Infant age, context, and family system influences on the interactive behavior of mothers of infants with mental delay. *Research in Nursing and Health, 20*(1), 39-50.

Bellman, M.H., & Paley, C.E. (1993). Pain control in children: Parents underestimate children's pain. *British Medical Journal, 307*, 1563-1563.

Bensing, J. (1991). *Doctor-patient communication and the quality of care: An observation study into affective and instrumental behavior in general practice.* Utrecht: NIVEL (Thesis Erasmus University Rotterdam).

Bergner, M., Bobbitt, R.A., Carter, W.B., & Gilson, B.S. (1981). The Sickness Impact Profile: Development and final revision of a health status measure. *Medical Care, 19*(8), 787-805.

Berth-Jones, J., Finlay, A.Y., Zaki, I., Tan, B., Goodyear, H., Lewis-Jones, S., Cork, M.J., Bleehen, S.S., Salek, M.S., Allen, B.R., Smith, S., & Graham-Brown, R.A. (1996). Cyclosporine in severe childhood atopic dermatitis: a multicenter study. *J Am Acad Dermatol, 34*(6), 1016-1021.

Bland, J.M., & Altman, D.G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet,* (february 8), 307-310.

Bland, J.M., & Altman, D.G. (1997). Cronbach's alpha. *BMJ., 314*(7080), 572.

Bosch, J.D., Veerman, J.W., Samson, J.F., & Smith, F. (1995). Ex-neonatale couveusepatieentjes vijf en tien jaar later. Een follow-up-onderzoek naar mogelijke gedragsproblemen [Dutch: Ex- neonatal intensive-care-unit (NICU) patients five and ten years later: A follow-up study of risk for behavioral problems]. *Kind en Adolescent, 16*(1), 1-15.

References

Bowling, A., & Brazier, J. (1995). 'Quality of Life' in social science and medicine: Introduction. *Social Science & Medicine, 41*(10), 1337-1338.

Bradburn, N.M. (1969). *The structure of psychological well-being.* Chicago: Aldine Publishing Company.

Browne, J.P., McGee, H.M., & O'Boyle, C.A. (1997). Conceptual approaches to the assessment of quality of life. *Psychology and Health, 12*(6), 737-751.

Brugman, E., Goedhart, H., Vogels, T., & van-Zessen, G. (1995). *Jeugd en seks [Dutch: Youth and sex].* Utrecht: Uitgeverij SWP .

Bruil, J., & Maes, S. (1995). Assessing quality of life among children with a chronic illness: The development of a questionnaire [Abstract]. *Poster presented at the first Dutch conference on Psychology & Health, Kerkrade, The Netherlands, 6-8 November,* 1-1.

Bruil, J. (1999). *Development of a quality of life instrument for children with a chronic illness.* Leiden: Doctoral dissertation Leiden University.

Bruinse, H.W., & Cats, B.P. (1994). *Te vroeg op de wereld: Over prematuren in Nederland [Dutch: Too early in this world: about prematures in the Netherlands].* Utrecht: Lemma BV.

Buchanan, N. (1995). The efficacy of lamotrigine on seizure control in 34 children, adolescents and young adults with intellectual and physical disability. *Seizure, 4*(3), 233-236.

Bullinger, M., & Hasford, J. (1991). Evaluating quality-of-life measures for clinical trials in Germany. *Controlled Clinical Trials, 12*(4 Suppl), 91S-105S.

Bullinger, M., von-Mackensen, S., & Kirchberger, I. (1994). Ein Fragebogen zur Erfassung der Gesundheitsbezogene Lebensqualitaet von Kindern [German: a questionnaire for assessing the health related quality of life of children]. *Zeitschrift für Gesundheitspsychologie, 2*(1), 64-77.

Bullinger, M., & Ravens-Sieberer, U. (1995). Stand der Forschung zur gesundheitsbezogenen Lebensqualität von Kindern: eine Literaturanalyse [German: State of the art of health-related quality of life research in children: a literature review]. *Prävention und Rehabilitation, 7*(3), 106-121.

Bullinger, M., & Ravens-Sieberer, U. (1995). Health related QOL assessment in children: A review of the literature. *European Review of Applied Psychology/Revue Européenne de Psychologie Appliquée , 45*(4), 245-254.

Busschbach, J.J.V., Rikken, B., Grobbee, D.E., DeCharro, F.T., & Wit, J.M. (1998). Quality of life in short adults. *Hormone Research, 49*(1), 32-38.

Calman, K.C. (1984). Quality of life in cancer patients--an hypothesis. *J.Med.Ethics., 10*(3), 124-127.

Campbell, D.T., & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Cantani, A., Arcese, G., Lucenti, P., Gagliesi, D., & Bartolucci, M. (1997). A three-year prospective study of specific immunotherapy to inhalant allergens: evidence of safety and efficacy in 300 children with allergic asthma. *J Investig Allergol Clin Immunol, 7*(2), 90-97.

Caputo, D.V., Goldstein, K.M., & Taub, H.B. (1981). Neonatal compromise and later psychological development: A 10-year longitudinal study. In S. L. Friedman & M. Sigman (Eds.), *Preterm birth and psychological development.* (pp. 353-385). New York, etc: Academic Press.

Carpay, H.A., Arts, W.F., Vermeulen, J., Stroink, H., Brouwer, O.F., Peters, A.C., Van-Donselaar, C.A., & Aldenkamp, A.P. (1996). Parent-completed scales for measuring seizure severity and severity of side-effects of antiepileptic drugs in childhood epilepsy: development and psychometric analysis. *Epilepsy Research, 24*(3), 173-181.

Cella, D.F., & Tulsky, D.S. (1993). Quality of life in cancer: Definition, purpose, and method of measurement. *Cancer Investigation, 11*(3), 327-336.

Christie, M.J., French, D., Weatherstone, L., & West, A. (1991). The patients' perception of chronic disease and its management: psychosomatics, holism and quality of life in contemporary management of childhood asthma. *Psychother.Psychosom., 56*, 197-203.

Cleary, P.D., Morrissey, G., Yver, A., & Oster, G. (1994). The effects of rG-CSF on health-related quality of life in children with congenital agranulocytosis. *Quality of Life Research, 3*(5), 307-315.

Cleary, P.D. (1996). Future directions of quality of life research. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trails.* (pp. 73-78). Philadelphia: Lippincott-Raven Publishers.

Cohen, S.E., & Parmelee, A.H. (1983). Prediction of five-year Stanford-Binet scores in preterm infants. *Child Development , 54*(5), 1242-1253.

Collings, J.A. (1990). Psychosocial well-being and epilepsy: an empirical study. *Epilepsia, 31*(4), 418-426.

Collings, J.A. (1990). Epilepsy and well-being. *Social Science & Medicine, 31* (2), 165-170.

Croyle, R.T., & Uretsky, M.B. (1987). Effects of mood on self-appraisal of health status. *Health Psychol, 6* (3), 239-253.

Croyle, R.T., & Ditto, P.H. (1990). Illness cognition and behavior: an experimental approach. *J Behav.Med., 13* (1), 31-52.

De-Leeuw, R., & Bakker, L. (1989). *Te vroeg geboren [Dutch: Born too early].* Ede: Zomer en Keuning Boeken B.V.

den-Ouden, A.L., Rijken, M., Brand, R., Verloove-Vanhorick, S.P., & Ruys, J.H. (1991). Is it correct to correct? Developmental milestones in 555 "normal" preterm infants compared with term infants. *Journal of Pediatrics, 118* (3), 399-404.

den-Ouden, A.L., Hille, E.T.M., Bauer, L., & Verloove-Vanhorick, S.P. (1993). School Performance in very preterm children: a report from the collaborative project on Preterm and Small for gestational age (POPS) infants in the Netherlands. *The Lancet, 342* (August), 550-551.

Deyo, R.A., Diehr, P., & Patrick, D.L. (1991). Reproducibility and responsiveness of health status measures: Statistics and strategies for evaluation. *Controlled Clinical Trials, 12*, 142S-158S.

Donadieu, J., Boutard, P., Tchernia, G., Oster, G., Gordon, S.E., Philippe, N., Le-Galle, E., Nivelon, J.L., Dopfer, R., Babin, B.A., Coulombel, L., Pocidalo, M.A., Yver, A., & Griscelli, C. (1994). [French: A phase II study of recombinant human granulocyte-colony stimulating factor (rHuG-CSF, lenograstim) in the treatment of agranulocytosis in children]. *Nouv.Rev.Fr.Hematol., 36* (6), 441-448.

Donadieu, J., Boutard, P., Bernatowska, E., Tchernia, G., Couillaud, G., Philippe, N., & Le-Galle, E. (1997). A European phase II study of recombinant human granulocyte colony-stimulating factor (lenograstim) in the treatment of severe chronic neutropenia in children. Lenograstim Study Group. *European Journal of Pediatrics, 156* (9), 693-700.

Donahue, M.L., & Pearl, R. (1995). Conversational interactions of mothers and their preschool children who had been born preterm. *Journal of Speech and Hearing Research, 38* (5), 1117-1125.

Dongen-Melman, J.E.W.M.v., Koot, H.M., & Verhulst, F.C. (1993). Cross-cultural validation of Harter's self-perception profile for children in a Dutch sample. *Educational and Psychological Measurement, 53*, 739-753.

Dossetor, D.R., Liddle, J.L.M., & Mellis, C.M. (1996). Measuring health outcome in paediatrics: Development of the RAHC measure of function. *Journal of Paediatrics and Child Health, 32* (6), 519-524.

Downie, A.B., Mulligan, J., McCaughey, E.S., Stratford, R.J., Betts, P.R., & Voss, L.D. (1996). Psychological response to growth hormone treatment in short normal children. *Archives of Disease in Childhood,* 32-35.

Downie, A.B., Mulligan, J., Stratford, R.J., Betts, P.R., & Voss, L.D. (1997). Are short normal children at a disadvantage? The Wessex growth study. *British Medical Journal, 314* (7074), 97-100.

Drotar, D. (1998). *Measuring health-related quality of life in children and adolescents: Implications for research and practice.* Mahwah, New Jersey: Lawrence Erlbaum Associates.

Dunn, G., Everitt, B., & Pickles, A. (1994). Multitrait-multimethod and multiple indicator multiple cause models. In G. Dunn, B. Everitt, & A. Pickles (Eds.), *Modelling, covariances and latent variables using EQS.* (pp. 77-87). London: Chapman & Hall.

Ebrahim, S. (1995). Clinical and public health perspectives and applications of health-related quality of life measurement. *Social Science & Medicine, 41* (10), 1383-1394.

Edgerton, R.B. (1990). Quality of life from a longitudinal research perspective. In L. S. Robert & J. B. Michael (Eds.), *Quality of life: Perspectives and issues.* (pp. 149-160). Washington,DC: American Association on Mental Retardation.

Eisen, M., Ware Jr., J.E., Donald, C.A., & Brook, R.H. (1979). Measuring components of children's health status. *Medical Care, 17* (9), 902-921.

Eiser, C. (1994). Making sense of chronic disease. The eleventh Jack Tizard memorial lecture. *Journal of Child Psychology and Psychiatry, 35* (8), 1373-1389.

Eiser, C., Havermans, T., Craft, A., & Kernahan, J. (1995). Development of a measure to assess the perceived illness experience after treatment for cancer. *Archives of Disease in Childhood, 72* (4), 302-307.

References

Elbers, E. (1995). Interactie tussen volwassenen en kinderen: Theorie, onderzoek, toepassing [Dutch: Interaction between adults and children: Theorie, reseach, application]. In L. Meeuwesen & H. Houtkoop-Steenstra (Eds.), *Studies naar mondelinge interacties [Dutch: Studies about verbal interactions].* (pp. 13-30). Utrecht: ISOR.

Ennett, S.T., DeVellis, B.M., Earp, J.A., Kredich, D., Warren, R.W., & Wilhelm, C.L. (1991). Disease experience and psychosocial adjustment in children with juvenile rheumatoid arthritis: Children's versus mothers' reports. *Journal of Pediatric Psychology, 16*(5), 557-568.

Epkins, C.C. (1995). Teachers' ratings of inpatient children's depression, anxiety, and aggression: A preliminary comparison between inpatient- facility and community-based teachers' ratings and their correspondence with children's self-reports. *Journal of Clinical Child Psychology, 24*(1), 63-70.

Epstein, A.M., Hall, J.A., Tognetti, J., Son, L.H., & Conant, L. (1989). Using proxies to evaluate quality of life. *Medical Care, 27*(3), S91-S98.

Erling, A., Wiklund, I., & Albertsson-Wikland, K. (1994). Prepubertal children with short stature have a different perception of their well-being and stature than their parents. *Quality of Life Research, 3*, 425-429.

Fairclough, D.L., & Gelber, R.D. (1996). Quality of life: Statistical issues and analysis. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials.* (pp. 427-435). Philadelphia: Lippincott-Raven Publishers.

Fallsberg, M. (1991). *Reflections on medicines and medication: a qualitative analysis among people on long-term drug regimens.* Linköping: Thesis Linköping University.

Fawer, C.L., Besnier, S., Forcada, M., Buclin, T., & Calame, A. (1995). Influence of perinatal, developmental and environmental factors on cognitive abilities of preterm children without major impairments at 5 years. *Early Human Development, 43*(2), 151-164.

Fayers, P.M., & Jones, D.R. (1983). Measuring and analysing quality of life in cancer clinical trials: a review. *Stat.Med., 2*(4), 429-446.

Fedrizzi, E., Inverno, M., Botteon, G., Anderloni, A., Filippini, G., & Farinotti, M. (1993). The cognitive development of children born preterm and affected by spastic diplegia. *Brain Development, 15*(6), 428-432.

Feeny, D., Furlong, W., Barr, R.D., Torrance, G.W., Rosenbaum, P., & Weitzman, S. (1992). A comprehensive multiattribute system for classifying the health status of survivors of childhood cancer. *J.Clin.Oncol., 10*(6), 923-928.

Feeny, D.H., Torrance, G.W., & Furlong, W.J. (1996). Health utilities index. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials.* (pp. 239-252). Philadelphia: Lippincott-Raven Publishers.

Ferell, B.R., Hassey Dow, K., & Grant, M. (1995). Measurement of the quality of life in cancer survivors. *Quality of Life Research, 4*, 523-531.

Finkelstein, J.W. (1998). Methods, models, and measures of health-related quality of life for children and adolescents. In D. Drotar (Ed.), *Measuring health-related quality of life in children and adolescents: implications for research and practice.* (pp. 39-52). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Fitzpatrick, R., Fletcher, A., Gore, S., Jones, D., Spiegelhalter, D., & Cox, D. (1992). Quality of life measures in health care. I: Applications and issues in assessment. *BMJ., 305*(6861), 1074-1077.

French, D.J., Christie, M.J., & Sowden, A.J. (1994). The reproducibility of the Childhood Asthma Questionnaires: Measures of quality of life for children with asthma aged 4-16 years. *Quality of Life Research, 3*(3), 215-224.

Gemke, R.J., Bonsel, G.J., & van-Vught, A.J. (1995). Long-term survival and state of health after paediatric intensive care. *Archives of Disease in Childhood, 73*(3), 196-201.

Gill, L.J., Shand, P.A.X., Fuggle, P., Dugan, B. , & Davies, S.C. (1997). Pain assessment for children with sickle cell disease: Improved validity of diary keeping versus interview ratings. *British Journal Of Health Psychology, 2 Part 2*, 131-140.

Gill, T.M., & Feinstein, A.R. (1994). A critical appraisal of the quality of quality-of-life measurements. *JAMA, 272*(8), 619-626.

Goldstein, K.M., Caputo, D.V., & Taub, H.B. (1976). The effects of prenatal and perinatal complications on development at one year of age. *Child Development, 47*(3), 613-621.

Gordeev, V.I., & Ul'rikh, E.V. (1997). Kachestvo zhizni detei, perenesshikh neotlozhnye khirurgicheskie operatsii v periode novorozhdennosti [Russian: The quality of life of children who have undergone emergency surgical operations in the neonatal period]. *Vestn.Khir.Im.I.I.Grek., 156*(2), 73-76.

Gross, S.J., Slagle, T.A., D'Eugenio, D.B., & Mettelman, B.B. (1992). Impact of a matched term control group on interpretation of developmental performance in preterm infants. *Pediatrics, 90*(5), 681-687.

Grunau, R.V.E., Whitfield, M.F., & Petrie, J.H. (1994). Pain sensitivity and temperament in extremely low-birth-weight premature toddlers and preterm and full-term controls. *Pain, 58*(3), 341-346.

Grunberg, S.M., Groshen, S., Steingass, S., Zaretsky, S., & Meyerowitz, B. (1996). Comparison of conditional quality of life terminology and visual analogue scale measurements. *Quality of Life Research, 5*, 65-72.

Guillemin, F., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J.Clin.Epidemiol., 46*(12), 1417-1432.

Guyatt, G.H., Jaeschke, R., Feeny, D.H., & Patrick, D.L. (1996). Measurements in clinical trials: Choosing the right approach. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials.* (pp. 41-48). Philadelphia: Lippincott-Raven Publishers.

Guyatt, G.H., Juniper, E.F., Griffith, L.E., Feeny, D.H., & Ferrie, P.J. (1997). Children and adult perceptions of childhood asthma. *Pediatrics, 99*(2), 165-168.

Hall, A., McLeod, A., Counsell, C., Thomson, L. , & Mutch, L. (1995). School attainment, cognitive ability and motor function in a total Scottish very-low-birthweight population at eight years: a controlled study. *Developmental Medicine and Child Neurology, 37* (12), 1037-1050.

Halsey, C.L., Collin, M.F., & Anderson, C.L. (1993). Extremely low birth weight children and their peers: a comparison of preschool performance. *Pediatrics, 91*(4), 807-811.

Haverkamp, F., & Noeker, M. (1998). 'Short stature in children - a questionnaire for parents': a new instrument for growth disorder-specific psychosocial adaptation in children. *Quality of Life Research, 7*, 447-455.

Hays, R.D., Vickrey, B.G., Hermann, B.P., Perrine, K., Cramer, J., Meador, K., Spritzer, K., & Devinsky, O. (1995). Agreement between self reports and proxy reports of quality of life in epilepsy patients. *Quality of Life Research, 4*, 159-168.

Heiser, W.J., & Meulman, J.J. (1994). Homogeneity analysis: Exploring the distribution of variables and their nonlinear relationships. In M. Greenacre & J. Blasius (Eds.), *Correspondence analysis in the social sciences: Recent developments and applications.* (pp. 179-209). London, San Diego, New York, Boston, Sydney, Tokyo, Toronto: Academic Press, Harcourt Brace & Company.

Hille, E.T.M., den-Ouden, L., Bauer, L., Van-den-Oudenrijn, C., Brand, R., & Verloove-Vanhorick, S.P. (1994). School performance at nine years of age in very premature and very low birth weight infants: Perinatal risk factors and predictors at five years of age. *Journal of Pediatrics, 125*, 426-434.

Hoppe-Hirsch, E., Renier, D., Lellouch, T.A., Sainte, R.C., Pierre, K.A., & Hirsch, J.F. (1990). Medulloblastoma in childhood: progressive intellectual deterioration. *Childs Nerv Syst , 6*(2), 60-65.

Horne, R. (1997). Representations of medication and treatment: Advances in theory and measurement. In K. J. Petrie & J. Weinman (Eds.), *Perceptions of health and illness: current research and applications.* (pp. 155-188). Singapore: Harwood Academic Publishers.

Hosli, E.J. (1998). *Ziektebegrip bij kinderen: de ontwikkeling van het denken over ziekte [Dutch: Understanding illness: The development of children's thinking about illness].* Utrecht: Thesis Utrecht University.

Huisman, J., Slijper, F.M.E., Sinnema, G., Akkerhuis, G.W., Brugman, A., Feenstra, J., Den Hartog, L., & Heuvel, F. (1991). Psychosocial functioning of short prepubertal children and girls with Turner's syndrome. *Acta Paediatrica Scandinavia, 377*(Suppl), 162-162.

Hyllienmark, L., & Ludvigsson, J. (1992). Insulinpump--ett realistiskt alternativ i behandlingen av diabetes hos barn och ungdom [Swedish: Insulin pump--a realistic alternative in the treatment of diabetes in children and adolescents]. *Läkartidningen, 89*(13), 1057-1062.

Iorio, R., Pensati, P., Botta, S., Moschella, S., Impagliazzo, N., Vajro, P., & Vegnente, A. (1997). Side effects of alpha-interferon therapy and impact on health- related quality of life in children with chronic viral hepatitis. *Pediatr Infect.Dis J, 16*(10), 984-990.

Jordan, T.E. (1993). Estimating the quality of life for children around the world: NICQL '92. *Social Indicators Research, 30*, 17-38.

Juniper, E.F., Guyatt, G.H., Feeny, D.H., Ferrie, P.J., Griffith, L.E., & Townsend, M. (1996). Measuring quality of life in children with asthma. *Quality of Life Research, 5*(1), 35-46.

Juniper, E.F., Guyatt, G.H., Feeny, D.H., Griffith, L.E., & Ferrie, P.J. (1997). Minimum skills required by children to complete health-related quality of life instruments for asthma: comparison of measurement properties. *European Respiratory Journal, 10*(10), 2285-2294.

Juniper, E.F., Howland, W.C., Roberts, N.B., Thompson, A.K., & King, D.R. (1998). Measuring quality of life in children with rhinoconjunctivitis. *Journal of Allergy and Clinical Immunology, 101*(2 Pt 1), 163-170.

Kalmar, M. (1996). The course of intellectual development in preterm and fullterm children: An 8-year longitudinal study. *International Journal of Behavioral Development, 19*(3), 491-516.

Kamp, G., Waelkens, J.J., Delemarre-Vandewaal, H.E., DeMuinck Keizer-Schrama, S.M., & Wit, J.M. (1999). Biochemical characterisation of growth hormone sensitivity in idiopathic short stature: Increments of alkaline phosphatase and PIIINP predict the response to GH treatment. *Hormone Research, 51*(Suppl 2), 54-54.

Kamphuis, R.P. (1987). The concept of quality of life in pediatric oncology. In N. K. Aaronson & J. Beckman (Eds.), *The quality of life of cancer patients*. New York: Ravenpress.

Kaplan, R.M., & Anderson, J.P. (1988). A general health policy model: update and applications. *Health Services Research, 23*(2), 203-235.

Kazak, A.E., Penati, B., Boyer, B.A., Himelstein, B., Brophy, P., Waibel, M.K., Blackall, G.F., Daller, R., & Johnson, K. (1996). A randomized controlled prospective outcome study of a psychological and pharmacological intervention protocol for procedural distress in pediatric leukemia. *J Pediatr Psychol, 21*(5), 615-631.

Kekomäki, M., Vilkki, P., Gordin, A., & Salo, H. (1981). Loperamide as a symptomatic treatment in pediatric surgery: A double-blind cross-over study. *Z. Kinderchir. Grenzgeb., 32*(3), 237-243.

Kerkhof, A.J.F.M., van-der-Wal, J., & Hengeveld, M.W. (1988). Typology of persons who attempted suicide with predictive value for repetition: a prospective cohort study. In H. J. Möller, A. Schmidtke, & R. Welz (Eds.), *Current issues of suicidology*. (pp. 194-203). Berlin: Springer-Verlag.

Kloosterman, G.J. (1969). Over intra-uteriene groei en de intrauteriene groeicurve [Dutch: Intrauterine growth and intrauterine growth curves]. *Maandschr. Kindergeneeskd., 37*(7), 209-225.

Konishi, T., Masuko, K., Naganuma, Y., Hongou, K., & Yagi, S. (1995). Flunitrazepam for sleep disturbance in children with intractable epilepsy. *Brain Development, 17*(1), 69-72.

Koopman, H.M. (1993). *De kijk van kinderen: Een cognitief structurele visie op de ontwikkeling van het denken over ziekte [Dutch: the children's view: a cognitive structural approach to the development of thinking about illness]*. Amsterdam: Thesis Publishers.

Koopman, H.M., Theunissen, N.C.M., Vogels, A.G.C., Zwinderman, A.H., Verloove-Vanhorick, S.P., & Wit, J.M. (1999). Health related quality of life of children with a chronic illness: Parent versus child report. *Gedrag & Gezondheid, 27*(1/2), 118-125.

Landgraf, J.M., Abetz, L., & Ware, J.E. (1996). *The CHQ user's manual*. (1st edition ed.). Boston, MA: The Health Institute, New England Medical Centre.

Landgraf, J.M., & Abetz, L. (1996). Measuring health outcomes in pediatric populations: Issues in psychometrics and application. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trails*. (pp. 793-802). Philadelphia: Lippincott-Raven Publishers.

Landgraf, J.M., & Abetz, L.N. (1997). Functional status and well-being of children representing three cultural groups: Initial self-reports using the CHQ-CF87. *Psychology and Health, 12*(6), 839-854.

Lane, S.J., Attanasio, C.S., & Huselid, R.F. (1994). Prediction of preschool sensory and motor performance by 18- month neurologic scores among children born prematurely. *American Journal of Occupational Therapy, 48*(5), 391-396.

Langfitt, J.T. (1995). Comparison of the psychometric characteristics of three quality of life measures in intractable epilepsy. *Qual. Life Res., 4*(2), 101-114.

Largo, R.H., Graf, S., Kundu, S., Hunziker, U., & Molinari, L. (1990). Predicting developmental outcome at school age from infant tests of normal, at-risk and retarded infants. *Developmental Medicine and Child Neurology, 32*(1), 30-45.

Largo, R.H., Molinari, L., Kundu, S., Lipp, A., & Duc, G. (1990). Intellectual outcome, speech and school performance in high risk preterm children with birth weight appropriate for gestational age. *European Journal of Pediatrics, 149*(12), 845-850.

Leiberman, E., Pilpel, D., Carel, C.A., Levi, E. , & Zadik, Z. (1993). Coping and Satisfaction with Growth Hormone Treatment among Short-Stature Children. *Hormone Research, 40*, 128-135.

Leventhal, H., & Colman, S. (1997). Quality of life: A process view. *Psychology and Health, 12*(6), 753-767.

Levi, R., & Drotar, D. (1998). Critical issues and needs in health-related quality of life assessment of children and adolescents with chronic health conditions. In D. Drotar (Ed.), *Measuring health-related quality of life in children and adolescents: implications for research and practice.* (pp. 3-20). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Levy-Shiff, R., Einat, G., Har Even, D., Mogilner, M., Mogilner, S., Lerman, M., & Krikler, R. (1994). Emotional and behavioral adjustment in children born prematurely. *Journal of Clinical Child Psychology, 23*(3), 323-333.

Lewis, C.C., Pantell, R.H., & Kieckhefer, G.M. (1989). Assessment of children's health status. Field test of new approaches. *Medical Care, 27*(3 Suppl), S54-S65.

Lindström, B. (1995). Measuring and improving quality of life for children. In B. Lindström & N. Spencer (Eds.), *Social Paediatrics.* (pp. 571-585). Oxford: Oxford University Press.

Loras, D.I., De-Potters, S., Pharaon, I., Olives, J.P., & Hermier, M. (1993). Qualité de vie des enfants en nutrition parenterale a domicile et de leurs parents [French: Quality of life of children with home parenteral nutrition and of their parents]. *Pediatrie., 48*(7-8), 555-560.

Magaziner, J., Simonsick, E.M., Kashner, T.M., & Hebel, J.R. (1988). Patient-proxy response comparability on measures of patient health and functional status. *Journal of Clinical Epidemiology, 41*(11), 1065-1074.

Maillé, A.R., Kaptein, A.A., deHaes, J.C.J.M., & Everaerd, W.T.A.M. (1996). Assessing quality of life in chronic non-specific lung disease - A review of empirical studies published between 1980 and 1994. *Quality of Life Research, 5*(2), 287-301.

Marrero, D.G., Vandagriff, J.L., Kronz, K., Fineberg, N.S., Golden, M.P., Gray, D., Orr, D.P., Wright, J.C., & Johnson, N.B. (1995). Using telecommunication technology to manage children with diabetes: the Computer-Linked Outpatient Clinic (CLOC) Study. *Diabetes Educ., 21*(4), 313-319.

Marshall, G.N., Hays, R.D., & Nicholas, R. (1994). Evaluating agreement between clinical assesment methods. *International Journal of Methods in Psychiatric Research, 4*, 249-257.

McCormick, M.C., Workman Daniels, K., & Brooks Gunn, J. (1996). The behavioral and emotional well-being of school-age children with different birth weights. *Pediatrics, 97*(1), 18-25.

McDonald, R.P. (1981). The dimensionality of tests and items. *Britisch Journal of Mathematical and Statistical Psychology, 34*, 100-117.

Meisels, S.J., Plunkett, J.W., Roloff, D.W., Pasick, P.L., & Stiefel, G.S. (1986). Growth and development of preterm infants with respiratory distress syndrome and bronchopulmonary dysplasia. *Pediatrics, 77*(3), 345-352.

Meltzer, E.O., Orgel, H.A., Biondi, R., Georgitis, J., Milgrom, H., Munk, Z., Van-Bavel, J., Wood, C.C., & Drda, K. (1997). Ipratropium nasal spray in children with perennial rhinitis. *Ann Allergy Asthma Immunol, 78*(5), 485-491.

Menyuk, P., Chesnick, M., Liebergott, J.W., & Korngold, B. (1991). Predicting reading problems in at-risk children. *Journal of Speech and Hearing Research, 34*(4), 893-903.

Micheli, R., Giordano, L., & Balestrini, M.R. (1996). Tumori cerebrali in neurofibromatosi tipo 1 in eta pediatrica [Italian: Cerebral tumors in children with neurofibromatosis type 1]. *Minerva Pediatr, 48*(3), 89-97.

Middle, C., Johnson, A., Alderdice, F., & Petty, T. (1996). Birthweight and health and development at the age of 7 years. *Child Care, Health and Development, 22*(1), 55-71.

Miller, G., Dubowitz, L.M., & Palmer, P. (1984). Follow-up of pre-term infants: Is correction of the developmental quotient for prematurity helpful? *Early Human Development, 9*(2), 137-144.

Morris, K.P., Sharp, J., Watson, S., & Coulthard, M.G. (1993). Non-cardiac benefits of human recombinant erythropoietin in end stage renal failure and anaemia. *Archives of Disease in Childhood, 69*(5), 580-586.

Naughton, M.J., Shumaker, S.A., Anderson, R.T., & Czajkowski, S.M. (1996). Psychological aspects of health-related quality of life measurement: Tests and scales. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials.* (pp. 117-131). Philadelphia: Lippincott-Raven Publishers.

References

Neely, E.K., & Rosenfeld, R.G. (1994). Use and abuse of human growth hormone. *Annu.Rev.Med, 45*, 407-420.

Nunnally, J.C. (1967). *Psychometric theory*. New York: McGraw-Hill.

O'Boyle, C. (1995). Making subjectivity scientific. *The Lancet, 345*(8950), 602-602.

Orenstein, D.M., & Kaplan, R.M. (1991). Measuring the quality of well-being in cystic fibrosis and lung transplantation. The importance of the area under the curve. *Chest, 100*(4), 1016-1018.

Pal, D.K. (1996). Quality of life assessment in children: A review of conceptual and methodological issues in multidimensional health status measures. *Journal of Epidemiology and Community Health, 50*, 391-396.

Papile, L.A., Burstein, J., Burstein, R., & Koffler, H. (1978). Incidence and evolution of subependymal and intraventricular hemorrhage: a study of infants with birth weights less than 1,500 gm. *Journal of Pediatrics, 92*(4), 529-534.

Parkin, P.C., Kirpalani, H.M., Rosenbaum, P.L., Fehlings, D.L., Van-Nie, A., Willan, A.R., & King, D. (1997). Development of a health-related quality of life instrument for use in children with spina bifida. *Quality of Life Research, 6*(2), 123-132.

Paty, E., De-Blic, J., Brunet, D., Le-Bourgeois, M., Garcelon, M., Paupe, J., & Scheinmann, P. (1990). Desensibilisation accelerée aux dermatophagoides pteronyssinus chez l'asthmatique severe. Bilan au bout d'un an d'immunotherapie [French: Accelerated desensitization with dermatophagoides pteronyssinus in severe asthmatic children. Evaluation after one year of immunotherapy]. *Arch Fr Pediatr , 47*(3), 173-179.

Pearl, R., & Donahue, M. (1995). Brief report: Four years after a preterm birth: Children's development and their mothers' beliefs and expectations. *Journal of Pediatric Psychology, 20*(3), 363-370.

Pichlmayr, R. (1991). Transplantations hépatiques. L'experience de Hanovre [French: Liver transplantation. The Hannover experience]. *Chirurgie., 117*(3), 225-229.

Pilpel, D., Leiberman, E., Zadik, Z., & Carel, C.A. (1995). Effect of growth hormone treatment on quality of life of short- stature children. *Hormone Research, 44*(1), 1-5.

Plomin, R. (1989). Developmental behavioral genetics: stability and instability. In M. H. Bornstein & N. A. Krasnegor (Eds.), *Stability and continuity in mental development.* (pp. 273-291). Hillsdale: Lawrence Erlbaum Associates.

Portnoy, S., Callias, M., Wolke, D., & Gamsu, H. (1988). Five-year follow-up study of extremely low-birthweight infants. *Developmental Medicine and Child Neurology, 30*(5), 590-598.

Post, M. (1997). *Living with spinal cord injury*. Utrecht: Elinkwijk.

Raboud, J.M., Singer, J., Thorne, A., Schechter, M.T., & Shafran, S.D. (1998). Estimating the effect of treatment on quality of life in the presence of missing data due to drop-out and death. *Quality of Life Research, 7*(6), 487-494.

Ramey, C.T., Lee, M.W., & Burchinal, M.R. (1989). Developmental plasticity and predictability: Consequences of ecological change . In M. H. Bornstein & N. A. Krasnegor (Eds.), *Stability and continuity in mental development.* (pp. 217-233). Hillsdale: Lawrence Erlbaum Associates.

Ranke, M.B. (1996). Towards a consensus on the definition of idiopathic short stature. Summary. *Hormone Research, 45*(Suppl 2), 64-66.

Reijden-Lakeman van der, I.E.A. (1996). *Growing Pains? Psychological evaluation of children with short stature after intrauterine growth retardation, before and after two years of growth hormone treatment.* Rotterdam: Thesis Erasmus University Rotterdam.

Reijden-Lakeman van der, I.E.A., Slijper, F.M.E., Dongen-Melman van, J.E.W.M., Waal de, W.J., & Verhulst, F.C. (1996). Self-concept before and after two years of growth hormone treatment in intrauterine growth-retarded children. *Hormone Research, 46*, 88-94.

Rekers-Mombarg, L.T.M., Busschbach, J.J.V., Massa, G.G., Dicke, J., & Wit, J.M. (1998). Quality of life of young adults with idiopathic short stature: effect of growth hormone treatment. *Acta Paediatrica, 87*(8), 865-870.

Rekers-Mombarg, L. (1996). Spontaneous growth in idiopathic short stature. *Archives of Disease in Childhood, 75*, 175-180.

Rendo, P., Biscigniano, L., Turconi, A., Sanchez, A.J., & Greco, J. (1994). Tratamiento de la anemia en ninos sometidos a hemodialisis cronica con eritropoyetina humana recombinante [Spanish: Treatment of aenaemia in children undergoing chronic hemodialysis with recombinant human erythropoietin]. *Medicina.B Aires., 54*(5 Pt 1), 392-398.

Robertson, C.M., Etches, P.C., & Kyle, J.M. (1990). Eight-year school performance and growth of preterm, small for gestational age infants: a comparative study with subjects matched for birth weight or for gestational age. *Journal of Pediatrics, 116* (1), 19-26.

Roede, M.J., & van-Wieringen, J.C. (1985). Growth diagrams 1980. Netherlands third nationwide survey. *Tijdschrift voor Sociale Gezondheidszorg, 63* (Suppl.), 1-34.

Rogerson, R.J. (1995). Environmental and health-related quality of life: conceptual and methodological similarities. *Social Science & Medicine, 41* (10), 1373-1382.

Rosenbaum, P., Cadman, D., & Kirpalani, H. (1990). Pediatrics: assessing quality of life. In B. Spilker (Ed.), *Quality of life assessments in clinical trials.* (pp. 205-215). New York: Raven Press, Ltd.

Rosenberg, R. (1995). Health-related quality of life between naturalism and hermeneutics. *Social Science & Medicine, 41* (10), 1411-1415.

Rosenfeld, R.M., Goldsmith, A.J., Tetlus, L., & Balzano, A. (1997). Quality of life for children with otitis media. *Arch Otolaryngol Head Neck Surg, 123* (10), 1049-1054.

Ross, G., Lipper, E.G., & Auld, P.A. (1990). Social competence and behavior problems in premature children at school age. *Pediatrics, 86* (3), 391-397.

Saigal, S., Szatmari, P., & Rosenbaum, P.L. (1992). Can learning disabilities in children who were extremely low birth weight be identified at school entry? *Journal of Developmental and Behavioral Pediatrics, 13* (5), 356-362.

Saigal, S., Rosenbaum, P., Stoskopf, B., Hoult, L., Furlong, W., Feeny, D., Burrows, E., & Torrance, G. (1994). Comprehensive assessment of the health status of extremely low birth weight children at eight years of age: Comparison with a reference group. *Journal of Pediatrics, 125* (3), 411-417.

Saigal, S., Feeny, D., Furlong, W., Rosenbaum, P., Burrows, E., & Torrance, G. (1994). Comparison of the health-related quality of life of extremely low birth weight children and a reference group of children at age eight years. *Journal of Pediatrics, 125* (3), 418-425.

Saigal, S., Feeny, D., Rosenbaum, P., Furlong, W., Burrows, E., & Stoskopf, B. (1996). Self-perceived health status and health-related quality of life of extremely low-birth-weight infants at adolescence. *JAMA, 276* (6), 453-459.

Saigal, S., Feeny, D., Rosenbaum, P., Furlong, W., Stoskopf, B., & Hoult, L. (1996). Extremely low-birth-weigth infants at adolescence: Health status and quality of life: reply to a letter to the editor. *JAMA, 276*, 1722-1723.

Sandberg, D.E., Brook, A.E., & Campos, S.P. (1994). Short stature: a psychosocial burden requiring growth hormone therapy? *Pediatrics, 94* (6 Pt 1), 832-840.

Sandberg, D.E., & Michael, P. (1998). Psychosocial stressses related to short stature: Does their presence imply psychological dysfunction? In D. Drotar (Ed.), *Measuring health-related quality of life in children and adolescents: implications for research and practice.* (pp. 287-312). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Saylor, C.F., Casto, G., & Huntington, L. (1996). Predictors of developmental outcomes for medically fragile early intervention participants. *Journal of Pediatric Psychology, 21* (6), 869-887.

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype environment effects. *Child Development, 54* (2), 424-435.

Schermann Eizirik, L., Hagekull, B., Bohlin, G., Persson, K., & Sedin, G. (1997). Interaction between mothers and infants born at risk during the first six months of corrected age. *Acta Paediatrica, 86* (8), 864-872.

Schipper, H., Clinch, J.J., & Olweny, C.L.M. (1996). Quality of life studies: Definitions and conceptual issues. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials.* (pp. 11-23). Philadelphia: Lippincott-Raven Publishers.

Schmidt, L.R., & Dlugosch, G.E. (1992). Entwicklungspsychologische Aspekte der Gesundheitspsychologie [German: Developmental aspects of health psychology]. *Zeitschrif für Klinsche Psychologie, 21* (1), 36-47.

Schmitt, N., & Stults, D.M. (1986). Methodology review: Analysis of multitrait-multimethod matrices. *Applied Psychological Measurement, 10* (1), 1-22.

Schoenfelder, J.R., & Spilker, B. (1996). Presentation of quality of life data. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials.* (pp. 477-487). Philadelphia: Lippincott-Raven Publishers.

Schols, M. (1993). *Betekenissen van een premature geboorte: Effecten van interventies in de relatie tussen moeder en prematuur kind in de eerste twee levensjaren [Dutch: Meanings of a premature birth: Effects of interventions in the relationship between mother and premature child in the first two years of life].* Enschede: Febo druk bv.

Scholtz, W., Haubrock, M., Lob, C.T., Gebert, N. , Wahn, U., & Szczepanski, R. (1996). Kosten-Nutzen-Untersuchungen bei ambulanten Schulungsmassnahmen für asthmakranke Kinder und ihre Familien [German: Cost-effectiveness studies of ambulatory educational programs for children with asthma and their families]. *Pneumologie, 50*(8), 538-543.

Schothorst, P.F., & van-Engeland, H. (1996). Long-term behavioral sequelae of prematurity. *J.Am.Acad.Child Adolesc.Psychiatry, 35*(2), 175-183.

Siegel, L.S. (1982). Reproductive, perinatal, and environmental variables as predictors of development of preterm (less than 1501 grams) and fullterm children at 5 years. *Seminars in Perinatology, 6*(4), 274-279.

Siegel, P.T., Clopper, R., & Stabler, B. (1991). Psychological impact of significantly short stature. *Acta Paediatrica Scandinavia, 377*(Suppl), 14-18.

Silva, P.A., McGee, R., & Williams, S. (1984). A longitudinal study of the intelligence and behavior of preterm and small for gestational age children. *Journal of Developmental and Behavioral Pediatrics, 5*(1), 1-5.

Spencer, N.J., & Coe, C. (1996). The development and validation of a measure of parent-reported child health and morbidity: the Warwick Child Health and Morbidity Profile. *Child Care, Health and Development, 22*(6), 367-379.

Spieth, L.E., & Harris, C.V. (1996). Assessment of health-related quality of life in children and adolescents: An integrative review. *Journal of Pediatric Psychology, 21*(2), 175-193.

Spiker, D., Kraemer, H.C., Constantine, N.A., & Bryant, D. (1992). Reliability and validity of behavior problem checklists as measures of stable traits in low birth weight, premature preschoolers. *Child Development, 63*(6), 1481-1496.

Spilker, B. (1996). Introduction. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials.* (pp. 1-10). Philadelphia: Lippincott-Raven Publishers.

Spilker, B. (1996). *Quality of life and pharmacoeconomics in clinical trials.* (2th ed.). Philadelphia: Lippincott-Raven-Publishers.

Sprangers, M.A. (1996). Response-shift bias: a challenge to the assessment of patients' quality of life in cancer clinical trials. *Cancer Treatment Reviews, 22 Suppl A*, 55-62.

Sprangers, M.A.G., & Aaronson, N.K. (1992). The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: A review. *Journal of Clinical Epidemiology, 45*(7), 743-760.

Sprangers, M.A.G., Cull, A., Bjordal, K., Groenvold, M., & Aaronson, N.K. (1993). The European Organization for Research and Treatment of Cancer approach to quality of life assessment: guidelines for developing questionnaire modules. *Quality of Life Research, 2*, 287-295.

Sprangers, M.A.G., Rozemuller, N., van-den-Berk, M.B.P., Boven, S.V., & van-Dam, F.S.A.M. (1994). Response shift bias in longitudinal quality of life research. *Quality of Life Research, 3*, 49-49.

SPSS Categories 8.0 (1998) Chicago: SPSS Inc.

Stabler, B., Clopper, R.R., Siegel, P.T., Nicholas, L.M., Silva, S.G., Tancer, M.E., & Underwood, L.E. (1996). Links between growth hormone deficiency, adaptation and social phobia. *Hormone Research, 45*(1-2), 30-33.

Stacy, A.W., Widaman, K.F., Hays, R., & DiMatteo, M.R. (1985). Validity of self-reports of alcohol and other drug use: A multitrait-multimethod assessment. *Journal of Personality and Social Psychology, 49*(1), 219-232.

Stanton, W.R., McGee, R., & Silva, P.A. (1991). Indices of perinatal complications, family background, child rearing, and health as predictors of early cognitive and motor development. *Pediatrics, 88*(5), 954-959.

Steiger, J.H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87*(2), 245-251.

Stein, R.E., & Jessop, D.J. (1990). Functional status II (R). A measure of child health status. *Medical Care, 28*(11), 1041-1055.

Strauss, A.L. (1987). *Qualitative analysis for social scientists.* Cambridge: Cambridge University Press.

Sweeny, A.J., & Creer, T.L. (1995). Health-related-quality-of-life assessment in medical care. *Disease-a-Month, 16*(1), 5-64.

Taylor, K.M., Macdonals, K.G., Bezjak, A., DePetrillo, P.N.G., & DePetrillos, A.D. (1996). Physicians' perspective on quality of life: an exploratory study of oncologists. *Quality of Life Research, 5*, 5-14.

Tennant, A. (1995). Quality of life - a measure too far? *Annals of the Rheumatic Diseases, 54*(6), 439-440.

Theunissen, N.C.M. (1994). Living with end-stage renal failure: A developmental task for the adolescent. In J. J. F. Ter-Laak (Ed.), *Developmental tasks: Research program developmental psychology Utrecht-Moscow.* (pp. 95-114). Utrecht: Utrecht University.

Theunissen, N.C.M., Vogels, A.G.C., Koopman, H.M., Verrips, G.H., Zwinderman, A.H., Verloove-Vanhorick, S.P., & Wit, J.M. (1998). The proxy problem: child report versus parent report in health-related quality of life research. *Quality of Life Research, 7*(5), 387-397.

Thompson, R.J., Gustafson, K.E., Oehler, J.M., & Catlett, A.T. (1997). Developmental outcome of very low birth weight infants at four years of age as a function of biological risk and psychosocial risk. *Journal of Developmental and Behavioral Pediatrics, 18*(2), 91-96.

Tin, W., Wariyar, U., & Hey, E. (1997). Changing prognosis for babies of less than 28 weeks' gestation in the north of England between 1983 and 1994. Northern Neonatal Network. *British Medical Journal, 314* (7074), 107-111.

Torrance, G.W. (1987). Utility approach to measuring health-related quality of life. *Journal of Chronic Diseases, 40*(6), 593-603.

Van-Damme-Lombaerts, R., Broyer, M., Businger, J., Baldauf, C., & Stocker, H. (1994). A study of recombinant human erythropoietin in the treatment of anaemia of chronic renal failure in children on haemodialysis. *Pediatric Nephrology, 8*(3), 338-342.

van-der-Ham, T., Meulman, J.J., van-Strien, D.C., & van-Engeland, H. (1997). Empirically based subgrouping of eating disorders in adolescents: A longitudinal perspective. *British Journal of Psychiatry, 170*, 363-368.

van-der-Heyden, Z. (1992). *Vanuit de couveuse de wereld in: vragen van ouders van een couveusekind beantwoord [Dutch: Out of the incubator, into the world: questions of parents with an incubator-child answered].* Ede: Zomer & Keuning Boeken B.V.

van-Westerlaak, J.M., Kropman, J.A., & Collaris, J.W.M. (1975). *Beroepenklapper [Dutch: Professional index].* Nijmegen: Instituut voor Toegepaste Sociologie [Institution for Applied Sociology].

Veen, S., Ens-Dokkum, M.H., Schreuder, A.M., Verloove-Vanhorick, S.P., Brand, R., & Ruys, J.H. (1991). Impairments, disabilities, and handicaps of very preterm and very-low-birthweight infants at five years of age. *The Lancet, 338*(6), 33-36.

Verhulst, F.C., & Akkerhuis, G.W. (1989). Agreement between parents' and teachers' ratings of behavioral/emotional problems of children aged 4-12. *Journal of Child Psychology and Psychiatry, 30*(1), 123-136.

Verhulst, F.C., Prince, J., Vervuurt, P.C., & de-Jong, J. (1989). Mental health in Dutch adolescents: self-reported competencies and problems for ages 11-18. *Acta Psychiatr. Scand. Suppl., 356*, 1-48.

Verloove-Vanhorick, S.P., Verwey, R.A., Ebeling, M.C.A., Brand, R., & Ruys, J.H. (1988). Mortality in very preterm and very low birth weight infants according to place of birth and level of care: Results of a national collaborative survey of preterm and very low birth weight Infants in the Netherlands. *Pediatrics, 81*(3), 404-411.

Verrips, G.H., Vogels, A.G.C., Verloove-Vanhorick, S.P., Fekkes, M., Koopman, H.M., Kamphuis, R.P., Theunissen, N.C.M., & Wit, J.M. (1997). Health-related quality of life measure for children -- the TACQOL. *Journal of Applied Therapeutics, 1*, 357-360.

Verrips, G.H., Vogels, A.G.C., Koopman, H.M., Theunissen, N.C.M., Kamphuis, R.P., Wit, J.M., & Verloove-Vanhorick, S.P. (1999). Measuring health-related quality of life in a child population. *European Journal of Public Health, 9*(3), 188-193.

Vivier, P.M., Bernier, J.A., & Starfield, B. (1994). Current approaches to measuring health outcomes in pediatric research. *Curr. Opin. Pediatr., 6*(5), 530-537.

Vogels, T., Theunissen, N.C.M., Verrips, G.H., Koopman, H.M., Verloove-Vanhorick, S.P., & Kamphuis, R.P. (1996). Het meten van kwaliteit van leven bij kinderen met chronische aandoeningen [Dutch: Assessing health related quality of life in chronically ill children]. *TIAZ, 3*, 104-111.

References

Vogels, T., Verrips, G.H., Verloove-Vanhorick, S.P., Fekkes, M., Kamphuis, R.P., Koopman, H.M., Theunissen, N.C.M., & Wit, J.M. (1998). Measuring health-related quality of life in children: the development of the TACQOL-Parent Form. *Quality of Life Research, 7*(5), 457-465.

Voss, L.D., Bailey, B.J.R., Mulligan, J., Wilkin, T.J., & Betts, P.R. (1991). Short Stature and School Performance: The Wessex Growth Study. *Acta Paediatrica Scandinavia, Suppl 370*, 29-31.

Voss, L.D., & Wiklund, I. (1995). Short stature and psychosocial assessment. *Acta Paediatrica Supplement, 411*, 69-74.

Wadsworth, M.E., Rodgers, B., & Wadsworth, M.E.J. (1989). Longterm follow-up studies; a critical overview. *Rev Epidemiol Sante. Publique., 37*(5-6), 533-547.

Weissman, M.M., Orvaschel, H., & Padian, N. (1980). Children's symptom and social functioning self-report scales. Comparison of mothers' and children's reports. *J. Nerv. Ment. Dis. , 168*(12), 736-740.

WHOQOL Group. (1995). The World Health Organization Quality of Life Assessment (WHOQOL): Position paper from the World Health Organization. *Social Science & Medicine, 41*(10), 1403-1409.

Wiklund, I., Albertsson-Wikland, K., & Wiren, L. (1991). How should the Psychosocial impact of Short Stature and Response to Growth Hormone Treatment be Assessed in Children? *Acta Paediatrica Scandinavia, Suppl. 377*, 174-175.

Wiklund, I., Wiren, L., Erling, A., Karlberg, J., & Albertsson-Wikland, K. (1994). A new self-assessment questionnaire to measure well-being in children, particularly those of short stature. *Quality of Life Research, 3*(6), 449-455.

Wiklund, I., Erling, A., & Albertsson-Wikland, K. (1998). Critical review of measurement issues in quality of life assessment for children with growth problems. In D. Drotar (Ed.), *Measuring health-related quality of life in children and adolescents: implications for research and practice.* (pp. 255-271). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Willner, A.E.E., & Rodewald, G.E. (1990). *Impact of cardiac surgery on the quality of life: Neurological and psychological aspects.* (xiii ed.). New York, NY, US: Plenum Press.

Wit, J.M., Kamp, G.A., & Rikken, B. (1996). Spontaneous growth and response to growth hormone treatment in children with growth hormone deficiency and idiopathic short stature. *Pediatric Research, 39*(2), 295-302.

Wolke, D., Ratschinski, G., Ohrt, B., & Riegel, K. (1994). The cognitive outcome of very preterm infants may be poorer than often reported: an empirical investigation of how methodological issues make a big difference. *European Journal of Pediatrics, 153*(12), 906-915.

World Health Organization. (1948). *Constitution of the World Health Organization.* Geneva: World Health Organization.

World Health Organization. (1980). *International classification of impairments, disabilities, and handicaps.* Geneva: World Health Organization.

Worthington, H.V. (1984). The suitability of the statistical techniques currently used to describe and analyse cross-sectional caries data. *Community. Dent. Health, 1*(2), 125-130.

Wray, J., Radley, S.R., & Yacoub, M. (1992). Effect of cardiac or heart-lung transplantation on the quality of life of the paediatric patient. *Quality of Life Research, 1*(1), 41-46.

Zivcic, I. (1993). Emotional reactions of children to war stress in Croatia. *J. Am. Acad. Child Adolesc. Psychiatry, 32*(4), 709-713.

Zwinderman, A.H. (1992). Statistical analysis of longitudinal quality of life data with missing measurements. *Quality of Life Research, 1*, 219-224.

Summary

In *Chapter 1* the aim of this thesis is introduced, which is the investigation of Health related Quality of Life in children aged 0-12 years. As medical successes have increased, a systematic outcome measure was needed that could account for the different ways children can react to a certain physical problem. Although the individual paediatrician already recognised these individual differences, a systematic and quantifiable outcome measure was missing. This gap was filled by the introduction of the construct *Health related Quality of Life (HRQoL)* also referred to as Quality of Life (QoL). As a generally accepted definition or theoretical framework is missing, we define HRQoL as: the individual's perception of problems in health status, combined with the affective responses to such problems. *Health Status (HS)* is the assessment by a person of his or her own health. The 'health' component involves physical, psychological as well as social functioning. Furthermore, we subscribe to the notion that HRQoL is multi-factorial (physical, psychological and social well-being), patient self-administered, subjective and variable over time. In this thesis, the usefulness of this definition was explored.

Chapters 2 to 6 describe how to define and obtain HRQoL in children. In *Chapter 2* the development of the 56-item TNO-AZL-Child-Quality-Of-Life (TACQOL) questionnaire is presented. The instrument is developed to meet the need for a reliable and valid instrument for measuring HRQoL in children. HRQoL was defined as HS in seven scales plus the emotional responses to problems in HS. The TACQOL explicitly offers respondents the possibility of differentiating between their functioning and the way they feel about it. A random sample of 1789 parents of 6-11 year olds completed the TACQOL, as well as 1159 8-11 year olds themselves. Multiple correspondence analyses showed that item response categories were ordinal, and that the TACQOL scales may be regarded as metric. Cronbach's alpha ranged from 0.65-0.84. Only 57% of reported HS problems were associated with negative emotions, which supports our definition of HRQoL. Intraclass correlation coefficients between Parent Forms and Child Forms ranged from 0.44-0.61. Pearson's correlation coefficients between TACQOL and the Dutch version of a German HRQoL instrument (KINDL) ranged from 0.24-0.60. Univariate analyses of variance showed that children with chronic diseases and children receiving medical treatment had lower TACQOL scores than healthy children. The study showed that with the TACQOL, children's

HRQoL can be measured in a reliable and valid way. In addition, the study provided a large norm group for children age 6 to 11 years.

In Chapters 3 and 4 the relation between child and parent report on children's HRQoL is studied. ***Chapter 3*** evaluates the agreement between child and parent reports on HRQoL in a representative sample of 1105 Dutch children (age 8 to 11 years old). Both children and their parents completed the TACQOL. The Pearson correlations between child and parent reports were between 0.44 to 0.61 ($p<0.001$). Intraclass correlations were between 0.39 to 0.62. Children on average reported significantly lower HRQoL than parents on the physical complaints, motor functioning, autonomy, cognitive functioning, and positive emotions scales (paired t-test: $p<0.05$). Agreement on all scales was related to the height of the HRQoL scores and on some scales to some background variables (gender, age, temporary illness, visiting a physician). According to multitrait-multimethod analyses, both child and parent reports on HRQoL proved to be valid. We recommended using both children's and parent's evaluations whenever possible.

In ***Chapter 4,*** the agreement is evaluated between child and parent reports on children's HRQoL in a sample of 416 Dutch children with a chronic disease (8 to 15 years). Both children and their parents completed the TACQOL. The correlations between child and parent reports varied from -0.10 to 0.99 amongst the various chronic conditions. Children reported lower HRQoL on the physical complaints, motor functioning and positive emotions scales. Parents reported lower HRQoL on the social, and negative emotions scales. Agreement on all scales was related to the type of chronic illness. The child and the parent each provide different information on HRQoL. Knowledge of both judgements will enhance the care of children with a chronic illness and their parents.

Chapter 5 describes the development of the 46-item TNO-AZL Pre-school Quality Of Life (TAPQOL) questionnaire. HRQoL was defined as HS in 13 scales plus the emotional responses to problems in HS. The TAPQOL has to be completed by the parents. A sample of 121 parents of preterm children completed the TAPQOL questionnaire as well as 362 parents of children from the open population. On the base of Cronbach's alpha, item-rest correlation, and factor-analysis, the TAPQOL scales were constructed on the data of the preterm children sample. The psychometric performance of these scales was evaluated for both the preterm children sample and the open population sample. Cronbach's alpha ranged from 0.66-0.88 for the preterm children sample and from 0.26-0.85 for the open population sample. The uni-dimensionality of the separate scales was confirmed by principal component analysis for

both samples. Pearson's correlations between scales were low. T-test analyses showed that very preterm children and children with chronic diseases had lower scores (indicating a poorer HRQoL) on the TAPQOL scales than healthy children. This study shows that the TAPQOL is a good instrument to measure HRQoL in pre-school children, but more research is needed to evaluate the psychometric performance of the TAPQOL in different clinical populations.

In *Chapter 6,* the relationship of preterm birth with HRQoL is examined for children aged 1-4. From the study groups in Chapter 5, three gestational age groups were selected, < 32 weeks (n=65) , 32-37 weeks (n=41), ≥37 weeks (n=54), and a reference group from the open population (n=50). The main instrument was the TAPQOL, which was completed by the parents, and provided HRQoL as well as HS scores. Other outcome measures obtained from parents or neonatologists were investigated in addition. It was shown that children born <32 weeks had significantly lower HRQoL than the reference group on the scales for lungs, stomach, eating disorders, motor functioning, communication and anxiety. We found differences between the neonatologist and the parent in perception of the child's situation, which can have clinical consequences. Parents of children born very preterm noticed many motor problems(=HS), but considered less emotional impact (=HRQoL) than would be expected from this number of problems. The judgement of the neonatologist related to the HS but not to the HRQoL as obtained by the parents. As a result, parents evaluated motor problems as being less of an emotional burden than the neonatologist does. In reverse, parents evaluated problems relating to lungs, stomach and sleeping as being more of an emotional burden than the neonatologist did. It can be stated that neonatal intensive care after birth has HRQoL implications for all children, particularly in children born after <32 weeks of gestation.

Chapters 7 to 10 deal with HRQoL in a longitudinal perspective. *Chapter 7* explores the time variability of QoL in children between 0 to 12 years of age. This was done by means of a systematic review of original studies, with at least two QoL assessments, and published between 1966 and 1998. The publications were identified from medical and psychological sources by computerised searches followed by manual selection. Thirty-two publications were selected and discussed according to their general characteristics, QoL assessment, longitudinal QoL research design and approaches to what changes QoL. It appeared that only two publications met all QoL assessment requirements (multi-factorial, self-administered, subjective) as well as longitudinal requirements (clear description of assessment period, recall period, sample

size at end of study, within-subject statistics). The approach to change that underlay the 32 publications can be described as: stable physical health gives stable QoL and changes in physical health changes QoL. It is rarely acknowledged that psychological, social and situational variables can change QoL as well. It can be concluded that there is a need for more studies that meet QoL assessment as well as longitudinal requirements. In the future, discussion is necessary about what exactly changes QoL, as this influences the planning of the assessments and guides the interpretation of changes.

Chapters 8 and 9 both study the same sample consisting of 688 children, born preterm in 1983 with a gestational age of less than 32 weeks and/or a birthweight of less than 1500 grams. Whereas Chapter 8 places the accent on the methodology, Chapter 9 focuses on the paediatric details. ***Chapter 8*** presents a strategy for analysing longitudinal QoL data that suffer from differences in measurement instruments over time. The strategy was applied to a set of longitudinal data from the cohort of preterm born children. HS data, differently defined at 5, 9 and 10 years of age, were prepared for longitudinal analyses with qualitative and quantitative item selection. Expert ratings fitted the data into physical, psychological and social HS domains. Principal component analysis (PCA) was used to match the data between measurements. Longitudinal PCAs were performed using the matched HS data. The impact of background variables such as gender and birthweight on HS changes was studied. It was concluded that this strategy to reconstruct and combine an imperfect data set, provided valuable information about the development of HS in preterm born children.

Chapter 9 describes the long-term effects of the complications that accompany a preterm birth. The HS at the age of 5 and 10 years in the cohort of children born preterm was studied to determine the impact of preterm birth on HS development. Prospectively collected HS variables, obtained from the parents, were analysed in a longitudinal perspective, using principal component analyses. One third of the sample had minor to severe HS problems at both ages of measurement. One third had problems on one assessment only. The remainder of the sample had no HS problems at either age. The analyses grouped the HS variables into three combinations: Problems in basic functioning, such as mobility or speech, decreased with age. Negative moods substantially increased, and concentration problems increased slightly. Specifically at risk (for HS problems) were preterm born children with handicaps, boys, and children born small for gestational age. In conclusion, according to the parents, one third of the cohort had no HS problems at either age. The pattern of HS problems of the preterm born children changed between 5 and 10 years of age.

In *Chapter 10,* changes in HRQoL and self-esteem Are studied in children with Idiopathic Short Stature (ISS) participating in a prospective randomised controlled study on the effect of Growth Hormone (GH) treatment. The sample consisted of forty prepubertal children (age 4 to 10 years old at start) with ISS (height <-2 SDS). The children were randomly assigned to a treatment or control group. HRQoL and self-esteem were assessed three times: shortly after randomisation (T1), and one (T2) and two years (T3). Children with ISS, their parents and the paediatrician completed questionnaires. At T1, children with ISS did not have a lower HRQoL and self-esteem than the norm population, except for the domain of social functioning as reported by children and parents. Children, parents and physician assessed HRQoL differently: At T3, children of the treatment group reported in some scales lower HRQoL and self-esteem than the control group did. The parent reports did not differ between groups, but the physician reported improved HRQoL in the treatment group. Changes in HRQoL and self esteem between T2 and T3 hardly related to growth (objectively measured or as perceived by the child). Instead, changes in several HRQoL and self-esteem scales related to the height appreciation by the child her/him self. The assumption that GH treatment improved HRQoL in children with ISS could not be supported in this study.

Chapter 11 contains the general discussion of the thesis, in which four themes are discussed that are relevant in the study of HRQoL in children. Firstly, it was shown that children as well as parents distinguish between the HS problems they observe, and the emotional appreciation of these problems incorporated in HRQoL. This justifies the distinction we made between the definitions of HRQoL and HS. Secondly, it was concluded that scientific tools are available to measure HRQoL in children. Two instruments were exhaustively described in this thesis (TACQOL and TAPQOL) and appeared reliable and valid. Thirdly, it appeared that children, parents and physicians assessed HRQoL differently. As all three reports had their value, it was recommended using all informants whenever possible. Finally, it was stressed that when the term *health related* QoL is used, it should be recognised that health contains physical, psychological as well as social health. This implies that when changes in physical health cannot be realised in children with a chronic illness, changes in psychological and social health may be able to improve the HRQoL of these children.

Samenvatting

Hoofdstuk 1 leidt het onderwerp van dit proefschrift in, te weten het bestuderen van de Gezondheidsgerelateerde Kwaliteit van Leven bij kinderen van 0 tot 12. Doordat het aantal medische successen toenam, ontstond er behoefte aan een uitkomstmaat die de verschillende manieren waarop kinderen kunnen reageren op bepaalde fysieke problemen tot uitdrukking brengt. Al houdt de individuele kinderarts wel rekening met de verschillen tussen kinderen, een systematisch en kwantificeerbare uitkomstmaat ontbrak. Daar is in voorzien door de introductie van het construct Gezondheidsgerelateerde Kwaliteit van Leven (GGKvL), ook wel genoemd Kwaliteit van Leven (KvL). Omdat een algemeen geaccepteerde definitie of theoretisch raamwerk ontbreekt, definiëren wij GGKvL als: de door de persoon waargenomen problemen in diens gezondheidstoestand, gecombineerd met de affectieve reactie op dergelijke problemen. Dat wat een persoon zelf waarneemt als diens gezondheid noemen we de Gezondheidstoestand (GT). De 'gezondheids' component in GGKvL bestaat uit het fysiek, psychologisch en sociaal functioneren. Tevens onderschrijven we dat GGKvL multi-factoreel is (fysiek, psychologisch en sociaal welzijn), door de patiënt zelf moet worden verschaft, subjectief is en veranderlijk over de tijd. In dit proefschrift werd de bruikbaarheid van deze definitie onderzocht.

Hoofdstuk 2 tot en met 6 beschrijven hoe GGKvL bij kinderen moet worden gedefinieerd en gemeten. In *Hoofdstuk 2* wordt de ontwikkeling van de 56-vragen bevattende TNO-AZL Child Quality Of Life (TACQOL) vragenlijst gepresenteerd. De vragenlijst is ontwikkeld om tegemoet te komen aan de behoefte naar een betrouwbaar en valide instrument voor het meten van GGKvL bij kinderen. GGKvL werd gedefinieerd als GT in zeven schalen plus de emotionele reactie op problemen in GT. De TACQOL geeft expliciet de mogelijkheid aan respondenten om onderscheid te maken tussen hun functioneren en hoe ze zich daarbij voelen. Een aselecte steekproef van 1789 ouders van 6 t/m 11 jarigen en 1159 kinderen van 8 t/m 11 jaar zelf, vulden de TACQOL in. Multipele correspondentie analyses tonen aan dat de antwoordcategorieën ordinaal zijn, en dat de TACQOL schalen als metrisch kunnen worden beschouwd. Cronbach's alpha varieerde van 0,65 t/m 0,84. Slechts 57% van de gerapporteerde GT problemen gingen samen met negatieve emoties, daarmee onze definitie van GGKvL ondersteunend. Intraclass correlatie coëfficiënten tussen ouder en kind versie varieerden van 0,44 t/m 0,61. Pearson correlatie coëfficiënten tussen de TACQOL en de Nederlandse versie van een Duits GGKvL instrument (KINDL)

varieerden van 0,24 t/m 0,60. Uni-variate variantie analyses lieten zien dat kinderen met chronische ziektes en kinderen onder medische behandeling lagere TACQOL scores hadden dan gezonde kinderen. Het onderzoek laat zien dat met de TACQOL de GGKvL van kinderen kan worden gemeten op betrouwbare en valide wijze. Bovendien leverde het onderzoek een grote norm groep op voor kinderen van 6 t/m 11 jaar. In

Hoofdstuk 3 en 4 is de relatie bestudeerd tussen kinder- en ouder rapportage aangaande de GGKvL van het kind. ***Hoofdstuk 3*** evalueert de overeenstemming tussen kinder- en ouder rapportage in een representatieve steekproef van 1105 Nederlandse kinderen (leeftijd 8 t/m 11). Zowel de kinderen als hun ouders vulden de TACQOL in. De Pearson correlaties tussen kind en ouder rapportage lagen tussen de 0,44 en 0,61 ($p<0,001$) en de intraclass correlations tussen de 0,39 en 0,62. Gemiddeld genomen rapporteerden de kinderen een significant lagere GGKvL dan ouders op de schalen: fysieke klachten, motorisch functioneren, zelfredzaamheid, cognitief functioneren en positieve emoties (gepaarde t-toets: $p<0,05$). Op alle schalen was de overeenstemming gerelateerd aan de hoogte van de GGKvL scores en bij sommige schalen tevens aan enkele achtergrond variabelen (sekse, leeftijd, tijdelijke ziektes, arts bezoek). Volgens Multitrait-Multimethod analyses zijn de GGKvL evaluaties van zowel kind als ouder valide. Het is daarom aan te bevelen om zo mogelijk beide evaluaties te gebruiken.

Hoofdstuk 4 bestudeert de overeenstemming tussen kinder- en ouder rapportage aangaande de GGKvL van kinderen, in een steekproef van 416 Nederlandse kinderen met een chronische ziekte (leeftijd 8 t/m 15). Zowel de kinderen als hun ouders vulden de TACQOL in. De correlaties tussen kind en ouder rapportage varieerden van -0,10 t/m 0,99 tussen de verschillende chronische condities. Kinderen rapporteerden een lagere GGKvL op de schalen: fysieke klachten, motorisch functioneren en positieve emoties. Ouders rapporteerden een lagere GGKvL op de sociaal functioneren en negatieve emotie schalen. Op alle schalen was de overeenstemming gerelateerd aan de soort chronische ziekte. Kind en ouder verschaffen ieder andere informatie over GGKvL. Door kennis te nemen van beide evaluaties kan de zorg voor kinderen met een chronische ziekte en hun ouders verbeterd worden.

Hoofdstuk 5 beschrijft de ontwikkeling van de 46-vragen bevattende TNO-AZL Pre-school Quality Of Life (TAPQOL) vragenlijst. GGKvL is hier gedefinieerd als GT in 13 schalen plus de emotionele reactie op problemen in GT. De TAPQOL dient te worden ingevuld door de ouders. Een steekproef van 121 ouders van prematuur geboren kinderen en 362 ouders van kinderen in de open populatie vulden de TAPQOL in. Op basis van Cronbach's alpha, itemrest correlaties en factor analyses zijn de TAPQOL schalen geconstrueerd, daarbij gebruik makend van de data van de

prematuren groep. De psychometrische prestaties van deze schalen werden geëvalueerd met zowel de premature groep als de open populatie groep. Cronbach's alpha varieerde van 0,66 t/m 0,88 bij de prematuren groep en van 0,26 t/m 0,85 voor de open populatie groep. De uni-dimensionaliteit van de afzonderlijke schalen werd in beide groepen bevestigd door de principale component analyse. De Pearson correlaties tussen de schalen waren laag. De T-toetsen tonen dat zeer prematuren en kinderen met chronische ziektes lagere scores hadden (dus slechtere GGKvL) op de TAPQOL schalen dan gezonde kinderen. Dit onderzoek heeft laten zien dat de TAPQOL een goed instrument is voor het meten van GGKvL bij peuters en kleuters. Desalniettemin is er meer onderzoek nodig waarbij de psychometrische prestaties van de TAPQOL ook in andere klinische groepen worden geëvalueerd.

In *Hoofdstuk 6* is de relatie tussen vroeggeboorte en GGKvL onderzocht bij kinderen van 1 t/m 4 jaar. Van de onderzoeksgroepen uit Hoofdstuk 5 zijn drie groepen geselecteerd volgens zwangerschapsduur: < 32 weken (n=65), 32-37 weken (n=41), ≥ 37 weken (n=54) en ook een referentie groep uit de open populatie (n=50). Hoofdinstrument is de TAPQOL, die zowel GGKvL als GT scores verschaft en werd ingevuld door de ouders. Ter aanvulling werden enkele andere uitkomstmaten verkregen van ouders of neonatologen. Kinderen geboren op < 32 weken hadden een significant lagere GGKvL dan de referentie groep voor de schalen longen, maag, eetproblemen, motorisch functioneren, communicatie en angst. We vonden verschillen tussen de neonatoloog en de ouder in hun waarneming van de situatie van het kind, wat klinische consequenties kan hebben. Ouders van zeer prematuur geboren kinderen vermeldden veel motoriek problemen (=GT), maar beschouwden de emotionele impact (=GGKvL) als minder dan zou worden verwacht op grond van het aantal problemen. De beoordeling van de neonatoloog relateerde aan GT maar niet aan GGKvL. Dit betekend dat ouders motoriek problemen als minder emotioneel belastend zagen dan de neonatologen. Daar tegenover staat dat ouders problemen gerelateerd aan longen, maag en slaap als meer emotioneel belastend zagen dan de neonatologen. Er kan worden geconcludeerd dat neonatale intensive-care na de geboorte gevolgen heeft voor alle kinderen, maar vooral voor kinderen geboren na een zwangerschap van minder dan 32 weken.

Hoofdstuk 7 t/m 10 behandeld GGKvL in een longitudinaal perspectief. *Hoofdstuk 7* exploreert de veranderlijkheid van KvL door de tijd bij kinderen tussen de 0 en 12 jaar. Dit is gedaan middels een systematisch overzicht van oorspronkelijke onderzoekingen, met minstens twee KvL metingen, gepubliceerd in de periode 1966-

1998. De publicaties werden geïdentificeerd uit medische en psychologische computerbronnen middels een gecomputeriseerde zoekstrategie, gevolgd door handmatige selectie. Tweeëndertig publicaties werden geselecteerd en bediscussieerd aan de hand van hun algemene karakteristieken, KvL meting, longitudinale KvL onderzoeksopzet en de ideeën over waardoor KvL veranderd wordt. Het bleek dat slechts twee publicaties tegemoet kwamen aan zowel alle voorwaarden voor een KvL meting (multi-factorieel, door de patiënt zelf verschaft, subjectief) als aan de voorwaarden voor een longitudinale studie (duidelijke beschrijving van afname periode, referentie periode, steekproef-grootte aan het einde van de studie, longitudinale statistische methoden). De benadering van verandering in KvL ten grondslag aan de 32 publicaties kan worden beschreven als: stabiele fysieke gezondheid geeft stabiele KvL, en veranderingen in fysieke gezondheid verandert de KvL. Er werd zelden rekenschap gegeven van het feit dat psychologische, sociale en situationele variabelen ook de KvL kunnen veranderen. Er is behoefte aan meer onderzoekingen die zowel voldoen aan de voorwaarden voor KvL metingen als aan die voor longitudinale metingen. In de toekomst is discussie nodig over wat precies KvL veranderd, omdat dit de planning van metingen beïnvloed en een leidraad is bij het interpreteren van veranderingen.

Hoofdstuk 8 en 9 bestuderen beide een steekproef bestaande uit 688 kinderen, te vroeg geboren in 1983 met een zwangerschapsleeftijd van minder dan 32 weken en/of een geboortegewicht van minder dan 1500 gram. Waar Hoofdstuk 8 de nadruk legt op de methodologie, benadrukt Hoofdstuk 9 de pediatrische details.

Hoofdstuk 8 presenteert een strategie voor het analyseren van longitudinale KvL-data, lijdend aan het gebruik van verschillende meetinstrumenten gedurende de tijd. De strategie is toegepast bij een longitudinale data set van een cohort te vroeg geboren kinderen. GT-data, verschillend gedefinieerd op 5, 9, en 10 jaar, is geprepareerd voor longitudinale analyses met behulp van kwalitatieve en kwantitatieve item selectie. Expertoordelen sorteerden de variabelen volgens fysieke, psychologische en sociale GT domeinen. Principale component analyse (PCA) is gebruikt voor het aan elkaar paren van data afkomstig van de verschillende meetmomenten. Longitudinale PCA's zijn uitgevoerd met de gepaarde GT-data. Er is bestudeerd wat de invloed was van achtergrond variabelen als sekse en geboortegewicht op GT veranderingen. Deze strategie voor het reconstrueren en combineren van een onvolkomen dataset verschaft waardevolle informatie over de ontwikkeling van GT bij te vroeg geboren kinderen.

Hoofdstuk 9 beschrijft het lange termijn effect van de complicaties die een vroeggeboorte vergezellen. De GT van een cohort van te vroeg geborenen is bestudeerd op 5 en 10 jarige leeftijd, teneinde de impact van vroeggeboorte op GT ontwikkeling

te bepalen. Prospectief verzamelde GT-variabelen, verkregen van de ouders, zijn geanalyseerd in een longitudinaal perspectief, daarbij gebruik makend van principale component analyses. Eenderde van de studiegroep had geringe tot ernstige GT problemen op beide gemeten leeftijden. Eenderde had problemen tijdens slechts een van de metingen. De rest van de groep had geen GT problemen op beide leeftijden. De analyses groepeerden de GT-variabelen in drie combinaties: Problemen in basaal functioneren, zoals mobiliteit en spraak, namen af met de leeftijd. Negatieve stemmingen namen substantieel toe, en concentratie problemen namen licht toe. Het meeste risico (op GT-problem) liepen te vroeg geboren kinderen met handicaps, jongens en kinderen die te klein voor de zwangerschapsduur waren. Concluderend, volgens de ouders had eenderde van het cohort geen GT problemen op beide leeftijden. Het patroon van GT problemen bij de te vroeg geboren kinderen veranderde tussen 5 en 10 jaar.

In *Hoofdstuk 10* zijn de veranderingen in GGKvL en zelfwaardering bestudeerd bij kinderen met een Idiopathische Kleine Gestalte (IKS) participerend in een prospectief gerandomiseerd en gecontroleerd onderzoek naar het effect van Groei Hormoon (GH) behandeling. De steekproef bestond uit veertig prepubertale kinderen (leeftijd 4 t/m 10 jaar bij aanvang) met IKS (lengte < -2SDS). De kinderen werden aselect toegewezen aan een behandel of controle groep. GGKvL en zelfwaardering werden drie keer gemeten: kort na de randomisering (T1), een jaar (T2) en twee jaar later (T3). Kinderen met IKS, hun ouders en de kinderarts vulden de vragenlijsten in. Op T1 bleken kinderen met IKS geen lagere GGKvL en zelfwaardering te hebben dan de norm-populatie, behalve voor de schaal sociaal functioneren zoals gerapporteerd door kinderen en ouders. Kinderen, ouders en arts beoordeelden de GGKvL verschillend: Op T3 rapporteerden kinderen uit de behandelgroep bij enkele schalen een lagere GGKvL en zelfwaardering dan de controlegroep. De ouder rapportage verschilde niet tussen de groepen, maar de arts rapporteerde verbeterde GGKvL bij de behandelgroep. Veranderingen in GGKvL en zelfwaardering tussen T2 en T3 relateerden nauwelijks aan groei (objectief gemeten of als waargenomen door het kind). Daar in tegen relateerden enkele GGKvL- en zelfwaarderings-schalen aan de waardering van de lengte door de kinderen zelf. De aanname dat GH-behandeling de GGKvL zou verbeteren bij kinderen met IKS, kan niet worden gesteund door dit onderzoek.

Hoofdstuk 11 bevat de algemene discussie van dit proefschrift. Vier thema's die relevant zijn bij het bestuderen van GGKvL van kinderen zijn besproken. Ten eerste is aangetoond dat zowel kinderen als ouders een onderscheid maken tussen de GT-

problemen die zij observeren en de emotionele waardering van deze problemen zoals opgenomen in GGKvL. Dit rechtvaardigt het onderscheid dat we maakten tussen de definities van GGKvL en GT. Ten tweede kan geconcludeerd worden dat wetenschappelijke instrumenten aanwezig zijn waarmee GGKvL bij kinderen kan worden gemeten. Twee instrumenten zijn uitputtend beschreven in dit proefschrift (TACQOL en TAPQOL) en blijken betrouwbaar en valide. Ten derde blijkt dat kinderen, ouders en artsen GGKvL verschillend inschatten. Omdat alle drie de rapportages hun waarde hebben wordt aanbevolen om, waar mogelijk. alle informanten te gebruiken. Tenslotte is benadrukt dat wanneer de term GGKvL wordt gebruikt men er rekenschap van dient te geven dat gezondheid zowel fysieke, psychologische als sociale gezondheid bevat. Dit impliceert dat wanneer de fysieke gezondheid bij kinderen met een chronische ziekte niet kan worden veranderd, dat dan veranderingen in psychologische en sociale gezondheid mogelijk de GGKvL van deze kinderen kunnen verbeteren.

Nawoord

In deze dissertatie wordt Gezondheidsgerelateerde Kwaliteit van Leven bij kinderen besproken. Een van de conclusies in het proefschrift is dat kwaliteit van leven niet slechts door fysieke veranderingen kan worden beïnvloed maar ook door psychologische, sociale en situationele variabelen. De leefomgeving en de mensen daarin kunnen dus grote invloed uitoefenen op iemands kwaliteit van leven. De volgende mensen wil ik dan ook bedanken voor hun bijdragen aan het verhogen van mijn kwaliteit van leven.

Allereerst wil ik de kinderen en ouders bedanken die deel uit maakten van de besproken onderzoeksgroepen. Zonder hun inspanningen was dit onderzoek niet mogelijk geweest. Verder wil ik de dames van het secretariaat Kindergeneeskunde en de heren van de foto-kamer bedanken. Als zij niet steeds voor me klaar stonden wanneer er iets geregeld moest worden, zou ik dit millennium niet meer kunnen promoveren. Ook de studenten die de afgelopen jaren hun bijdragen hebben geleverd dank ik hartelijk. Wat betreft de collegae van de Universiteit Leiden en van TNO Preventie & Gezondheid die betrokken waren bij de totstandkoming van dit proefschrift: ik hoop dat ik altijd mijn dankbaarheid heb laten blijken aan ieder die mijn kwaliteit van leven verhoogde.

Veel belangrijke gebeurtenissen in het leven hebben een 'er voor' en een 'er na'. Uit de periode vóór mijn promotieonderzoek, dank ik Peter Heymans van de Vakgroep Ontwikkelingspsychologie Utrecht omdat hij mij in het begin van mijn loopbaan enthousiast heeft gemaakt voor het doen van onderzoek. Uit de periode ná mijn promotie-onderzoek, dank ik Denise de Ridder en Jozien Bensing van de Capaciteitsgroep Gezondheidspsychologie Utrecht, die me de gelegenheid gaven mijn proefschrift als postdoc af te ronden. Mijn andere nieuwe collegae dank ik eveneens voor hun steun bij de 'laatste loodjes'.

Op deze plaats wil ik natuurlijk ook mijn familie en vrienden noemen die me gedurende dit promotie project met hun verbazing, trots, afgrijzen, enthousiasme, medeleven en verwondering hebben begeleid. In het bijzonder bedank ik Marijke Theunissen die niet alleen mijn moeder maar ook mijn eerste mentor was.

Als laatste richt ik mijn dankwoord (op z'n Twents) tot degene door wiens liefde ik me altijd gesteund heb gevoeld: Robbert, bedankt.

Curriculum Vitae

Nicolet Theunissen werd op 12 november 1964 geboren te Roosendaal. In 1983 behaalde zij het eindexamen Atheneum B aan het Norbertus College te Roosendaal. Daarna studeerde zij psychologie aan de Universiteit Utrecht en behaalde haar doctoraal examen in 1990 met als hoofdrichting Ontwikkelingspsychologie van de individuele levensloop. Van 1991 tot en met 1994 werkte ze als toegevoegd docent en onderzoeker bij de Vakgroep Ontwikkelingspsychologie van de Universiteit Utrecht. In 1992/1993 deed zij een literatuur studie naar de levenskwaliteit van jeugdige nierpatiënten, in opdracht van de Psychosociale afdeling van het Wilhelmina Kinderziekenhuis te Utrecht en de Nierpatiëntenvereniging LVD.

Van december 1994 tot en met februari 1999 had ze een aanstelling als assistent in opleiding, gefinancierd door de Vakgroep Kindergeneeskunde van de Universiteit Leiden en TNO Preventie & Gezondheid afdeling Jeugd. In die periode verrichtte zij het onderzoek dat leidde tot dit proefschrift.

Sinds april 1999 is zij als Postdoc-onderzoeker verbonden aan de Capaciteitsgroep Gezondheidspsychologie aan de Universiteit Utrecht, gelieerd aan de onderzoeksschool Psychologie & Health. De centrale vraag van haar huidige onderzoek is: Zijn hypertensie patiënten eerder geneigd de behandelingsadviezen van hun huisarts op te volgen, als de arts meer rekening houdt met hoe de patiënt tegen diens ziekte en behandeling aankijkt?