

A Brief Version of the Quality of Life in Short Stature Youth Questionnaire - the QoLISSY-Brief

Anja C. Rohenkohl¹ · Monika Bullinger¹ ·
Andreas M. Pleil² · Levente Kriston¹ ·
Julia H. Quitmann¹ · the QoLISSY Study Group

Accepted: 4 November 2015 / Published online: 21 November 2015
© Springer Science+Business Media Dordrecht 2015

Abstract The Quality of Life in Short Stature Youth (QoLISSY) questionnaire measures health-related quality of life (HrQoL) in short statured children (8–18 years) from patient and parent perspectives. To minimize respondent burden when assessing HrQoL in clinical practice, a brief version of the currently available 22-item QoLISSY is needed. The dataset of the European QoLISSY study ($N=268$) was divided into two subdatasets at random, one to identify the items of the brief questionnaire and the other to test the operating characteristics (reliability and validity). Concept-based construction involved the selection of three items per Quality of Life (QoL)-dimension (physical, social, emotional) according to the highest corrected item-scale correlation. Psychometric properties were inspected in terms of reliability and validity supplemented by testing item fit statistics to examine item response theory (IRT) compliance. Cronbach's alpha for the 9-item version was 0.89 for the patient- and the parent report. Pearson's correlations with the generic KIDSCREEN questionnaire were low to moderate (children: $r=0.17-0.58$; parents: $r=0.12-0.56$). Shorter children reported significantly poorer QoL (mean difference 15.39 points; $p<0.001$) than taller children, indicating known-groups validity. Results from IRT analysis showed an acceptable fit to Masters' Partial Credit Model. The 9-item QoLISSY brief version has satisfactory operating characteristics and is an efficient alternative for use in research and practice.

Electronic supplementary material The online version of this article (doi:10.1007/s12187-015-9350-2) contains supplementary material, which is available to authorized users.

✉ Anja C. Rohenkohl
a.rohenkohl@uke.de

¹ Department of Medical Psychology, University Medical Center Hamburg – Eppendorf, Martinistr. 52, Building W26, G-20246 Hamburg, Germany

² Global Health & Value Pfizer, Inc, San Diego, CA, USA

Keywords Children · Short stature · Quality of life · Questionnaire · Brief version

1 Introduction

Persons from a representative reference population with a height below -2 *SDS* (Standard Deviation Scores) meet the definition of short stature. About 125.000 children are born in Europe each year with a height below -2 *SDS* of the mean height of their peers according to age, gender and country of origin, therefore meeting the definition of short stature (SS) (Wit et al. 2008). In contrast to the statistical definition of SS, the clinical diagnosis is based on Growth Hormone Deficiency (GHD) as identified by a low Growth Hormone (GH) response to provocation tests. These children can be treated with daily injections of recombinant human GH to reach normal height in adulthood (Binder 2011). Treatment options for children with Idiopathic Short Stature (ISS) are limited because underlying pathology or lack of GH is not confirmed (Wit et al. 2008). Efforts were rarely made to assess the psychosocial burden of SS. The few existing studies which focus on HrQoL in SS youth describe consequences such as stigmatization, social isolation, “juvenilisation” and low self - esteem while other studies acknowledged the burden of the condition but found no impairments in mental health (Bullinger et al. 2009; Sandberg and Voss 2002; Voss and Mulligan 2000).

In general, HrQoL is defined as subjective physical, emotional and social aspects of well-being and functioning as perceived by patients and or observers (Bullinger 1991). HrQoL is a key treatment outcome that is increasingly considered important in pediatric endocrinology. It is a relevant treatment outcome from epidemiological and clinical perspectives and is broadly employed in health economic analyses as well (Bullinger et al. 2009). Although generic measures have been used to assess HrQoL in children with short stature, the **Quality of Life in Short Stature Youth (QoLISSY)** was the first condition specific instrument which focused on the perspectives of both the young patients and their parents and which could be applied across countries (Brütt et al. 2009).

The QoLISSY questionnaire was developed simultaneously in five European countries (the UK, Sweden, Spain, France and Germany) to assess HrQoL from the perspective of SS children and their parents (The European QoLISSY Group 2013). The aim was to provide a valid measure of outcome for use in both, clinical trials and clinical practice. However, a more efficient instrument for use in busy pediatric clinics prompted the development of a brief version of the QoLISSY. When developing the original QoLISSY questionnaire, care was taken to identify a core measure assessing physical (6 items), emotional (8 items) and social (8 items) domains of HrQoL (Bullinger 1991), as well as additional determinants of HrQoL such as perception of treatment (14 items), coping (10 items) and beliefs (4 items) about height to supplement the core measure. In the parent version, based on parent comments in the original content elicitation, two parent specific concepts were identified – the effect of the child’s SS on the family (11 items) and the perceptions of the parent regarding their child’s future (4 items). The QoLISSY core module with its 22 items showed satisfactory psychometric characteristics with a Cronbach’s α of 0.95 (range 0.84–0.88) for the child report and 0.95 (range 0.86–0.90) for the parent report across its different language versions. The QoLISSY questionnaire is fit for use in clinical research (Bullinger et al. 2013; Quitmann et al. 2013; The European QoLISSY Group 2013).

Since the length of instruments is perceived as a barrier against the implementation of measuring HrQoL in clinical practice, the development of short-form versions is recommended to simplify data collection (Varni et al. 2005). The challenge is to represent the content and domains of the original instrument in an abbreviated version without unduly affecting the instrument's psychometric performance and at the same time ensuring the maximum amount of shared variance between short and long versions (Muehlan 2010; Pollak et al. 2006).

A short form of a well-established questionnaire broadens its application potential in research and practice contexts by reducing the burden of data collection and the risk of item non-response (Jokovic et al. 2006; Ravens-Sieberer et al. 2010). The current paper describes the development and testing of a brief version of the QoLISSY questionnaire which aims at providing a valid measure of HrQoL e.g., in an outpatient pediatric clinic to be completed in less than 5 min. We were specifically interested in identifying which items best represent the QoLISSY core module with its 22 items. Furthermore we were interested in the psychometric properties of the QoLISSY brief version in terms of reliability and validity. Finally we wanted to characterize the operating characteristics of the brief version using Item Response Theory (IRT) performance. We hypothesize that the brief version is a psychometrically sound uni-dimensional scale, which can be used in place of the full 22 item version when a short form is indicated.

2 Materials and Methods

The original QoLISSY project focused on developing a condition-specific instrument across several European countries to capture the impact on HrQoL of short stature in children and adolescents (The European QoLISSY Group 2013). The objective was to provide a tool which could be used e.g., for measuring HrQoL as an outcome in pediatric clinical trials of growth promoting therapies, to give a voice to SS patients and their parents. Patients and parents can communicate their experience with SS, and this information can be used to improve patient care.

2.1 Development of the QoLISSY Questionnaire

The QoLISSY questionnaire was developed through focus groups, pilot testing including a cognitive debriefing process, and a field test with re-test (The European QoLISSY Group 2013). The psychometric properties of the child- and parent- reported versions of QoLISSY were examined in the field test in a total of 268 patients (134 children aged 8–12 years and 134 adolescents aged 13 to 18 years). A classical test-theoretical approach was chosen to construct the QoLISSY questionnaire, predicated on distributional characteristics, reliability, and the tenets of content and construct validity (The European QoLISSY Group 2013).

The original QoLISSY core module includes three dimensions, reflecting *physical* (6 items), *emotional* (8 items) and *social* (8 items) aspects of quality of life. Together the 22 items form the QoLISSY total score. All items are to be answered on a 5-point Likert scale (5 “not at all/ never”, 4 “slightly/ seldom”, 3 “moderately/ quite often”, 2 “very/ very often”, 1 “extremely/ always”) with higher values reflecting a more positive evaluation of HrQoL.

2.2 Sample Recruitment and Description

Children and adolescents diagnosed with GHD or ISS and one of their parents were recruited from participating clinical centers in the five European countries (the UK, Sweden, Spain, France and Germany) to assess the child's quality of life. Recruitment of families was performed by clinical centers according to the inclusion criteria age (young patients between 8 and 18 years and their parents as well as parents of young children (4–7 years old)), diagnosis (GHD or ISS) and treatment (GH treatment yes or no). Only data from participating families having given informed consent was forwarded, information on non-participants was not available. At date of diagnosis all children met the definition of short stature with a height of ≤ -2 standard deviations (*SDS*) below the mean, adjusted for age and gender (Ranke 1996). To be eligible, children had to be diagnosed as short statured according to patient files. Several children/ adolescents had achieved a normal height at the time of assessment due to growth with or without GH treatment earlier.

A total of 268 children and adolescents were included in the current analysis (129 children of 8 to 12 years and 139 adolescents of 13 to 18 years). Sample characteristics are shown in Table 1. Paired child and parent data was included in the analyses to compare the child and parent report in the brief version. Since eight parents did not answer the QoLISSY parent version in the field test, 260 parent–child dyads were used to test parent–child agreement. General analyses were conducted with a sample of $N=134$ children for the development as well as an N of 134 for the test sample (see Table 1). Results of a chi-square test indicated a similar distribution across the development and test samples in terms of patient age, gender, diagnosis, treatment status, and *SDS* height.

Table 1 Sample characteristics

		Total sample $N=268$	Sample A (development) $N=134$	Sample B (testing) $N=134$
Age group	8–12 years	129 (48.13 %)	60 (44.78 %)	69 (51.49 %)
	13–18 years	139 (51.87 %)	74 (55.22 %)	65 (48.51 %)
Gender	Female	114 (42.54 %)	55 (48.25 %)	59 (51.75 %)
	Male	154 (57.46 %)	79 (51.30 %)	75 (48.70 %)
Diagnosis	GHD	109 (40.67 %)	59 (54.13 %)	50 (45.87 %)
	ISS	159 (59.33 %)	75 (47.17 %)	84 (52.83 %)
Treatment status	Treated	142 (52.99 %)	72 (50.70 %)	70 (49.30 %)
	Untreated	126 (47.01 %)	62 (49.21 %)	64 (50.79 %)
<i>SDS</i> height*	$> -2SDS$	131 (48.88 %)	58 (44.27 %)	73 (55.73 %)
	$\leq -2SDS$	130 (48.51 %)	73 (56.15 %)	57 (43.85 %)

Annotation. *standard deviation score (*SDS*; taller ($> -2.0 SDS$), and shorter ($\leq -2.0 SDS$) children); seven children of the 268 did not report their actual height; therefore the n is 261 in this variable; *GHD* Growth Hormone Deficiency, *ISS* Idiopathic Short Stature

2.3 Item Selection - QoLISSY Brief Version

The field test dataset of the original validation study was used to develop and examine the psychometric properties of a brief version of the QoLISSY questionnaire, supplemented by testing the brief version for item response test (IRT) performance. For the current analysis, the total field test dataset was randomly split into two subsets; the “development sample” was used to generate the QoLISSY brief version and the “test sample” to examine its preliminary psychometric properties.

The initial task in developing the brief version was to identify within each of the domains those items that best represent the three core HrQoL subscales. This selection process was conducted using reliability indicators to identify items with the highest corrected item-scale correlations (*rtt*); i.e., correlation between the item and the remaining items within the scale and the lowest impact on alpha coefficients when items were omitted in the developmental sample (Coste et al. 1997; Wong et al. 2013). To facilitate comparisons between child or adolescent and parent ratings, child report items were used to construct the parent-reported QoLISSY brief version, items were only changed in wording to reflect the parent perspective. In the case of the parent version, the term “I” was changed to “My child” with the remaining item text and response categories unchanged. Just like the original full version, the parent-report version thus includes the identical items, i.e., nine core items in the QoLISSY- brief. All analyses reported above were conducted using the Statistical Package for the Social Sciences version 18 (SPSS Inc 2009).

2.4 Testing for IRT Assumptions

Once the proposed items for the brief version had been identified from the development sample, adequate functioning of rating scale categories was examined in the test sample (Las Hayas et al. 2010; Linacre 2009). The procedure was based on the following IRT assumptions regarding tests for uni-dimensionality: presence of more than 10 observations per answer category; a smooth distribution of category frequency (the frequency distribution is not jagged); clearly advancing average measures; and sufficient model fit (i.e., congruence between observed and expected values). Uni-dimensionality was assumed because of high scale inter-correlations indicating that the correlations relate to the same construct. To test for these assumptions, the frequency distribution across the five answer categories was inspected, which showed significant skewness with the majority of the responses in categories 4 and 5, reflecting better functioning.

In several cases, observed frequencies remained below 10 cases across categories and therefore did not meet the distribution assumptions outlined above. There was a marked skewness in the data with the majority of respondents indicating little or no problems consistently across the nine items. To allow for a more balanced distribution across responses, answers 1, 2 and 3 were collapsed into one category.

The assumed uni-dimensional structure of the QoLISSY brief was tested by conducting a confirmatory factor analysis using ordered categorical indicators and a weighted least square estimator. A normed chi-square value of 2.412, a Comparative Fit Index of 0.972, a Tucker Lewis Index of 0.963, and statistically significant ($p < 0.001$) positive factor loadings for all indicators suggested that the brief instrument meets the IRT criteria for uni-dimensionality (Las Hayas et al. 2010; Linacre 2009).

2.5 Psychometric Properties

Internal consistency was assessed using Cronbach's alpha, values above $\alpha=0.70$ were interpreted as acceptable (Cronbach 1951). Convergent validity was assessed based on the brief version's total score correlation with the KIDSCREEN 52 generic measure of QoL (The KIDSCREEN Group Europe 2006), expecting moderate correlations with thematically similar scales. Known groups validity analyses based on t-tests were conducted to test for differences according height ($> -2 SDS/ \leq -2 SDS$), both for child- and parent- reported data. In relation to t-tests, standardized effect sizes (Cohen's d), where a value of 0.2 is considered a small difference, 0.5 a medium or moderate difference, and 0.8 a large difference (Cohen 1988) were used as an indicator of the magnitude of potential clinically significant differences.

2.6 Concordance Between the 22-Item Version and the Brief Version and Between Parent-child QoL Ratings

Finally the Pearson's product-moment correlation between the QoLISSY 9-item brief version and the QoLISSY original 22-items total score was inspected to determine the amount of explanatory power retained in the brief version. In addition, Pearson's correlation and intraclass correlation coefficients (ICC) between child and parent report were inspected to examine the degree of agreement across respondent perspectives and within child-parent dyads.

2.7 IRT Performance

To test IRT performance, we fitted Masters' partial credit model (PCM; (Masters 1982) and performed general model tests. We calculated itemfit statistics (infit and outfit), with a good fit indicated by values between 0.7 and 1.3 (Bond and Fox 2001). Item difficulties and differential item functioning (DIF) for age, gender, treatment status, diagnosis, and *SDS* height were examined with Andersen's likelihood ratio test (Mair et al. 2012).

Missing values, present in less than 3 % of all items across the patient and parent samples, were replaced by mean substitution per case. Analyses were performed using SPSS version 18 (SPSS Inc 2009), Mplus 7.11 (Muthén and Muthén 2012), and R (R version 0.97.318, The R Foundation for Statistical Computing) which also included the eRm package (Mair et al. 2012).

3 Results

3.1 Item Selection - QoLISSY Brief Version

Item difficulties (means; range 1–5), standard deviations, corrected item-scale correlations, multiple correlations and Cronbach's alpha after item deletion for each of the original QoLISSY items in the core module were calculated. Results showed that for *physical aspects of quality of life* items 1, 3 and 4 met the item selection criteria in terms of the highest item-total correlations, the highest multiple correlation with the

remaining items, and the lowest impact on alpha coefficients when items were omitted. For *social aspects of quality of life*, criteria were met by items 2, 4 and 5, and for *emotional aspects of quality of life* items 1, 5 and 8 were selected. The items for the QoLISSY brief version were chosen on the basis of results from the child dataset (see Tables 2, 3 and 4) and these items were also used for the parent brief version.

3.2 Psychometric Properties

Inspecting the reliability of the items selected in the test sample indicated acceptably high item-total correlations with values between $r_{tt}=0.59$ and $r_{tt}=0.70$ for the child version and values between $r_{tt}=0.55$ and $r_{tt}=0.73$ for the parent reported version. Together, these results support the validity of a total score from the QoLISSY brief version, in line with the results of testing for uni-dimensionality (see Table 5).

Scores were transformed to a 0-100 scale, with higher scores indicating better QoL. The total mean score in the test sample was 63.93 ($SD=29.58$) for children and 58.21 for parents ($SD=29.71$). Cronbach's alpha for the brief version total score was 0.89 with a split half reliability of 0.87 for the child version. Item and total score reliability indices were nearly identical for the parent version in the test sample (Cronbach's α : 0.89; Split-half: 0.86).

Statistically significant differences between groups according to *SDS* status were found. Taller children (with a lower height *SDS*) reported higher HrQoL in the brief version (see Table 6). Differences in height deviation were found when inspecting parent ratings as well (see Table 6). The standardized effect sizes (Cohen's d) are shown in the last column of the tables and showed a large effect size for the difference between

Table 2 Reliability analysis – development sample: physical aspects of QoL

No.	Items*	Mean	SD	Corrected item-scale correlation (discrimination)	Cronbach's α if item deleted
1	.. prevents me from doing things...	4.02	1.11	0.609	0.795
2	... problems everyday.	4.40	0.88	0.559	0.808
3	I have more trouble reaching things....	3.57	1.27	0.751	0.761
4	Because of my height I depend on others.	4.09	1.21	0.683	0.778
5	I have to look up...	3.44	1.30	0.499	0.818
6	...others my age can go on fairground rides...	3.91	1.44	0.520	0.818

Annotation. *Items selected for the QoLISSY Brief Version are highlighted. Only shortened abstractions of the items are displayed

Table 3 Reliability analysis – development sample: social aspects of QoL

No.	Items*	Mean	SD	Corrected item- scale correlation (discrimination)	Cronbach's α if item deleted
1	.. height at school bothers me.	3.83	1.38	0.643	0.833
2	I feel small...	3.68	1.37	0.711	0.824
3	...mistake me for being younger	2.99	1.49	0.616	0.837
4	... I get laughed at or teased.	4.26	1.07	0.666	0.833
5	... I am treated differently.	4.32	.99	0.651	0.836
6	... only thing others notice...	4.11	1.18	0.563	0.842
7	...problems getting the clothes...	3.92	1.25	0.432	0.857
8	It hurts to be left ...	4.12	1.33	0.559	0.843

Annotation. *Items selected for the QoLISSY Brief Version are highlighted. Only shortened abstractions of the items are displayed

Table 4 Reliability analysis – development sample: emotional aspects of QoL

No.	Items*	Mean	SD	Corrected item- scale correlation (discrimination)	Cronbach's α if item deleted
1	... I feel different	3.96	1.26	0.725	0.848
2	I am fed....	3.50	1.62	0.677	0.855
3 I am shy.	4.36	1.07	0.563	0.865
4	I am happy	2.95	1.54	0.615	0.861
5	I am insecure	4.32	1.02	0.746	0.850
6	I am sad	4.12	1.11	0.655	0.857
7	... I feel comfortable	3.88	1.40	0.393	0.884
8	...height bothers ...	3.81	1.33	0.794	0.840

Annotation. *Items selected for the QoLISSY Brief Version are highlighted. Only shortened abstractions of the items are displayed

Table 5 Reliability analysis of the brief version (test sample)

Item	N	Mean (1–3)		SD		Corrected item-total correlation			
		Children	Parents	Children	Parents	Children	Parents		
Physical aspects of QoL	1	134	130	2.28	2.32	0.81	0.81	0.658	0.610
	3	134	130	1.87	2.08	0.80	0.76	0.653	0.551
	4	134	130	2.26	2.35	0.80	0.80	0.650	0.623
Social aspects of QoL	2	134	130	2.13	1.84	0.87	0.85	0.699	0.725
	4	134	130	2.43	2.26	0.76	0.81	0.692	0.694
	5	134	130	2.45	2.26	0.77	0.81	0.637	0.638
Emotional aspects of QoL	1	134	130	2.41	2.26	0.79	0.85	0.680	0.725
	5	134	130	2.41	2.26	0.79	0.83	0.592	0.586
	8	134	130	2.27	1.84	0.85	0.82	0.705	0.620

the two height groups in the parent report ($d=0.91$) and medium or moderate effect size for the other results ($d=0.56-0.69$).

Correlation coefficients of the QoLISSY- brief with the generic KIDSCREEN instrument were highest for self-perception and social acceptance (above $r=0.50$, see Table 7). As expected, correlations were lower with KIDSCREEN subscales not represented in the condition-specific QoLISSY (e.g., Autonomy, Parent-relation & Home Life and Financial Resources).

3.3 Concordance Between the 22-Item Version and the Brief Version and Between Parent–child QoL Ratings

The correlations between the brief and the full versions are examined to provide evidence that the short form, obtained after item reduction, explains a significant amount of variance in scores of the full form. The correlation between the original QoLISSY total score and the brief version total score was $r=0.95$ for both comparing the full and brief version in children and comparing the two versions in parents.

Table 6 Testing for group differences in SDS height

T-test (Student)		<i>M</i>	<i>SD</i>	<i>T</i>	<i>df</i>	<i>P</i>	<i>d</i>
Children	$>-2SDS$	71.14	27.43	-3.86	128	<0.001	0.69
	$\leq-2SDS$	51.27	30.04				
Parents	$>-2SDS$	67.01	28.05	-4.87	125	<0.001	0.91
	$\leq-2SDS$	42.39	26.20				

Annotation. Standard deviation score (*SDS*; taller (>-2.0 *SDS*), and shorter (≤-2.0 *SDS*) children)

Table 7 Correlations between the QoLISSY brief version total score and domains of the Kidscreen 52

KIDSCREEN 52	Children Brief version	Parents Brief version
Physical wellbeing	0.183	0.198
Psychological wellbeing	0.253	0.241
Moods & emotions	0.339	0.341
Self perception	0.517	0.556
Autonomy	0.172	0.269
Parent-relation & home life	0.222	0.123
Financial resources	0.182	0.121
Peers & social support	0.256	0.291
School environment	0.221	0.217
Social acceptance (Bullying)	0.583	0.514

Pearson's correlation between children and parents' brief version scores was $r=0.68$ and the intraclass correlation coefficient (ICC) was 0.70 for parent-child dyads.

3.4 IRT Performance

Results of the IRT analyses showed that the partial credit model fits the data well (Andersen's likelihood ratio test $p=0.477$, Loef's likelihood ratio test $p=0.504$). Outfit (range 0.784 to 1.290) and infit (range 0.798 to 1.201) item statistics were within acceptable limits (Bond and Fox 2001). Item thresholds were monotonically increasing for all items, but distances between category thresholds were below 1.4 (logit scale) for all items but item 2, suggesting limited discrimination (see Table 8).

Differential item functioning (DIF; alpha level=0.01) was not present across different subgroups (age, gender, diagnosis, *SDS* height, treatment status) which implies measurement invariance – the QoLISSY brief version is measuring the same thing for all respondents (Walker 2011).

Table 8 Item statistics for the QoLISSY brief version according to the partial credit model

Item	Chisquare	<i>df</i>	<i>p</i> -value	Outfit MSQ	Infit MSQ	Threshold 1	Threshold 2
physical#1	93.520	109	0.855	0.850	0.949	-0.371	0.604
physical#3	90.773	109	0.897	0.825	0.798	0.474	2.256
physical#4	141.860	109	0.019	1.290	1.201	-0.411	0.715
social#2	90.607	109	0.899	0.824	0.844	0.414	0.720
social#4	91.447	109	0.888	0.831	0.829	-0.811	0.059
social#5	108.847	109	0.486	0.990	0.935	-0.626	-0.140
emotion#1	91.826	109	0.882	0.835	0.868	-0.503	-0.018
emotional#5	127.041	109	0.114	1.155	1.127	-0.503	-0.018
emotional#8	86.207	109	0.947	0.784	0.817	0.126	0.244

4 Discussion

The question regarding which approaches and strategies should be employed to construct short forms of instruments has been debated in the literature (Jokovic et al. 2006; Newcombe et al. 2013). Different methods are available, and choice of the appropriate approach depends on the priorities for development and intended use (Coste et al. 1997; Ravens-Sieberer et al. 2010; World Health Organization 1996). Even though the construction of abbreviated versions of instruments can be diverse, the basis of construction should be clearly outlined and relevant test statistics should be used and results reported. The present paper describes the development of an IRT conform 9-item brief version of the QoLISSY questionnaire which conceptually reflects the underlying three-dimensional domain structure of the 22-item full version.

The current approach was based on a theoretical model of quality of life represented by three domains (physical, social and emotional aspects of QoL) that had been suggested in the WHO definition of health (World Health Organization 1946). This model was empirically substantiated in the original QoLISSY study (The European QoLISSY Group 2013; World Health Organization 1946). Following a theoretically based construction principle, the decision was made to retain this structure rather than derive the short form by identifying items that correlate highest with the original 22 items total score.

The results of the QoLISSY brief version are in line with the results found with the 22-item version still the nine items were selected using a priori criteria for short form construction (Coste et al. 1997; Wong et al. 2013) with the aim of ensuring an optimal representation of each of the three QoLISSY core domains. One half of the original QoLISSY validation sample was used to determine which QoLISSY items should form the brief version, and the other half to test its psychometric properties and IRT performance. These findings must be considered in light of methodological limitations. The correlation between the total scores of the brief and full version was high, indicating that the brief version explains a major proportion of variance ($r=0.95$, $p<0.01$) of the full version. This information however should not be mistaken to imply concurrent validity of the brief version, which has to be tested in an independent sample. It should be noted that testing the brief version within the original QoLISSY sample is not additional evidence of validity but rather an indication of successful development of the abbreviated instrument. Also the composition of the sample was unbalanced because of a covariation between clinical characteristics such as diagnosis and treatment status. Therefore, the brief version needs to be examined and tested in independent samples in order to demonstrate its value as a reliable and valid measure of HrQoL in SS children and adolescents and their parents. In addition, more information about the families who gave their informed consent compared to the ones who rejected participation in the study would be needed to control for selection bias.

Though the distribution of responses in this sample suggests that the response category choices could be reduced from 5 to 3 levels, at this time the 5 level response choice will be retained in the instrument as seen by the respondent. The scoring of the 9-item brief version could be based on the collapsed category set if the data are sparse in the higher impairment categories as we have done here. But the ability to differentiate in populations who report a different distribution of their answers than ours (i.e., are more affected in their quality of life) will be lost by the reduction of the 5-point answer category.

Results, however, suggest that the QoLISSY brief version adequately represents the full version, demonstrates acceptable psychometric characteristics in terms of reliability and validity and fulfills the IRT criterion of uni-dimensionality. Once replicated in an independent clinical sample, the QoLISSY – brief can be used in situations where a brief version is preferable to the full QoLISSY instrument. The psychometric testing of the long and brief version is relevant to the five European countries of the QoLISSY study. Further testing is needed in a wider range of linguistic, cultural and national contexts to examine the stability of the results in other countries.

While the original QoLISSY questionnaire with its 22- item core QoL module provides information on physical, social, and emotional subscales, the QoLISSY - brief version yields one uni-dimensional total score. This version was developed on the basis of a three- dimensional conceptual model for Item selection and IRT approach to examine uni-dimensionality, with subsequent psychometric testing for reliability and validity.

Its brevity and ease of administration as well as minimal respondent burden provides a benefit to clinical practitioners desiring to assess impairment in short statured patients. The QoLISSY - brief may be used to screen for individuals with an impaired height-related HrQoL so that appropriate treatment options can be considered. Ongoing management of patients under treatment may be enhanced by the ability to monitor the HrQoL impacts over time in a simple fashion. As a research tool, its quality and feasibility make it well suited for clinical studies, epidemiological cohort studies and population surveys. As such, the QoLISSY-brief contributes to broadening the application potential of patient- and parent reported outcomes in pediatric growth disorders.

4.1 Access to the QoLISSY Instrument

QoLISSY is a joint initiative between Pfizer Limited and the University Medical Center Hamburg - Eppendorf. Copyright Pfizer Limited all rights reserved. The QoLISSY instrument, together with comprehensive information of its development and validation process is published in the QoLISSY's User's Manual (Pabst Science Publishers, Lengerich, 2013). The Manual, which is available upon request, includes QoLISSY child and parent forms in all existing language versions, as well as scoring information. The QoLISSY 22-Item version as well as the brief version will be made available for bona fide research and clinical purposes via the Pfizer Patient Reported Outcome website (www.pfizerpatientreportedoutcomes.com). For those interested in conducting collaborative research projects directly with the University of Hamburg, please contact Dr. Quitmann (j.quitmann@uke.de) or Prof. Dr. Bullinger (bullinger@uke.de).

Acknowledgments The QoLISSY study group, in addition to the German team of Monika Bullinger and Julia Quitmann (University Medical Centre Hamburg-Eppendorf, Department of Medical Psychology, Hamburg, Germany), includes John E. Chaplin (Sahlgrenska Academy at University of Gothenburg, Department of Paediatrics, Växthuset, Queen Silvia's Children's Hospital, Gothenburg, Sweden), Michael Herdman (IMIM University, Insight Consulting and Research, Barcelona, Spain), Emmanuelle Mimoun (Children's Hospital of Toulouse, Department of Pediatric Endocrinology, Toulouse, France), Andreas Pleil (Pfizer, Inc. Outcomes and Evidence, Global Health & Value, San Diego, CA), Hartmut Wollmann (University of Tuebingen formerly

of Pfizer Ltd, Walton Oaks, UK) and Michael Power (Rivers Trauma Centre, Royal Edinburgh Hospital, Edinburgh). We thank Dr. Joseph C. Cappelleri from Pfizer Inc for his review of the manuscript. The original QoLISSY Development study was funded by Pfizer, Inc. Dr. Andreas Pleil is a full time employee of Pfizer. No authors received payment for their contributions to this manuscript.

References

- Binder, G. (2011). Growth hormone deficiency: new approaches to the diagnosis. *Pediatric Endocrinology Reviews*, 9, 535–537.
- Bond, T. G., & Fox, C. M. (2001). *Applying the rash model: Fundamental measurement in the human sciences*. Mahwah: Erlbaum.
- Brütt, A. L., Sandberg, D. E., Chaplin, J., Wollmann, H., Noeker, M., Koltowska-Haggstrom, M., & Bullinger, M. (2009). Assessment of health-related quality of life and patient satisfaction in children and adolescents with growth hormone deficiency or idiopathic short stature - part 1: a critical evaluation of available tools. *Hormone Research*, 72, 65–73. doi:10.1159/000232158.
- Bullinger, M. (1991). Quality of life: definition, conceptualization and implications—a methodologist's view. *Theoretical Surgery*, 6, 143–148.
- Bullinger, M., Koltowska-Haggstrom, M., Sandberg, D., Chaplin, J., Wollmann, H., Noeker, M., & Brutt, A. L. (2009). Health-related quality of life of children and adolescents with growth hormone deficiency or idiopathic short stature - part 2: available results and future directions. *Hormone Research*, 72, 74–81. doi: 10.1159/000232159.
- Bullinger, M., Quitmann, J., Power, M., Herdman, M., Mimoun, E., Debusk, K., Chaplin, J. E. (2013). Assessing the quality of life of health-referred children and adolescents with short stature: development and psychometric testing of the QoLISSY instrument. *Health and Quality of Life Outcomes*, 11. doi: 10.1186/1477-7525-11-76.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: Lawrence Erlbaum Associates, Inc.
- Coste, J., Guillemin, F., Pouchot, J., & Fermanian, J. (1997). Methodological approaches to shortening composite measurement scales. *Journal of Clinical Epidemiology*, 50, 247–252.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334.
- SPSS Inc. (2009). *PASW for windows, version 18.0*. Chicago: SPSS Inc.
- Jokovic, A., Locker, D., Guyatt, G. (2006). Short forms of the Child Perceptions Questionnaire for 11-14-year-old children (CPQ11-14): development and initial evaluation. *Health and Quality of Life Outcomes*, 4. doi: 10.1186/1477-7525-4-4.
- Las Hayas, C., Quintana, J. M., Padierna, J. A., Bilbao, A., Muñoz, P. (2010). Use of Rasch methodology to develop a short version of the health related quality of life for eating disorders questionnaire: a prospective study. *Health Qual Life Outcomes*, 29. doi: 10.1186/1477-7525-8-29.
- Linacre, J. (2009). *A user's guide to WINSTEPS*. Chicago: MESA Press.
- Mair, P., Hatzinger, R., Maier, M. J. (2012). eRm: Extended Rasch Modeling. R package version 0.15-1. Retrieved 30.06.2014, 2014, from <http://CRAN.R-project.org/package=eRm>.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149–174.
- Muehlan, H. (2010). *Developing the DCGM-12: A short-form of the DISABKIDS condition-generic module assessing health related quality of life in children and adolescents with chronic conditions* (Doctoral Thesis, University of Hamburg, Germany). Retrieved from http://ediss.sub.uni-hamburg.de/volltexte/2010/4923/pdf/DISS_HM_DCGM12_FIN_UPLOAD.pdf.
- Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide* (7th ed.). Los Angeles: Muthén & Muthén.
- Newcombe, P. A., Sheffield, J. K., & Chang, A. B. (2013). Parent cough-specific quality of life: development and validation of a short form. *Journal of Allergy and Clinical Immunology*, 131, 1069–1074. doi:10.1016/j.jaci.2012.10.004.
- Pollak, E., Muehlan, H., von Mackensen, S., & Bullinger, M. (2006). The Haemo-QoL index: developing a short measure for health-related quality of life assessment in children and adolescents with haemophilia. *Haemophilia*, 12, 384–392.
- Quitmann, J., Rohenkohl, A., Bullinger, M., Chaplin, J., Herdman, M., Dolores, S., & Pleil, A. (2013). Parental perception of health-related quality of life in children and adolescents with short stature: literature review and introduction of the parent-reported QoLISSY instrument. *Pediatric Endocrinology Reviews*, 11, 147–160.

- Ranke, M. B. (1996). Towards a consensus on the definition of idiopathic short stature. *Hormone Research*, 45, 64–66.
- Ravens-Sieberer, U., Erhart, M., Rajmil, L., Herdman, M., Auquier, P., Bruil, J., & European KIDSCREEN Group. (2010). Reliability, construct and criterion validity of the KIDSCREEN-10 score: a short measure for children and adolescents' well-being and health-related quality of life. *Quality of Life Research*, 19, 1487–1500. doi:10.1007/s11136-010-9706-5.
- Sandberg, D. E., & Voss, L. D. (2002). The psychosocial consequences of short stature: a review of the evidence. *Best Practice & Research Clinical Endocrinology & Metabolism*, 6, 449–463.
- The European QoLISSY Group. (2013). *Quality of life in short stature youth. The QoLISSY questionnaire – user's manual*. Lengerich: Pabst Science Publishers.
- The KIDSCREEN Group Europe. (2006). *The KIDSCREEN questionnaires: Quality of life for children and adolescents - Handbook*. Lengerich: Pabst Science Publishers.
- Vami, J. W., Burwinkle, T. M., Lane, M. M. (2005). Health-related quality of life measurement in pediatric clinical practice: an appraisal and precept for future research and application. *Health and Quality of Life Outcomes*, 3. doi: 10.1186/1477-7525-3-34.
- Voss, L. D., & Mulligan, J. (2000). Bullying in school: are short pupils at risk? Questionnaire study in a cohort. *British Medical Journal*, 320, 612–613.
- Walker, C. M. (2011). What's the DIF? Why differential item functioning analyses are an important part of instrument development and validation. *Journal of Psychoeducational Assessment*, 29, 364–376.
- Wit, J. M., Clayton, P. E., Rogol, A. D., Savage, M. O., Saenger, P. H., & Cohen, P. (2008). Idiopathic short stature: definition, epidemiology, and diagnostic evaluation. [Review]. *Growth Hormone & IGF Research*, 18, 89–110. doi:10.1016/j.ghir.2007.11.004.
- Wong, G., Wai Lam, S., Ngai, K., Wong, A., Poon, W. S., Mok, V. (2013). Development of a short form of stroke-specific quality of life scale for patients after aneurysmal subarachnoid hemorrhage. *Journal of the Neurological Sciences*, Article in press.
- World Health Organization. (1946). *Preamble to the constitution of the world health organization*. New York: Paper presented at the International Health Conference.
- World Health Organization. (1996). *WHOQOL-BREF. Introduction, administration, scoring and generic version of the assessment*. Geneva.