

Cross-cultural equivalence of the patient- and parent-reported Quality of Life in Short Stature Youth (QoLISSY) Questionnaire

Monika Bullinger^{1*}, Julia Quitmann^{1*}, Neuza Silva^{1, 7}, Anja Rohenkohl¹, John E. Chaplin², Kendra DeBusk³, Emmanuelle Mimoun⁴, Eva Feigerlova⁴, Michael Herdman⁵, Dolores Sanz⁵, Hartmut Wollmann⁶, Andreas Pleil⁶, & Michael Power³

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

² Sahlgrenska Academy at University of Gothenburg, Department of Paediatrics, Växthuset, Queen Silvia's Childrens Hospital, 41685 Gothenburg, Sweden

³ University of Edinburgh, Department of Clinical Psychology, Teviot Place, EH8 9AG Edinburgh, Scotland

⁴ University of Toulouse, Department of Paediatric Endocrinology, CHU, 330 avenue de Grande-Bretagne FR-31059 Toulouse, France

⁵ IMIM University, Insight Consulting and Research, Cami Ral 266, Mataró, Barcelona, Spain

⁶ Pfizer Ltd., Specialty Care MDG, Endocrinology, Dorking Road, Walton-on-the-Hill, Tadworth Surrey, KT20 7NS Walton Oaks, UK

⁷ University of Coimbra, Faculty of Psychology and Educational Sciences, Rua do Colégio Novo, Apartado 6153, 3001-802 Coimbra, Portugal

Abstract

Background: Testing cross-cultural equivalence of Patient Reported Outcomes requires sufficiently large samples per country, which is difficult to achieve in rare endocrine paediatric conditions. We describe a novel approach to cross-cultural testing of the Quality of Life in Short Stature Youth (QoLISSY) questionnaire in five countries by sequentially Taking-One-Country-Out (TOCO) from the total sample and iteratively comparing the resulting psychometric performance. **Methods:** Development of the QoLISSY proceeded from focus group discussions over pilot testing to field testing in 268 short statured patients and parents. In exploring cross-cultural equivalence the iterative TOCO technique was used to examine and compare validity, reliability and convergence of QoLISSY patient- and parent-ratings in the field test data set, and to predict QoLISSY scores from clinical, socio-demographic and psychosocial variables. **Results:** Validity and reliability indicators were satisfactory for each sample after iteratively omitting one country. Comparisons with the total sample revealed cross-cultural equivalence in internal consistency and construct validity for patients and parents, high inter-rater agreement and a substantial proportion of QoLISSY variance explained by predictors. **Conclusion:** The TOCO technique is a powerful method to overcome problems of country specific testing of PRO instruments. It provides an empirical support to QoLISSY's cross-cultural equivalence and is recommended for future research.

Keywords

Health related quality of life, short stature, growth hormone deficiency, idiopathic short stature, patient and parent assessment, and outcome research.

Bullinger, M., Quitmann, J., Silva, N., Rohenkohl, A., Chaplin, J. E., DeBusk, K., Mimoun, E., Feigerlova, E., Herdman, M., Sanz, D., Wollmann, H., Pleil, A., & Power, M. (2014). Cross-cultural equivalence of the patient- and parent-reported Quality of Life in Short Stature Youth (QoLISSY) Questionnaire. *Hormone Research in Paediatrics*, 82(1), 18-30. doi:10.1159/000358832

1. Introduction

Assessment of health related quality of life in children and adolescents is increasingly regarded as an important topic in paediatrics. Measures to assess health related quality of life (HrQoL) in young persons have been developed for population studies and for clinical studies of children with acute or chronic health conditions [1].

A recent review of the paediatric literature shows that generic measures have primarily been used to assess child self-reported health. Instruments have also been targeted to children and adolescents with specific health problems [2, 3]. The psychological development of children with short stature and associated impairments in HrQoL, as well as the impact of Growth Hormone (GH) treatment has received substantial, however not unequivocal, attention in the literature [4].

Using generic measures, several studies have not found differences in the HrQoL of children and adolescents with or without GH treatment, while others report more psychological impairment in GH treatment naïve patients with an achieved height below -2 standard deviation (SD) as compared to GH treated patients within the normal range for age and gender group [5].

These results suggest that generic measures could be less sensitive for detecting subtle but clinically relevant changes in HrQoL and highlight the importance of including specific assessments in short stature. However, few instruments have been developed for conditions related to physical impairments, disabilities or handicaps, such as those associated with short stature. A literature review identified five condition-specific HrQoL measures in adolescents and children with short stature. Most of them, however, have been developed in one country or language context only, have not been translated to other cultural contexts, have not undergone linguistic validation and have not been subjected to cross-cultural psychometric testing [6]. This limits the ability to conduct rigorous cross-cultural research to evaluate outcomes from new treatment paradigms and precludes generalization of observed treatment effects beyond national borders.

The Quality of Life in Short Stature Youth (QoLISSY) project aimed at developing an instrument for short statured children and adolescents with growth hormone deficiency (GHD) or idiopathic short stature (ISS) and their parents [4, 7], responding to the necessity to cross-culturally develop an appropriate measure rather than relying on the translation of an instrument developed in just one country or language. The QoLISSY project respected the simultaneous approach to making HrQoL instruments available cross-culturally which aims at a common conceptual development of an instrument across several cultures and languages simultaneously [8]. The international guidelines for the development of quality of life measures have been proposed which describe three steps to be taken to arrive at a cross-culturally valid instrument [8, 9]:

- Conducting focus groups with children/adolescents and parents in order to identify relevant domains and items and a conceptual model,
- Pilot testing with cognitive debriefing of the preliminary items of the measure and
- Field testing with a re-test of the instrument in children and one of their parents, in order to examine psychometric properties, namely validity and reliability.

The aim of the QoLISSY instrument development was to construct a psychometrically sound and cross culturally valid tool that covers the impacts of short stature on HrQoL in children and adolescents from their own perspective and the added perspective of their parents [10, 11]. Development and psychometric testing of the QoLISSY instrument in a large sample of children and adolescents from 5 European countries (France, Germany, Spain, Sweden and the UK) were described in a previous paper [12]. The current paper specifically intends to employ a truly cross-cultural approach to instrument development and testing by iteratively omitting data from one country from the pooled data set and comparing the resulting psychometric characteristics across the data sets. This "take-one-country-out (TOCO) technique" aims at testing cross cultural stability in instrument psychometric performance.

2. Methods

Participants and Procedures

The current analyses use data from the field test phase of the QoLISSY project. The sample consisted of short statured patients in two age groups (children aged 8-12, and adolescents aged 13-18 years) and their parents, recruited from the collaborating growth clinics in France, Germany, Spain, Sweden and the UK. Recruitment focused on including subjects by diagnosis (GHD and ISS), treatment status (GH treated and untreated), age (children and adolescents) and gender (boys and girls). Upon ethics committee approvals, informed consent from all parents and informal assent from children and adolescents were obtained, together with the permission to extract medical data from patients charts through their clinicians at the study centres. Questionnaires were mailed to parents and children, in which, in addition to HrQoL assessments, socio-demographic, clinical and psychosocial information was collected. Distribution by mail was conducted through the growth centres to patients and parents separately with a prepaid return envelope. Completed documents, including clinical data, were anonymised, quality-checked and processed at the German study centre.

The QoLISSY Questionnaire

The QoLISSY questionnaire is a condition-specific instrument to assess the patient-reported HrQoL of children and adolescents with short stature, aged 8 to 18 years as well as the parent-reported HrQoL of children aged 4 to 18 years. The questionnaire was simultaneously developed in five European countries using focus-groups with item generation, a pilot test with cognitive debriefing and a field test with re-test.

The patient- and parent-reported version of the QoLISSY questionnaire as applied in the field test phase, consists of a core module with 22 items, assessing three HrQoL domains: physical (6 items, e.g. “Because of my height I depend on others.”), social (8 items, e.g. “Because of my height I get laughed at or teased.”) and emotional (8 items, e.g. “I am happy with my height.”). Additionally three scales assessing determinants of HrQoL: coping (10 items, e.g. “I tell myself it is ok to be short.”), height-related beliefs (4 items, e.g. “I believe that tall people are better liked.”) and treatment experiences (14 items, e.g. “I feel good because of my treatment.”) are included. The parent-reported version included two additional scales assessing concerns about the child’s future (5 items, e.g. “My child worries about whether (s)he will grow enough.”) and the effects of their child’s short stature on the family (11 items, e.g. “My child’s growth problems make me feel anxious.”). All items are to be answered with a 5-point Likert scale, whose interval properties had been examined in earlier research. Response options ranged from 1 (*not at all/ never*) to 5 (*extremely/ always*) considered the “past week” as time reference. Standardized (0-100) scores were computed for each scale separately, and summed up to a total score of HrQoL representing the three scales of the core domain [12]. The QoLISSY instrument is freely available for academic use. Access to the instrument is provided upon request through a published manual, webpages and individual orders (see information at end of manuscript).

Data analyses

Statistical analyses of field test data were performed with the Statistical Package for the Social Sciences 18.0 (SPSS, v.18; Chicago, IL, USA) and with the Analysis of Moments Structures (AMOS, v. 18). Descriptive statistics were calculated for socio-demographic and clinical variables. Scale scores were calculated if less than 25% of responses were missing. The homogeneity of the sample’s characteristics across countries was tested through chi-square test and analysis of variance (ANOVA) with post-hoc tests, depending on whether the variable was categorical or continuous.

Second-order confirmatory factor analyses (CFA) were performed to examine the factorial structure of the patient-reported and parent-reported version of the QoLISSY questionnaire. The second-order

CFA model comprised 6 observed indicators (i.e., 6 items) loading on the physical domain, 8 observed indicators loading on the social domain and 8 observed indicators loading on the emotional domain; in turn, the three domains loaded on a HrQoL total score. The models' goodness of fit was assessed using the main fit indexes: maximum-likelihood χ^2 p-value and $\chi^2/\text{degrees of freedom}$, comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root mean squared residual (SRMR). The reference values of $\chi^2/\text{df} \leq 2$, $\text{CFI} \geq .95$, $\text{RMSEA} \leq .06$ and $\text{SRMR} \leq .08$ were considered indicators of model's good fit to the data; and the reference values of $\chi^2/\text{df} \leq 5$, $\text{CFI} \geq .90$ and $\text{RMSEA} \leq .10$ were considered the threshold for assessing the model's fit as acceptable [15-17].

Given the small sample size in each country separately, the measurement invariance of the factorial model across the five countries was tested using multi-group analyses with the Take one country out (TOCO approach). This technique consists of comparing the psychometric results and the socio-demographic, clinical and psychosocial determinants of HrQoL between the total sample from the 5 participating countries and subsamples. From the total sample of five countries, one was removed at a time and results were compared. The Chi-square difference method ($\Delta\chi^2$) was used to compare the unconstrained model with a model in which factor loadings (measurement invariance) and structural weights (structural invariance) were constrained to be equal across groups. [13].

Once a factorial structure had been identified for both patient- and parent-reported version of the QoLISSY questionnaires, its psychometric properties were further analysed for each country independently and for the subsamples removing one country at a time. Reliability of the QoLISSY subscales and total score was assessed by using the Cronbach's alpha value and the composite reliability (CR) value, calculated from the squared sum of standardized factor loading divided by the squared sum of standardized factor loading plus the sum of the error variance terms. Good construct reliability was established if CR was higher than .70 [14].

Agreement between patients' and parents' reports across different countries was examined using Intraclass Correlation Coefficients [ICC] (two-way mixed model, absolute agreement, 95% confidence interval). The levels of agreement were calculated for each country independently and for the subsamples, again taking-one-country-out. By transforming the correlation coefficients in Z-Fisher scores [15], the levels of agreement of each subsample were compared with the total sample.

To identify the determinants of HrQoL total score, hierarchical multiple regression analyses were performed with two different approaches for the cross-cultural analyses. First, a regression analyses for the total sample was conducted, including the "country" as a predictor, after this categorical variable had been recoded into dummy-variables. Second, the regression analyses were conducted for the TOCO subsamples. To minimize the effect of sample dissimilarities across countries, socio-

demographic and clinical characteristics were controlled by including them in the first block of the regression equation. Children's and parents' HrQoL determinants (coping, beliefs and treatment experiences) were entered in the second block. A third block was added only for the parent report including the two supplementary determinants for parents (future of the child and effects on parents).

3. Results

Sample Characteristics

A total of 268 children and adolescents with short stature, aged between 8 and 18 years and one of their parents were included in the QoLISSY field test used for the current analyses. An overview of socio-demographic and clinical characteristics of the samples per country is displayed in Table 1.

Table 1. Country-specific socio-demographic and clinical characteristics of the sample

Children and adolescents		Total <i>N</i> = 268	France <i>n</i> = 53	Germany <i>n</i> = 67	Spain <i>n</i> = 44	Sweden <i>n</i> = 73	the UK <i>n</i> = 31	Differences across countries
Gender	Male	155	29	43	24	42	17	$\chi^2 = 1.63$
	Female	113	24	24	20	31	14	
Age	8-12	127	30	41	27	14	15	$\chi^2 = 33.69^{***}$
	13-18	141	23	26	17	59	16	
Diagnosis	GHD	111	17	26	20	33	15	$\chi^2 = 3.44$
	ISS	157	36	41	24	40	16	
Treatment status	Treated	145	21	49	14	39	22	$\chi^2 = 33.69^{***}$
	Untreated	123	32	18	30	34	9	
Current height deviation	Above -2 <i>SD</i>	130	3	37	24	48	18	$\chi^2 = 33.69^{***}$
	Below -2 <i>SD</i>	132	50	29	16	25	12	
	Missing	6	0	1	4	0	1	
Parents		Total <i>N</i> = 268	France <i>n</i> = 53	Germany <i>n</i> = 67	Spain <i>n</i> = 44	Sweden <i>n</i> = 73	the UK <i>n</i> = 31	Differences across countries
Kinship	Mother	217	43	61	30	57	26	$\chi^2 = 35.96^{***}$
	Father	33	4	5	7	13	4	
	Both parents together	6	0	0	6	0	0	
	Missing	12	6	1	1	3	1	
Age	<i>M</i> (<i>SD</i>)	43.58 (5.73)	42.04 (6.15)	42.61 (5.49)	^a	44.93 (5.38)	44.90 (5.67)	$F = 3.79^*$

^a There is no available information on age or birthdate for Spanish parents.

*** $p < .001$, two-tailed. ** $p < .01$, two-tailed. * $p < .05$, two-tailed.

The frequency distribution of gender and diagnosis was homogeneous across countries for the patient sample. Sweden, however, had recruited a larger number of adolescents as compared to other countries. The distribution of treatment status and height deviation also differed across countries, with higher percentages of untreated children in France and Spain, and higher percentage of children with height deviation above -2 SD in France. Significant differences in the parents' socio-demographic characteristics across countries were found: the Spanish sample included more fathers or both parents; regarding age, Swedish parents were older than parents from other countries.

Confirmatory Factor Analysis and Reliability of the QoLISSY Patient-reported Version

The second-order CFA model of the patient-reported version of the QoLISSY questionnaire had an acceptable fit, with $\chi^2(206, n = 263) = 615.35, p < .05, \chi^2/df = 2.99, CFI = .88, RMSEA = .087$ and $SRMR = .054$. All items showed factorial validity, with significant factor loadings ($p < .001$) and, except for two items, standardized regression weights were above the threshold of .50. Using the TOCO approach, the factorial model was re-run for the five subsamples, sequentially removing the data from one country at a time. As presented in Table 2, the second-order CFA model had an acceptable fit for the five TOCO subsamples. The results from multi-group analyses showed no significant χ^2 differences in factor loadings and structural weights between the total sample and the TOCO subsamples in which one country was removed, confirming the measurement and structural invariance of the structural model across countries.

Table 2. Cross-cultural comparison of the factorial structure of the patient-reported version of QoLISSY questionnaire

	<i>n</i>	Model's goodness of fit					Model comparison			
		$\chi^2(206), p$	χ^2/df	CFI	RMSEA	SRMR	$\Delta\chi^2$	Δdf	<i>p</i>	
Taking-out France	213	581.72, $p < .05$	2.82	.87	.093	.058	Measurement invariance	8.07	19	.986
							Structural invariance	.23	2	.891
Taking-out Germany	197	567.77, $p < .05$	2.76	.86	.095	.058	Measurement invariance	10.30	19	.945
							Structural invariance	.24	2	.971
Taking-out Spain	219	525.76, $p < .05$	2.55	.89	.084	.053	Measurement invariance	.99	19	1.00
							Structural invariance	.16	2	.922
Taking-out Sweden	191	493.96, $p < .05$	2.40	.86	.086	.062	Measurement invariance	3.96	19	1.00
							Structural invariance	.14	2	.932
Taking-out the UK	232	605.83, $p < .05$	2.94	.87	.092	.057	Measurement invariance	.94	19	1.00
							Structural invariance	.03	2	.987

Inspection of the reliability of the three core domains (physical, social and emotional) and of the total score showed acceptable Cronbach's α values [16, 17] for each country independently (ranging from .72 to .96) and for the total sample removing one country at a time (ranging from .83 to .95) (see Table 3). Additionally, the composite reliability values for the TOCO subsamples were higher than the threshold of .70 [14], indicating good reliability for the subscales and for the HrQoL total score.

Confirmatory Factor Analysis and Reliability of the QoLISSY Parent-reported Version

The factorial model for the parent-reported version of the QoLISSY questionnaire was built in strict equivalence to the patient-reported version. The model had an acceptable fit to the total sample, with $\chi^2(206, n = 259) = 718.81, p < .05, \chi^2/df = 3.49, CFI = .87, RMSEA = .098$ and $SRMR = .052$, and with all factor loadings being statistically significant. The comparison across countries with the TOCO approach was replicated for parents' reports (see Table 4). The second-order CFA model had an acceptable fit for the TOCO subsample in which Spain data were removed, and a nearly acceptable fit for the remaining TOCO subsamples. The absence of significant χ^2 differences in the factor loadings and structural weights between the total sample and the TOCO subsamples indicates that the factorial structure of the parent-reported version of the QoLISSY questionnaire was invariant across countries.

High internal consistency values were observed for all three core domains and HrQoL total score for each country independently (Cronbach's alpha values ranging from .78 to .96). Moreover, very good Cronbach's alphas (ranging from .83 to .96) and composite reliability values (ranging from .85 and .97) were found for the five TOCO subsamples where one country was sequentially removed (see Table 5).

Table 3. Reliability (*Cronbach's* alphas and composite reliability) of the three domains and total QoL score of the patient-reported version of QoLISSY questionnaire across countries

		Physical domain (6 items)		Social domain (8 items)		Emotional domain (8 items)		HrQoL total score (22 items/ 3 domains)	
		<i>Cronbach's</i> alpha	Composite reliability	<i>Cronbach's</i> alpha	Composite reliability	<i>Cronbach's</i> alpha	Composite reliability	<i>Cronbach's</i> alpha	Composite reliability
France	Independently	.80	-	.83	-	.85	-	.92	-
	Take-out	.85	.86	.87	.88	.88	.88	.95	.97
Germany	Independently	.87	-	.85	-	.89	-	.95	-
	Take-out	.84	.85	.88	.88	.88	.88	.95	.96
Spain	Independently	.72	-	.81	-	.83	-	.91	-
	Take-out	.86	.87	.88	.89	.89	.89	.95	.96
Sweden	Independently	.88	-	.92	-	.91	-	.96	-
	Take-out	.83	.84	.84	.85	.86	.87	.94	.95
the UK	Independently	.87	-	.89	-	.85	-	.95	-
	Take-out	.84	.86	.87	.87	.88	.89	.94	.96
Total sample		.85	.86	.87	.88	.88	.88	.95	.96

Table 4. Cross-cultural comparison of the factorial structure of the parent-reported version of QoLISSY questionnaire

	<i>n</i>	Model's goodness of fit					Model comparison			
		$\chi^2(206), p$	χ^2/df	CFI	RMSEA	SRMR	$\Delta\chi^2$	Δdf	<i>p</i>	
Taking-out France	212	700.22, <i>p</i> < .05	3.40	.85	.107	.055	Measurement invariance	9.07	19	.972
							Structural invariance	.02	2	.991
Taking-out Germany	194	649.97, <i>p</i> < .05	3.16	.86	.106	.057	Measurement invariance	10.34	19	.944
							Structural invariance	.63	2	.729
Taking-out Spain	216	628.24, <i>p</i> < .05	3.05	.87	.098	.052	Measurement invariance	1.90	19	1.00
							Structural invariance	.09	2	.956
Taking-out Sweden	186	600.74, <i>p</i> < .05	2.92	.84	.102	.062	Measurement invariance	7.52	19	.991
							Structural invariance	.31	2	.856
Taking-out the UK	228	698.02, <i>p</i> < .05	3.39	.86	.103	.053	Measurement invariance	2.69	19	1.00
							Structural invariance	.15	2	.927

Agreement between Patients'- and Parents'-Reports

Moderate to strong levels of agreement between children/ adolescents' and parents' reports of the HrQoL domains and total score were found for all countries independently (ICCs ranging from .58 to .80) and for the TOCO subsamples (ICCs ranging from .66 to .78). These high levels of agreement across countries supported the cross-cultural concurrent validity of the QoLISSY instruments (see Table 6).

Determinants of HrQoL in Short Stature

The QoLISSY additional modules (coping, beliefs and treatment as patient-reported HrQoL determinants; coping, beliefs, treatment, future and effect on parents as parent-reported HrQoL determinants) presented good reliabilities, with *Cronbach's alpha* values ranging from .77 (coping) to .91 (treatment) for the children/ adolescents' reports, and from .72 (coping) to .96 (beliefs) for the parents reports (see Table 7).

Table 5. Reliability (*Cronbach's* alphas and composite reliability) of the three domains and total score of the parent-reported version of QoLISSY questionnaire across countries

		Physical domain (6 items)		Social domain (8 items)		Emotional domain (8 items)		HrQoL total score (22 items/ 3 domains)	
		<i>Cronbach's</i> alpha	Composite reliability	<i>Cronbach's</i> alpha	Composite reliability	<i>Cronbach's</i> alpha	Composite reliability	<i>Cronbach's</i> alpha	Composite reliability
France	Independently	.82	-	.90	-	.86	-	.94	-
	Take-out	.86	.88	.91	.91	.88	.88	.95	.97
Germany	Independently	.86	-	.89	-	.89	-	.95	-
	Take-out	.86	.88	.92	.93	.89	.89	.96	.97
Spain	Independently	.78	-	.91	-	.92	-	.95	-
	Take-out	.88	.89	.91	.92	.88	.88	.96	.97
Sweden	Independently	.91	-	.92	-	.86	-	.96	-
	Take-out	.83	.85	.90	.90	.88	.88	.95	.95
the UK	Independently	.85	-	.90	-	.82	-	.94	-
	Take-out	.87	.88	.91	.92	.89	.90	.96	.97
Total sample		.86	.88	.91	.91	.88	.89	.95	.97

Table 6. Agreement between patients'- and parents'-report (Intraclass Correlation Coefficients) for the three subscales and total score of the QoLISSY questionnaire

		Physical	Z- score ^a	Social	Z- score ^a	Emotional	Z- score ^a	Total HrQoL	Z- score ^a
France	Independently	.58***	-1.46	.62***	-1.61	.67***	-.22	.67***	-1.39
	Take-out	.73***	.23	.77***	.26	.66***	-.59	.78***	.00
Germany	Independently	.75***	.46	.77***	.17	.63***	-.75	.76***	-.34
	Take-out	.72***	.00	.76***	.00	.71***	.41	.78***	.00
Spain	Independently	.63***	-.98	.70***	-.76	.69***	.00	.76***	-.29
	Take-out	.73***	.23	.77***	.26	.69***	.00	.78***	.00
Sweden	Independently	.70***	-.30	.75***	-.17	.66***	-.41	.75***	-.53
	Take-out	.71***	-.21	.73***	-.69	.67***	-.38	.76***	-.50
the UK	Independently	.78***	.69	.80***	.51	.63***	-.54	.79***	.13
	Take-out	.71***	-.22	.76***	.00	.69***	.00	.78***	.00
Total sample		.72***		.76***		.69***		.78***	

^a Test of the difference between the ICCs of the subsamples of each country independently/ removed and the ICC of the total sample (Preacher, 2002). * $p < .05$; ** $p < .01$; *** $p < .001$, two-tailed.

To identify the determinants of HrQoL, we conducted hierarchical regression analyses for the total sample, entering socio-demographic (age group, gender and country), and clinical (diagnosis, treatment status, current height deviation) characteristics in the first block, and psychosocial determinants as assessed by the QoLISSY additional subscales in the second (coping, beliefs and treatment experiences) and third (future and effects on parents; only for parent-reported data) blocks of the regression equation. The regression model explained 59% of the patient-reported HrQoL variance and 72% of the parent-reported HrQoL variance (see Table 8).

Socio-demographic and clinical characteristics explained 28% of the patient -reported and 32% of the parent-reported HrQoL variance. Here height deviation was the variable with the greatest individual contribution, indicating a negative association between current short stature and HrQoL, both patient- and parent-reported. No significant effects of country were found for patient-reported HrQoL. For parent-reported HrQoL, having Swedish nationality was associated with better HrQoL. Controlling for socio-demographic and clinical characteristics, the HrQoL determinants explained 32% of the patient-reported and 25% of the parent-reported HrQoL. Specifically, adaptive general beliefs about height were positively associated with better patient- and parent-reported HrQoL. The two extra HrQoL determinants included in the parents' version additionally explained 16% of the parent-reported HrQoL variance: fewer worries about the child's future and reduced impact of the growth problems on parents' feelings were significantly linked to higher parental ratings of child's HrQoL.

Table 7. Reliability (*Cronbach's* alphas) of the QoLISSY subscales assessing the determinants of patient- and parent-reported HrQoL across countries

Patients' reports		Coping (10 items)	Beliefs (4 items)	Treatment (14 items)		
France	Independently	.77	.85	.87		
	Take-out	.83	.86	.86		
Germany	Independently	.79	.84	.88		
	Take-out	.82	.86	.86		
Spain	Independently	.79	.86	.85		
	Take-out	.82	.86	.87		
Sweden	Independently	.87	.89	.83		
	Take-out	.79	.84	.88		
the UK	Independently	.83	.84	.91		
	Take-out	.81	.86	.86		
Total sample		.82	.86	.87		
Parents' reports		Coping (10 items)	Beliefs (4 items)	Treatment (14 items)	Future (5 items)	Effects on parents (11 items)
France	Independently	.73	.87	.77	.91	.89
	Take-out	.84	.91	.89	.91	.91
Germany	Independently	.80	.86	.86	.87	.90
	Take-out	.84	.91	.90	.92	.91
Spain	Independently	.72	.92	.87	.92	.86
	Take-out	.84	.90	.89	.90	.92
Sweden	Independently	.92	.96	.85	.91	.94
	Take-out	.77	.87	.88	.90	.89
the UK	Independently	.83	.89	.92	.93	.90
	Take-out	.83	.90	.87	.90	.91
Total sample		.83	.90	.89	.91	.91

Regressions analyses rerun for the TOCO subsamples showed consistent results across countries. For the patient-reported HrQoL, the portion of variance which was explained by socio-demographic and clinical characteristics together with HrQoL determinants ranged from 52% to 61% across countries (see Table 9). Similar to the results found for the total sample, height deviation and beliefs presented significant individual contributions for all TOCO subsamples: current height deviation above -2 SD (achieved normal height) and more adaptive beliefs about stature were associated with better patient-reported HrQoL.

Table 8. Regression analyses of the HrQoL determinants on patients-reports and on parents'-reports of HrQoL total score for the total sample

	Patient-reported HrQoL $R^2 = .59; F_{(12, 130)} = 15.86^{***}$		Parent-reported HrQoL $R^2 = .72; F_{(14, 111)} = 20.74^{***}$	
	β	t	β	t
1st step: socio-demographic and clinical variables	$\Delta R^2 = .28; F_{(9, 133)} = 5.63^{***}$		$\Delta R^2 = .32; F_{(9, 116)} = 5.94^{***}$	
Age group ^a	.06	.79	.06	.77
Gender ^b	.10	1.29	.06	.71
Diagnosis ^c	-.05	-.56	-.14	-1.47
Treatment status ^d	-.19	-1.93	-.08	-.72
Height deviation ^e	-.29	-3.27**	-.31	-3.31**
Country ^f				
Germany vs. France	-.05	-.52	-.04	-.36
Germany vs. Spain	.04	.46	.07	.77
Germany vs. Sweden	.15	1.56	.25	2.37*
Germany vs. the UK	-.04	-.45	-.06	-.65
2nd step: HrQoL determinants	$\Delta R^2 = .32; F_{(3, 130)} = 33.99^{***}$		$\Delta R^2 = .25; F_{(3, 113)} = 21.65^{***}$	
	β	t	β	t
Age group ^a	.02	.35	.08	1.21
Gender ^b	.15	2.55*	.09	1.45
Diagnosis ^c	.05	.72	-.10	-1.25
Treatment status ^d	-.21	-2.77**	-.10	-1.13
Height deviation ^e	-.22	-3.15**	-.16	-2.05*
Country ^f				
Germany vs. France	-.03	-.45	-.04	-.53
Germany vs. Spain	.04	.52	.15	1.86
Germany vs. Sweden	.07	.86	.12	1.37
Germany vs. the UK	-.07	-1.02	-.06	-.74
Coping	-.03	-.44	-.01	-.15
Beliefs	.61	10.07***	.53	7.90***
Treatment experiences	.10	1.55	.12	1.65
3rd step: Additional HrQoL determinants for parents			$\Delta R^2 = .16; F_{(2, 111)} = 31.77^{***}$	
			β	t
Age group ^a			.02	.41
Gender ^b			.09	1.69
Diagnosis ^c			-.03	-.43
Treatment status ^d			-.04	-.55
Height deviation ^e			-.12	-1.82
Country ^f				
Germany vs. France			-.06	-.89
Germany vs. Spain			.11	1.68
Germany vs. Sweden			.05	.68
Germany vs. the UK			-.11	-1.72
Coping			.01	.20
Beliefs			.16	2.24*
Treatment experiences			.07	1.20
Future			.37	4.55***
Effects on parents			.29	3.87***

^a Reference group: o = children 8-12 years-old; ^b Reference group: o = girl; ^c Reference group: o = Growth Hormone Deficiency; ^d Reference group: o = treated; ^e Reference group: o = height deviation above -2 SD (achieved normal stature); ^f The variable "country" was recoded into four dummy-variables; Reference group: o = Germany.

* $p < .05$; ** $p < .01$; *** $p < .001$, two-tailed.

Table 9. Regression analyses of the HrQoL determinants on patients-reports of HrQoL total score taking-one-country-out

	Total health-related quality of life (patient-reported)									
	Take-France-out		Take-Germany-out		Take-Spain-out		Take-Sweden-out		Take-the UK-out	
	$R^2 = .57; F_{(8, 112)} = 18.78^{***}$		$R^2 = .61; F_{(8, 86)} = 16.10^{***}$		$R^2 = .61; F_{(8, 118)} = 23.04^{***}$		$R^2 = .52; F_{(8, 98)} = 13.28^{***}$		$R^2 = .60; F_{(8, 113)} = 21.07^{***}$	
1st step: socio-demographic and clinical variables	$\Delta R^2 = .26; F_{(5, 115)} = 8.01^{***}$		$\Delta R^2 = .26; F_{(5, 101)} = 5.94^{***}$		$\Delta R^2 = .30; F_{(5, 121)} = 10.59^{***}$		$\Delta R^2 = .17; F_{(5, 101)} = 4.06^{**}$		$\Delta R^2 = .25; F_{(5, 116)} = 7.71^{**}$	
	β	t	β	t	β	t	β	t	β	t
Age group ^a	.13	1.53	.08	.81	.12	1.55	.09	1.00	.13	1.54
Gender ^b	.11	1.40	.16	1.68	.08	1.09	.05	.57	.08	.92
Diagnosis ^c	-.05	-.48	.00	.01	-.08	-.84	-.05	-.44	-.08	-.83
Treatment status ^d	-.20	-2.01*	-.22	-1.79	-.10	-.96	-.13	-1.25	-.16	-1.58
Height deviation ^e	-.34	-3.93 ^{***}	-.33	-3.34 ^{**}	-.42	-5.00 ^{***}	-.30	-3.15 ^{**}	-.33	-3.88 ^{***}
2nd step: HrQoL determinants	$\Delta R^2 = .32; F_{(3, 112)} = 27.51^{***}$		$\Delta R^2 = .35; F_{(3, 83)} = 24.81^{***}$		$\Delta R^2 = .31; F_{(3, 118)} = 30.77^{***}$		$\Delta R^2 = .35; F_{(3, 98)} = 24.02^{***}$		$\Delta R^2 = .35; F_{(3, 113)} = 32.79^{***}$	
	β	t	β	t	β	t	β	t	β	t
Age group ^a	.07	1.10	.02	.23	.04	.69	.05	.64	.03	.49
Gender ^b	.17	2.66 ^{**}	.19	2.71 ^{**}	.11	1.84	.12	1.71	.14	2.22*
Diagnosis ^c	.06	.74	.09	.98	.00	.02	.04	.51	.04	.48
Treatment status ^d	-.21	-2.78 ^{**}	-.19	-2.11*	-.11	-1.37	-.19	-2.29*	-.20	-2.72 ^{**}
Height deviation ^e	-.25	-3.70 ^{***}	-.22	-2.78 ^{**}	-.27	-3.95 ^{***}	-.24	-3.21 ^{**}	-.25	-3.82 ^{***}
Coping	-.05	-.67	-.02	-.23	-.02	-.31	.01	.12	-.05	-.72
Beliefs	.60	9.07 ^{***}	.62	8.58 ^{***}	.60	9.43 ^{***}	.62	8.32 ^{***}	.63	9.91 ^{***}
Treatment experiences	.12	1.69	.09	1.11	.16	2.46*	.15	1.87	.09	1.44

^a Reference group: o = children 8-12 years-old; ^b Reference group: o = girl; ^c Reference group: o = Growth Hormone Deficiency; ^d Reference group: o = treated; ^e Reference group: o = height deviation above -2 SD (achieved normal stature).

* $p < .05$; ** $p < .01$; *** $p < .001$, two-tailed.

Table 10. Regression analyses of the HrQoL determinants on parents' reports of HrQoL total score taking-one-country-out

	Total health-related quality of life (parent-reported)									
	Take-France-out		Take-Germany-out		Take-Spain-out		Take-Sweden-out		Take-the UK-out	
	$R^2 = .71; F_{(10, 104)} = 16.83^{***}$	$R^2 = .68; F_{(10, 68)} = 14.71^{***}$	$R^2 = .75; F_{(10, 110)} = 30.71^{***}$	$R^2 = .59; F_{(10, 87)} = 14.88^{***}$	$R^2 = .73; F_{(10, 99)} = 26.56^{***}$	$\Delta R^2 = .27; F_{(5, 107)} = 7.89^{***}$	$\Delta R^2 = .25; F_{(5, 73)} = 4.78^{**}$	$\Delta R^2 = .31; F_{(5, 110)} = 9.93^{***}$	$\Delta R^2 = .17; F_{(5, 92)} = 3.83^{**}$	$\Delta R^2 = .26; F_{(5, 104)} = 7.42^{**}$
1st step: socio-demographic and clinical variables	β	t	β	t	β	t	β	t	β	t
Age group ^a	.19	2.26*	.11	1.05	.13	1.63	.07	.66	.18	2.14*
Gender ^b	.07	.79	.06	.62	.06	.71	.06	.64	.03	.39
Diagnosis ^c	-.13	-1.23	-.09	-.68	-.20	-1.99*	-.16	-1.45	-.17	-1.67
Treatment status ^d	-.09	-.85	-.08	-.63	.09	.89	.01	.06	-.02	-.23
Height deviation ^e	-.36	-4.09***	-.39	-3.58**	-.46	-5.19***	-.34	-3.33**	-.36	-3.99***
2nd step: HrQoL determinants	$\Delta R^2 = .30; F_{(3, 104)} = 23.45^{***}$		$\Delta R^2 = .26; F_{(3, 70)} = 12.49^{***}$		$\Delta R^2 = .28; F_{(3, 107)} = 24.85^{***}$		$\Delta R^2 = .29; F_{(3, 89)} = 15.63^{***}$		$\Delta R^2 = .29; F_{(3, 101)} = 21.87^{***}$	
	β	t	β	t	β	t	β	t	β	t
Age group ^a	.14	2.09*	.06	.70	.09	1.39	.08	1.03	.14	1.98*
Gender ^b	.09	1.31	.08	.90	.11	1.75	.10	1.22	.08	1.26
Diagnosis ^c	-.12	-1.52	-.09	-.85	-.15	-1.91	-.15	-1.61	-.13	-1.53
Treatment status ^d	-.05	-.62	-.06	-.54	.00	.02	-.00	-.04	-.03	-.38
Height deviation ^e	-.19	-2.59*	-.16	-1.61	-.23	-3.09**	-.26	-3.04**	-.20	-2.67**
Coping	-.04	-.54	-.02	-.18	-.03	-.52	.04	.50	.00	.01
Beliefs	.51	7.54***	.50	5.59***	.54	8.17***	.52	6.49***	.54	7.77***
Treatment experiences	.24	3.59**	.23	2.51*	.19	2.93**	.17	2.05*	.13	1.85
3rd step: Additional HrQoL determinants for parents	$\Delta R^2 = .26; F_{(2, 102)} = 26.24^{***}$		$\Delta R^2 = .35; F_{(2, 68)} = 18.79^{***}$		$\Delta R^2 = .15; F_{(2, 105)} = 31.16^{***}$		$\Delta R^2 = .17; F_{(2, 87)} = 20.39^{***}$		$\Delta R^2 = .18; F_{(2, 99)} = 31.97^{***}$	
	β	t	β	t	β	t	β	t	β	t
Age group ^a	.04	.72	-.01	-.15	.02	.44	.02	.21	.03	.56
Gender ^b	.10	1.79	.05	.71	.10	2.03*	.12	1.79	.10	1.78
Diagnosis ^c	-.01	-.19	-.03	-.34	-.08	-1.24	-.07	-.94	-.03	-.51
Treatment status ^d	-.01	-.13	-.00	-.04	.02	.37	.03	.38	-.00	-.05
Height deviation ^e	-.13	-2.11*	-.12	-1.51	-.15	-2.47*	-.19	-2.61*	-.15	-2.50*
Coping	.00	.06	.01	.09	-.02	-.41	.03	.43	.03	.54
Beliefs	.15	1.96*	.11	1.05	.16	2.25*	.16	1.82	.15	-2.00*
Treatment experiences	.19	3.43**	.15	2.03*	.16	2.96**	.16	2.29*	.10	1.80
Future	.38	4.29***	.32	2.83**	.38	4.50***	.36	3.61**	.39	4.43***
Effects on parents	.27	3.51**	.36	3.58**	.26	3.25**	.29	3.36**	.30	3.66***

^a Reference group: o = children 8-12 years-old; ^b Reference group: o = girl; ^c Reference group: o = Growth Hormone Deficiency; ^d Reference group: o = treated; ^e Reference group: o = height deviation above -2 SD (achieved normal stature). * $p < .05$; ** $p < .01$; *** $p < .001$, two-tailed.

For the parents' reports of their children's HrQoL, clinical and socio-demographic variables and the QoLISSY determinants explained between 59% and 75% of the variance across the TOCO subsamples (see Table 10). Significant individual contributions of the current height deviation, beliefs about stature, worries about the future and effects on parents were consistently found for the total sample and for all TOCO subsamples. In addition, treatment experiences also had a significant effect on parent-reported HrQoL in four of the five TOCO subsamples. Specifically, and except when the UK was taken out, more positive treatment experiences were associated with better parent-reported HrQoL. Fewer worries about the child's future and reduced impact of the growth problems on parents were also associated with better parent-reported HrQoL, explaining an additional portion of HrQoL variance ranging from 15% and 35% across TOCO subsamples.

4. Discussion

Using a simultaneous strategy in instrument development across countries makes it possible to incorporate cross-cultural input from the outset of the process and provides insight into similarities and differences between countries with respect to perceptions and evaluations of the HrQoL concept.

The QoLISSY project simultaneously developed a condition specific questionnaire for children and adolescents with short stature and their parents. Researchers and clinicians from five countries cooperated to create a quality of life measure through a systematic qualitative and quantitative process involving focus groups (for item generation and formation of the conceptual model), a pilot test (for preliminary analysis and cognitive debriefing of items and scales) and a field test with retest (with final psychometric analysis of instrument performance across all countries) yielding rich information about the assessment of HrQoL of children. The present paper examined the last of these steps with regard to the cross-cultural psychometric performance of the QoLISSY quality of life core module and its determinant modules from field test data. The process respected lessons learned from one of the first international HrQoL instrument development exercises: the World Health Organization (WHO) quality of life measure: the WHO-QoL [18].

One major issue in the analysis of a cross-cultural data set is the choice of the strategy used to derive psychometric information [19]. Analyses conducted on the international level involving all countries cannot be easily broken down to the individual countries both for reasons of sample size as well as sample composition. Between 5 and 10 respondents per items would be necessary to conduct principal component factor analysis on a country level. Especially in rare health conditions, such as GHD, and limited patient samples, such as health-referred ISS patients, it is difficult to recruit a sufficient number of participants into a study. Therefore the use of sophisticated psychometric strategies such as confirmatory factor analyses with structural equation modelling, requiring larger sample sizes, is

usually not possible on the single country level.

One alternative strategy is to use a combined international total data set, then remove subsequently and iteratively remove one country at a time and examine the correspondence between the resulting models. This “Take-One-Country-Out – TOCO- strategy was chosen for testing the QoLISSY questionnaire, after having examined and taken into account differences in the sample composition across countries.

Confirmatory factor analysis of the patient-reported combined QoLISSY data set showed that the factorial model including physical, social and emotional quality of life with a global score had an acceptable fit. Within the TOCO approach, CFA results, as indicators of construct validity, remained stable after iteratively omitting one country. Fit indices were comparable across countries and comparisons of each data set with the total data set did not yield statistically significant differences, suggesting cross-cultural comparability in terms of factorial validity.

Reliability was inspected for each country individually as well as following the iterative TOCO approach. Results showed good overall reliability with high convergence of coefficients between individual countries as well as good comparability between the TOCO and total data sets.

Similarly, the parent-reported version, subjected to CFA and reliability testing, yielded acceptable overall fit indices for the factorial model. Comparisons of each iterative TOCO data set with the total data set did not yield substantial differences.

Convergence of children’s/ adolescents’ and parents’ reports was supported using both individual countries and TOCO subsample ICCs. The moderate to high levels of parent-child agreement in assessing paediatric HrQoL found in the present study were in line with previous research reporting higher levels of agreement between parents and children with chronic health conditions than between parents and healthy children [20]. Moreover, Upton, et al. [21] advocated that the levels of agreement would be dependent on the relevance of different domains for a specific clinical group. Since parent and children completed their field test questionnaires at home, a parental influence on children’s answers’ cannot be ruled out. As the QoLISSY questionnaire is a disease-specific measure, the high levels of agreement may be explained by the pertinence of its items for the particular group of children with short stature and by the higher likelihood of parents to be more alert to the condition-specific issues, than to general aspects of HrQoL. Lack of significant differences in level of agreement suggested cross culturally comparable correspondence between parents’ and children’s ratings [22].

Predictor models of HrQoL which include clinical, socio-demographic and psychosocial

characteristics were tested. A hierarchical regression was performed for the total data set and for the TOCO subsamples. Since sample characteristics differed across countries, age and treatment were included as covariates in the analysis. Even though fit statistics were not optimal in all tests, the resulting models explained similar proportions of HrQoL variance across countries, suggesting again that the QoLISSY instrument performance across countries is promising.

This was also true for the QoLISSY determinant modules, additionally included in regression models to predict HrQoL total scores. Regression results indicated a large contribution of psychosocial determinants to the proportion of variance, which was already well explained by socio-demographic and clinical factors. Interestingly, including country as dummy variable in the regression predicting both patient- and parent-reported HrQoL did not alter the results, supporting again the cultural invariance of the QoLISSY tool identified within the TOCO approach. In order to test the responsiveness of the measure to psychosocial or endocrine interventions, longitudinal studies, best randomized clinical trials, are needed.

The TOCO approach chosen proved to be a helpful and transparent technique to examine the cross-cultural equivalence of patient-reported outcome measures, which can serve as an alternative to repeated psychometric analysis in each national sample. It assumes that if taking out one country of the total data set does not lead to significant changes in psychometric performance, this suggests cross-cultural comparability of the measure. However, if sufficient sample sizes can be reached in individual countries, the cross-cultural equivalence of the measure should additionally be assessed in a country-wise fashion, so that the psychometric performance can be critically examined comparing the TOCO approach to the sequential country-wise procedure.

A word of caution is necessary regarding the parent and patient sample included in the QoLISSY field test. Recruitment was at the discretion of participation clinics with the possible consequence of a selection bias and reduced generalizability of findings. Future studies should more clearly specify respective samples and recruitment procedures. They should also use longitudinal designs to investigate the responsiveness of the measure, within and across countries. This measure has been targeted to GHD and ISS patient because these conditions are primarily related to short stature and degree of GH deficiency is a common feature which is not the case with other conditions.

The QoLISSY project is an example for simultaneous cross-cultural instrument development in paediatric health care, ready to be included in epidemiological, clinical, health economic as well as quality assurance studies in Europe. The instrument has been tested for cross-cultural equivalence in five European countries so far, which share substantial societal commonalities. Extension of testing to other countries, especially outside of Europe, is ongoing and is expected to yield more information about variability or stability in QoLISSY's psychometric performance.

Acknowledgments

This study was funded by Pfizer Limited. Andreas Pleil is at the time of this study full-time employee of Pfizer Inc; Hartmut Wollmann was a full-time employee of Pfizer Inc. No authors received compensation for their contributions to the manuscript.

Authors contributed as follows:

MB was the coordinating principle investigator and provided direct input into the design and execution of the study and the preparation and review of the manuscript. JQ provided direct oversight of the study conduct, conducted child and parent interviews and debriefs, organized the qualitative and quantitative content analysis, and contributed to the preparation and review of the manuscript. NS and AR provided analysis and interpretation of the study results and assisted in the preparation and review of the manuscript. JC, MH, EM, EF, and MP were the principle investigators in Sweden, Spain, and France (EM and EF) and the UK respectively (MP additionally provided guidance and support in preparing and reviewing the manuscript). KD and DS assisted the PIs in the UK and Spain respectively in the coordinated and administered the study including the interviews and cognitive debriefs and field test, participated in the item content analysis and instrument design, and contributed to the preparation and review of the manuscript. AP and HW provided oversight of the project and participated in the item content analysis and instrument design, execution of the field test, and contributed to the preparation and review of the final manuscript.

The authors would like to thank Dr. Meinolf Noeker from the University Medical Centre Bonn; Prof. Dr. Helmuth Dörr from the University Medical Center Erlangen; Prof. Dr. med. Ursula Kuhnle- Krahl from a joint practice in Gauting; Prof. Dr. med. Rolf Peter Willig, Dr. med. Achim Wüsthof, Prof. Dr. med. Carl-Joachim Partsch and Dr. med. Bele Jakisch from the Endocrinologicum Hamburg; Dr. Jovanna Dahlgren from the University of Gothenburg, Paediatric Growth Research Centre; Prof. Ángel Ferrández Longás from the Hospital Miguel Servet, University of Zaragoza, and Prof. Antonio Carrascosa Lezcano from the Hospital Vall d'Hebrón, Barcelona; Dr. Mélanie Glattard, Géraldine Sebie-Rachad, Carine Mantoulan and Prof. Maithé Tauber from the Hôpital des Enfants CHU Toulouse; and Prof. Chris Kelnar from the Royal Hospital for Sick Children in Edinburgh.

Competing interests [TK]

The project was funded by Pfizer and the authors were provided funding for its conduct at each site. There are no additional competing interests to report.

Access to the QoLISSY instrument

The Copyright for the QoLISSY is held jointly by the University Medical - Center of Hamburg-Eppendorf, on behalf of the QoLISSY Study Group, and Pfizer Inc. The QoLISSY instrument,

together with comprehensive information of its development and validation process is published in the QoLISSY's User's Manual (Lengerich: Pabst Science Publishers; 2013). The Manual, which is available upon request, includes the patient and parent reported QoLISSY forms in all existing language versions as well as scoring information.

The QoLISSY will be made available for bonafide research and clinical purposes via the Pfizer Patient Reported Outcome website (www.pfizerpatientreportedoutcomes.com). The measure will be available free of charge for non-commercial entities – specifically academic researchers and for clinical practice use. Commercial users may be asked to pay a small fee which helps fund the upkeep of the website. In order to maintain an overview over the use of the instrument and to facilitate potential data sharing, users are asked to agree to a cooperation agreement, which is to be signed before obtaining the QoLISSY manual. All requests for access incoming through the cooperation agreement will require the approval of the QoLISSY Study Group. For those interested in conducting collaborative research projects directly with the University, please contact Dr. Quitmann (j.quitmann@uke.de) or Prof. Dr. Bullinger (bullinger@uke.de) directly.

References

- 1 Eiser C, Jenney M: Measuring quality of life. *Arch Dis Child* 2007;92:348-350.
- 2 Matza LS, Swensen AR, Flood EM, Secnik K, Leidy NK: Assessment of health-related quality of life in children: A review of conceptual, methodological, and regulatory issues. *Value Health* 2004;7:79-92.
- 3 Payot A, Barrington KJ: The quality of life of young children and infants with chronic medical problems: Review of the literature. *Curr Probl Pediatr Adolesc Health Care* 2011;41:91-101.
- 4 Gardner M, Sandberg DE: Growth hormone treatment for short stature: A review of psychosocial assumptions and empirical evidence. *Pediatr Endocrinol Rev* 2011;9:579-588; quiz 589.
- 5 The European QoLISSY Group: Quality of life in short stature youth. The qolissy questionnaire – user's manual. Lengerich, Pabst Science Publishers, 2013.
- 6 Brütt AL, Sandberg DE, Chaplin J, Wollmann H, Noeker M, Koltowska-Haggstrom M, Bullinger M: 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. *Horm Res* 2009;72:65-73.
- 7 Bullinger M, Dellenmark Blom M, Feigerlova E, Herdman M, Lunde C, Mimoun E, Pleil A, Power M, Quitmann J, Rohenkohl A, Sanz D, Skoropadskaya A, Wollmann H, Chaplin J: Assessing quality of life in short stature youth-the qolissy project focus group and cognitive debriefing experience. *Value Health* 2010;13:63.
- 8 Bullinger M, Anderson R, Cella D, Aaronson N: Developing and evaluating cross-cultural instruments from minimum requirements to optimal models. *Qual Life Res* 1993;2:451-459.
- 9 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.
- 10 Bullinger M, Koltowska-Haggstrom M, Sandberg D, Chaplin J, Wollmann H, Noeker M, Brutt AL: Health-related quality of life of children and adolescents with growth hormone deficiency or idiopathic short stature - part 2: Available results and future directions. *Horm Res* 2009;72:74-81.
- 11 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.
- 12 Bullinger M, Quitmann J, Power M, Herdman M, Mimoun E, Debusk K, Feigerlova E, Lunde C, Dellenmark-Blom M, Sanz D, Rohenkohl A, Pleil A, Wollmann H, Chaplin JE: Assessing the quality of life of health-referred children and adolescents with short stature: Development and psychometric testing of the qolissy instrument. *Health Qual Life Outcomes* 2013;11:76.
- 13 Byrne BM: Structural equation modeling with amos: Basic concepts, applications, and programming. New York, NY, Routledge/Taylor & Francis Group, 2010.
- 14 Hair JF, Black WC, Babin BJ, Anderson RE: Multivariate data analyses, ed 7. New Jersey,

Prentice Hall, 2010.

- 15 Preacher K: Calculation for the test of the difference between two independent correlation coefficients [computer software], 2002, Available from <http://quantpsy.org>
- 16 DeVellis RF: Scale development: Theory and applications. Thousand Oaks, CA, Sage Publications, Inc, 1991.
- 17 Nunnally JC: Psychometric theory, ed 3. New York, NY, McGraw-Hill, 1994.
- 18 Skevington SM, Lotfy M, O'Connell KA: The world health organization's whoqol-bref quality of life assessment: Psychometric properties and results of the international field trial. A report from the whoqol group. *Qual Life Res* 2004;13:299-310.
- 19 Ravens- Sieberer U, Bettge S: Aktuelles zum kinder- und jugendsurvey des rki (kiggs): Vorstellung des moduls "psychische gesundheit". *Epid Bull* 2004;1
- 20 Russell KMW, Hudson M, Long A, Phipps S: Assessment of health-related quality of life in children with cancer: Consistency and agreement between parent and child reports. *Cancer* 2006;106:2267-2274.
- 21 Upton P, Lawford J, Eiser C: Parent-child agreement across child health-related quality of life instruments: A review of the literature. *Qual Life Res* 2008;17:895-913.
- 22 Rescorla LA, Ginzburg S, Achenbach TM, Ivanova MY, Almgvist F, Begovac I, Bilenberg N, Bird H, Chahed M, Dobrea A, Döpfner M, Erol N, Hannesdottir H, Kanbayashi Y, Lambert MC, Leung PW, Minaei A, Novik TS, Oh KJ, Petot D, Petot JM, Pomalima R, Rudan V, Sawyer M, Simsek M, Steinhausen HC, Valverde J, Ende J, Weintraub S, Metzke CW, Wolanczyk T, Zhang EY, Zukauskiene R, Verhulst FC: Cross-informant agreement between parent-reported and adolescent self-reported problems in 25 societies. *J Clin Child Adolesc Psychol* 2013;42:262-273.