Quality of Life Research (2019) 28:2553–2563
https://doi.org/10.1007/s11136-019-02210-z

Cross-cultural selection and validation of instruments to assess patient-reported outcomes in children and adolescents with achondroplasia

Janika Bloemeke[1] · Rachel Sommer[2] · Stefanie Witt[1] · Monika Bullinger[1] · Clementine Nordon[3] · Francisco Javier Badia[3] · Felipe Luna González[4] · Antonio Leiva-Gea[4] · Francisco de Borja Delgado Rufino[4] · Fermín Mayoral-Cleries[5] · Pablo Romero-Sanchiz[5,6,7] · Verónica Clamagirand Saiz[5] · Raquel Nogueira-Arjona[7] · Klaus Mohnike[8] · Julia Quitmann[1]

Accepted: 9 May 2019 / Published online: 15 May 2019

Abstract

Purpose Achondroplasia, as the most common form of disproportionate short stature, potentially impacts the health-related quality of life (HRQOL) and functioning of people with this condition. Because there are no psychometrically validated patient-reported outcome (PRO) condition-specific instruments for achondroplasia, this study selected and tested available generic, disease-specific and under development questionnaires for possible use in multinational clinical research.

Methods A three-step approach was applied. First, a literature review and clinician/expert opinions were used to select relevant PRO questionnaires. Second, focus group discussions, including a group cognitive debriefing for piloting of the questionnaires with children/adolescents with achondroplasia and their parents, were performed in Spain and Germany. Third, a field-test study was conducted to test the psychometric properties of these instruments.

Results Six questionnaires were identified as potentially relevant in children with achondroplasia. In each country, five focus groups including a cognitive debriefing were conducted, and the results narrowed the possibilities to three instruments as most appropriate to assess HRQOL (the generic PedsQL, the height-specific QoLISSY, and the achondroplasia-specific APLES). Results of the field study indicate the QoLISSY and the PedsQL questionnaires to be most appropriate for use in clinical research at this time.

Conclusion This selection study is a step forward in assessing the impact of achondroplasia on HRQOL. Of the instruments examined, the QoLISSY and the PedsQL both capture items relevant to children with achondroplasia and have met the psychometric validation criteria needed for use in research. The APLES instrument is a promising tool that should be revisited upon psychometric validation.

Keywords Achondroplasia · HRQOL assessment · Psychometric properties · Children · Parents

✉ Janika Bloemeke
j.bloemeke@uke.de

[1] Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Martinistraße 52 | W 26, 20246 Hamburg, Germany

[2] Center for Psychosocial Medicine, Institute of Health Care Research in Dermatology and Nursing (IVDP), University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[3] LASER Analytica, Tour CIT 3 rue de l'arrivée, 75015 Paris, France

[4] Unidad de Cirugía Ortopédica y Traumatología, Hospital Universitario Virgen de la Victoria Málaga, Málaga, Spain

[5] Unidad de Salud Mental, Hospital Universitario Regional de Málaga, Instituto de Investigación Biomédica de Málaga (IBIMA), Málaga, Spain

[6] Departamento de Personalidad, Evaluación y Tratamientos Psicológicos, Universidad de Málaga, Málaga, Spain

[7] Department of Psychology and Neuroscience, Dalhousie University, Halifax, Canada

[8] University Children's Clinic, Otto von Geuricke University in Magdeburg, Magdeburg, Germany

Introduction

Achondroplasia, the most common type of skeletal dysplasia, is characterized by disproportionate short stature and systematic abnormalities in cartilage and bone that can be painful, require surgery, and be life-threatening [1, 2]. Because of musculoskeletal impairments that are associated with this condition, patients have a need for assistance in daily life, especially during childhood [3, 4]. Skeletal abnormalities and biomechanical impairments such as shorter extremities or hypotonia contribute to developmental delays in motor skills and limit functional independence in children [5]. Moreover, achondroplasia is often related to obesity which might increase morbidity, particularly joint problems [3, 6].

In addition to physical impairments, children with achondroplasia face challenges in their social life, such as bullying and social stigmatization, which may result in lower self-esteem and an increased risk of depression in adulthood [7, 8].

These physical and psychosocial consequences have a negative effect on the health-related quality of life (HRQOL) of children with achondroplasia [7]. The concept of HRQOL was coined to describe the impact of a defined health condition on well-being and functioning and reflects the patients' own perception on physical, emotional, mental, and social areas of health [9]. HRQOL is increasingly being assessed using patient-reported outcome (PRO) instruments to quantify how patients feel and function without outside interpretation [10]. Different types of PRO measures exist, including generic and condition-specific measures. Generic instruments can be applied independently of the health condition and allow a comparison of health status across populations or diseases [11, 12]. Condition-specific measures are used to identify the impact of a specific disease.

Treatment options for achondroplasia are few and commonly include surgical lower limb lengthening. Despite being painful and associated with serious complications, it can lead to improvement in HRQOL [13]. However, if lengthening is only performed in the lower limbs, the resulting disproportion between the upper and lower limbs remains a major issue and might negatively affect HRQOL [6, 14, 15]. Nonsurgical strategies, such as human growth hormone therapy, have been used to increase growth velocity; however, long-term benefits and the effect on body proportionality are still unclear. Therefore, this therapy is not yet recommended for use in achondroplasia [16, 17].

Study objectives

There is a lack of cross-cultural validated PRO instrument for young people with achondroplasia. Hence, appropriate instruments relevant to assess the described psychological and physical burden of young achondroplasia patients need to be identified or developed and tested to provide insight of the treatment benefits on HRQOL. When selecting PRO measures for research or clinical trials, quality criteria like reliability and validity must be considered, as recommended in the COSMIN (Consensus-based Standards for the selection of health Measurement Instruments) checklist [18]. This study selected and tested psychometrically validated PRO measures for potential use in multinational research studies of children with achondroplasia.

Method

This cross-sectional study, performed in three stages using a mixed method approach, was approved by the ethic committee of the University Medical Hospital Magdeburg and Otto-von-Guericke University, Magdeburg in Germany and by the Ethic Committee in Malaga.

First stage—literature and expert review

A process of appraisal and elimination was used to select relevant PRO measures to be used in children with achondroplasia. Therefore, a literature review was conducted in the medical databases PubMed and Web of Science at the end of 2014, to identify validated PRO instruments used in children and adolescents with achondroplasia or skeletal dysplasia. The search strategy included word combinations using the terms 'HRQOL', 'wellbeing', 'QoL', 'skeletal dysplasia', 'achondroplasia', 'children', 'adolescents', 'youth', 'instruments', 'measures' in various combinations. The pre-selection of the literature search was reviewed by two clinical experts, experienced in managing and treating achondroplasia, who were consulted to determine which PRO questionnaires could be relevant for use in children with achondroplasia in a clinical research setting. Measures were selected based on the following criteria: (1) skeletal dysplasia or achondroplasia, (2) children and their parents aged 5–14 years, (3) HRQOL including different assessment perspectives (child and proxy report), and (4) language availability (German and Spanish).

Second stage—focus group cognitive debriefing

Focus group discussions were conducted in Germany and Spain for piloting of the selected questionnaires, including a group cognitive debriefing approach to explore clarity, relevance, and appropriateness of the tools [19, 20]. The overall objective of the group cognitive debriefings was to explore how well the selected questionnaires capture experiences important to children with achondroplasia. The focus groups were moderated by trained experts and organized in three phases: (1) the moderators explained the purpose and organization of the meeting, (2) the participants completed the PRO questionnaires individually within the focus group setting, (3) using the think-aloud approach of cognitive debriefing interviews, the participants were invited to share their thoughts about each instrument after completion of the respective PRO measure with the group in an open discussion. Retrospective verbal probing questions were used to elicit additional information from the participants on the instruments' interpretability, comprehensiveness, and appropriateness for achondroplasia (e.g., "How easy or difficult was it to answer the questions?"). Additional areas of interest and aspects that were not covered by the instruments' but are relevant to the respondents, were discussed in the group to gain more information about the experiences of living with the condition [19, 21]. During the final session, the PRO instruments were reviewed overall and final ranking questions (e.g., "Overall, which questionnaire is the most appropriate for you/which one do you prefer?") were asked.

Patients aged 5–14 years and their parents were recruited through patient organizations in Germany and Spain. For young patients below the age of 8 years, only parents participated. Inclusion criteria were a clinically confirmed diagnosis of achondroplasia by the clinician, understanding and speaking the local language, voluntary participation, and the provision of informed consent.

In each country, five focus group discussions with six participants were planned to be conducted, including one discussion with parents of children aged 5–7 years, 8–11 years, and 12–14 years and one discussion each with children aged 8–11 years and 12–14 years. Group discussions were audio taped and transcribed using the software f4 (Version v4.2—www.audiotranskription.de/english/f4). Sociodemographic and clinical data were collected in Germany in parallel by a clinician the same day. In Spain, parents reported anthropometric measures as part of the enrollment process.

A mixed-methods sequential explanatory design, which implies analyzing quantitative and then qualitative data in two consecutive phases was used for data analysis. Properties of each questionnaire were investigated in terms of missing values and floor/ceiling effects, using the analysis program SPSS (Version 21) [22]. Floor/ceiling effects were defined as present if more than 15% of the respondents scored the highest/lowest score [23]. MAXQDA (Version 11) was used to analyze qualitative data [24]. Statements from patients and parents were sorted in the original language and documented. Content validation was analyzed in an open deductive process by screening the transcripts and abstracting important statements into categories regarding the instrument's relation to achondroplasia, it's reflection of everyday life experiences, it's importance in relation to the disease, on the general opinion about each instrument and on aspects not covered by the respective instrument that would be relevant to achondroplasia. The answers of the final ranking questions were quantified and recorded in a ranking list from the most preferred tool to the least preferred tool. In the final reconciliation process, the most relevant PRO instruments were ranked using the ranking list from the most preferred tool to the least preferred tool as reported by the participants themselves. The tools performing best in terms of qualitative and quantitative results were selected to be further tested in the third stage of the study.

Third stage—field test and validation of PRO tools

A cross-sectional study was performed in both countries for psychometric testing. Per country, 25 children/adolescents and their parents in each age group (8–12 years, 13–14 years) and 25 parents of children aged 5–7 years were planned to be recruited. Data collection procedures differed between the countries. In Germany, the questionnaires were self-completed in the patients' homes and anthropometric measures were self-reported by parents. Clinical data in Spain were collected from the centers, and questionnaires were completed under supervision of psychologists.

After description of sociodemographic and clinical data, descriptive statistics and other relevant distribution variables were reported. Psychometric properties were tested using Classical Test Theory (CTT) [25] including internal consistency measured with Cronbach's Alpha and criterion validity using Spearman's rank correlation coefficient. To determine if the measures captured variability that was expected based on anthropometric measures, correlations between the sub-dimensions of the PRO questionnaires and height standardized deviation score (SDS), the body proportionality ratio and the BMI [weight (kg)/height^2 (m)] were calculated separately for each country, to minimize the limitation concerning different data collection procedures between the countries. Height SDS scores were calculated using the general population of the children of same age and gender, on the basis of WHO Growth Reference Program [26]. The body proportionality ratio was calculated using the upper segment/lower segment ratio, which is a known ratio to determine body proportions. The ratio derived from standing and sitting heights of children computed as: *sitting*

height in cm/(*standing height in cm* − *sitting height in cm*) and is increased in achondroplasia, representing a higher disproportionality [27, 28]. Validity of the instruments was further assessed to compare HRQOL dimensions across groups with a different body proportionality index using multivariate analysis of covariance (MANCOVA), with country as a covariate to control for country differences. For total scores and when multivariate effects were significant, univariate analysis was calculated, using country as a covariate as well. Since reference data and cutoff lines for proportionality ratio in achondroplasia are lacking [28, 29], body proportionality index (ranging from 0.91 to 2.34) data were used to form three groups based on the frequency distribution of the proportionality ratio including two smaller extreme groups each ($N = 28$) and one larger middle group ($N = 72$) (little disproportionate: body proportionality index < 1.57, very disproportionate: body proportionality index 1.57–1.94, severely disproportionate: body proportionality index > 1.94). The significance level was set at $p < 05$.

Results

First stage—pre-selection of PRO tools

A total of 15 questionnaires were identified in the literature search. This pre-selection of possible PRO instruments was reviewed by the project team based on the criteria mentioned in the method section, to determine which PRO questionnaires were most likely to capture the physical and psychosocial burden of illness experienced by children with achondroplasia in clinical research. For instance, the WeeFIM [5, 30] and the PEDI-CAT [31] capture physical burden but do not capture emotional and mood-related items and were thus excluded. Hence, six PRO tools were identified as potentially relevant for achondroplasia patients: three generic pediatric QoL questionnaires (KIDSCREEN-27, PedsQL and EQ-5D-Y), one musculoskeletal-specific questionnaire (PODCI), one short-stature-specific-questionnaire (QoLISSY) and the recently developed, but at the time of the study not psychometrically validated, achondroplasia-specific APLES.

In this study, the generic 27-item version of the KIDSCREEN surveying HRQOL in children and adolescents aged 8-18 years was used, including the dimensions *physical well-being*, *psychological well-being*, *autonomy & parents*, *peers & social support*, and *school environment* [32].

The generic 23-item version of the 'Pediatric Quality of Life Inventory' (PedsQL) designed to assess HRQOL in patients with chronic or acute conditions from the patient's and parents' perspectives was used including the dimensions *physical functioning*, *emotional functioning*, *social functioning*, and *school functioning* [33].

The generic EQ-5D-Y, a modified version of the EQ-5D to appropriately measure HRQOL in children and adolescents consists of five items that refer to the domains mobility, looking after oneself, usual activities, pain/discomfort, and feeling worried/sad/unhappy. A Visual Analog Scale rates the health status between 0 (worst) and 100 (best) [34].

Parents only filled out the musculoskeletal-specific 'Pediatric Outcomes Data Collection Instrument' (PODCI) assessing functioning and HRQOL of children and adolescents with 86 items [35, 36].

The short stature-specific Quality of Life in Short Stature Youth (QoLISSY) questionnaire assesses HRQOL of children and adolescents with short stature. It has been validated for children diagnosed with achondroplasia in Germany [37]. It consists of core dimensions (*physical*, *social*, *emotional*) that are accompanied by predictors of HRQOL (*coping*, *beliefs*, and *treatment)*. Domains included in the parent version only address aspects of the child's *future* and the *effects on parents* of the child's short stature. The child's version consists of 36 items, the parents' version of 52 items [38].

The recently developed 'Achondroplasia Personal Life Experience Scale' (APLES), a pediatric achondroplasia-specific instrument to assess HRQOL, includes the dimensions: *self-perception*, *friends*, *recreation*, *kindergarten/school*, and *physical function* [39, 40]. At the start of the current study, the APLES was still under development; thus, different versions were included in the pilot-test (15 items) and the field test (21 items).

Second stage—piloting and selection of PRO tools

The pilot-test sample included 27 participants in Germany and 28 individuals in Spain (Table 1).

Children aged 8–11 years and 12–14 years filled out the child version of the QoLISSY, the KIDSCREEN-27, the PedsQL, and EQ-5D-Y questionnaire, as well as a preliminary version of the APLES with 15 items, that was at the time being still under development. Parents of children aged 5–14 years were asked to complete the parent/proxy version of these questionnaires. In order to lower the respondent burden for the children, parents only were asked to fill out the PODCI.

Focus group discussions provided feedback on the content validity and preliminary feasibility of the instruments. The QoLISSY and APLES were favored most by children and parents in both countries because they focus on short stature and topics relevant to achondroplasia. The PODCI was considered not to be adequate in content, length, and wording. Patients and parents felt none of the generic instruments captured the specific effects of achondroplasia. However, some favored the PedsQL because it reflects daily life situations and is easy and quick to complete. EQ-5D-Y and the

Table 1 Sample characteristics of the FG discussions and pilot test in Germany and Spain (stage 2)

Age group	Country	N		Child gender		Age mean (SD)
		Parents	Children	Male	Female	
5–7[a]	Germany	12	–	5	7	5.6 (SD: 1.0)
	Spain	5	–	1	4	6.3 (SD: 0.8)
8–11	Germany	5[b]	4	3	1	10.2 (SD: 1.1)
	Spain	6	6	3	3	9.3 (SD: 1.1)
12–14	Germany	3	3	0	3	13.3 (SD: 0.72)
	Spain	5	6[c]	6	0	13.6 (SD: 0.9)

Note The gender distribution was not balanced across the age groups, especially in the age group 13–14 years, in which no male patients were represented in the German sample and no females were represented in the Spanish sample

[a]Only parents were interviewed in the age group 5–7 years

[b]One additional parent participated in this group

[c]Two of the participants in this group were twin brothers, so only five families were represented

Table 2 Overview of cognitive debriefing results in Germany and Spain

	Parents (5–7 years)	Children (8–11 years)	Parents (8–11 years)	Children (12–14 years)	Parents (12–14 years)
Most preferred tool	QoLISSY	APLES	APLES	PedsQL	QoLISSY
	APLES	QoLISSY	PedsQL	APLES	APLES
	KIDSCREEN27	EQ5D-Y	PODCI	QoLISSY	PODCI
	PedsQL		QoLISSY		PedsQL
Least preferred tool	PODCI		KIDSCREEN27		
	EQ-5D-Y		EQ-5D-Y		

KIDSCREEN-27 were also rated as too general, although the distinction between family, friends, and social life in the KIDSCREEN-27 was appreciated, and the *school* domain was highly valued by the parents of the young children. An overview of the rank order preference of the final session of the cognitive debriefing is shown in Table 2.

In both countries, the completion rates across all instruments were good. However, the missing data rate was high in the school domain of the KIDSCREEN (8.93%) and the PedsQL instrument (12.86%) in the German sample. Floor effects were not present, while ceiling effects (> 15%) were observed in the EQ-5D-Y and in some subdomains of the PODCI for both countries as well as in the *friends* scale of the KIDSCREEN instrument in the Spanish sample.

Based on the qualitative and quantitative results, the QoLISSY and APLES were considered as the tools that capture the main features of achondroplasia. One group (children 12–14 years) favored the PedsQL because it was easy to answer and required minimal clarification on the items, which present a good summary of their life. Hence, it was included in the study as a generic instrument. Overall, items of these three tools were rated as clear, understandable, and relevant for achondroplasia. Instruments' measurement properties in the pilot test were good, with no evidence of significant floor or ceiling effects in all three instruments. The instruments were included in the field-test to assess their psychometric properties and to validate these tools for achondroplasia.

Third stage—field test and validation of PRO tools

Sample description

A total of 133 children participated in the field test in both countries, 73 in Germany and 60 in Spain. There were no differences in sample composition between countries, except

Table 3 Sample characteristics field test (stage 3)

	5–7 years			8–12 years			13–14 years		
	GER $n=25$	SPAIN $n=20$	Total $n=45$	GER $n=34$	SPAIN $n=30$	Total $n=64$	GER $n=14$	SPAIN $n=10$	Total $n=24$
Age mean (SD)	6.36 (.86)	6.25 (1.04)	6.32 (.94)	10.46 (1.5)	10.94 (1.4)	10.68 (1.4)	14.04 (.70)	14.06 (.59)	14.05 (.64)
Sex male (%)	12 (48.0%)	10 (50.0%)	22 (48.9%)	19 (55.9%)	13 (43.3%)	32 (50.0%)	6 (42.9%)	7 (70.0%)	13 (54.2%)
Height mean	92.80	92.05	92.46	111.30	111.96	111.61	119.11*	130.20*	123.73
(SD)	(7.52)	(5.71)	(6.71)	(8.37)	(10.12)	(9.17)	(4.15)	(12.86)	(10.28)
Missing	–	–	–	1	–	1	–	–	–
Height SDS mean	−4.79	−4.92	−4.85	−4.53	−4.93	−4.72	−5.78	−4.52	−5.29
(SD)	(1.18)	(0.79)	(1.02)	(1.06)	(1.12)	(1.09)	(0.60)	(1.70)	(1.28)
Missing	–	–	–	2	–	2	–	–	1
Prop.ratio mean	1.90	1.82	1.86	1.79	1.68	1.73	1.73	1.43	1.61
(SD)	(.26)	(.17)	(.23)	(.29)	(.27)	(.28)	(.37)	(.41)	(.41)
Missing	–	–	–	5	–	5	–	–	–
BMI mean (SD)	20.34	20.34	20.34	22.90	22.00	22.47	23.72	22.59	23.25
(SD)	(4.35)	(1.82)	(3.43)	(6.63)	(3.48)	(5.32)	(3.43)	(3.51)	(3.44)
Missing	–	–	–	2	–	2	–	–	0

SD standard deviation

**T* test revealed a significant difference between the German and the Spanish sample ($p=0.024$)

a significant difference in the mean height between the German (119.11 cm) and Spanish (130.20 cm) participants in the age group 13–14 years ($p \leq 0.05$) (Table 3).

Psychometric properties of the PedsQL

Scale characteristics of the PedsQL in the child and parent's version had a skewed mean score distribution to the right in both countries indicating a high HRQOL within the range of 0–100 (optimal HRQOL). Missing items were below 1% indicating good feasibility and floor and ceiling effects were not present. Instrument reliabilities were excellent in the German sample ($\alpha > 0.7$), while scores in the Spanish sample were slightly lower and unsatisfactory reliability was found in the *social* scale in child report ($\alpha = 0.48$) (Table 4).

In both countries, very low and nonsignificant correlations between the PedsQL subscales and the clinical variables BMI, height SDS and body proportionality ratio, suggested an unsatisfactory criterion validity. In addition, the MANCOVA for the child and parent report yielded no significant main effect of body proportionality groups on PedsQL dimensions.

Psychometric properties of the QoLISSY

In the children's version, the scale means of the QoLISSY instrument were skewed to the right, indicating a high HRQOL for children with achondroplasia in both countries. Missing items were low (< 4%), suggesting satisfactory feasibility. While a ceiling effect in the *beliefs* dimension was observed in the Spanish and German sample in children aged 8–12 years and in Spanish parents of children aged 13–15 years in the *future* scale, other floor or ceiling effects were not remarkable. Internal consistency was satisfactory in both countries ($\alpha > 0.7$), except for the *coping* scale in Spanish children ($\alpha = 0.65$) and the *physical* scale in German parents ($\alpha = 0.69$) (Table 4).

Correlations between the QoLISSY subscales and the selected clinical variables height SDS, BMI, and body proportionality ratio were significant in the Spanish sample between the *believe* scale ($r = 0.32$, $p \leq 0.05$) and BMI and between the *future* scale and BMI ($r = 0.26$, $p \leq 0.05$) in parent report. Correlations with data of Spanish children were significant between the *physical* scale and the proportionality ratio ($r = 0.33$, $p \leq 0.05$), the *social* scale and the proportionality ratio ($r = 0.33$, $p \leq 0.05$) and between the *total score* and the proportionality ratio ($r = 0.32$, $p \leq 0.05$). Using data of German children, correlations between the *physical* scale and the proportionality ratio were significant ($r = -0.36$, $p \leq 0.05$), as well as between the *physical* scale and height SDS ($r = 0.31$, $p \leq 0.05$), the *coping* scale and height SDS ($r = 0.35$, $p \leq 0.05$) and between the *believe* scale and height SDS ($r = 0.32$, $p \leq 0.05$). No significant correlations were found for the German parent report. Results of the MANOVA yielded no significant main effect of body proportionality groups on QoLISSY dimensions.

Table 4 Reliability of PRO instruments (measured with Cronbach's Alpha)

PRO instrument	Domain	Germany				Spain				Both countries			
		Parents		Children		Parents		Children		Parents		Children	
		α	N	α	N	α	N	α	N	α	N	α	N
APLES	Self-perception	0.72	71	0.76	47	0.63	60	0.39	40	0.69	131	0.61	87
	Friends	0.83	71	0.76	43	0.88	58	0.80	38	0.82	129	0.77	81
	Recreation	0.75	72	0.78	46	0.66	60	0.32	40	0.71	132	0.63	87
	Kindergarten/School	0.72	72	0.61	46	0.82	59	0.80	40	0.77	131	0.72	86
	Physical function	0.76	73	0.80	44	0.44	58	0.65	39	0.67	131	0.75	83
	Total Score[a]	0.84	69	0.87	39	0.78	55	0.69	37	0.81	124	0.80	76
QoLISSY	Physical	0.69	55	0.86	46	0.78	58	0.83	40	0.79	113	0.85	86
	Social	0.72	73	0.87	44	0.83	59	0.89	39	0.78	132	0.87	83
	Emotional	0.84	70	0.89	46	0.85	60	0.84	39	0.84	130	0.87	85
	Coping	0.81	64	0.78	43	0.81	58	0.65	39	0.78	122	0.71	82
	Beliefs	0.89	70	0.85	47	0.90	60	0.88	40	0.89	130	0.86	87
	Future*	0.84	71	–	–	0.89	60	–	–	0.85	131		–
	Effects on parents[b]	0.85	47	–	–	0.76	60	–	–	0.80	107		–
	Total Score[c]	0.88	54	0.94	43	0.91	57	0.94	38	0.90	111	0.94	81
PedsQL	Physical	0.88	72	0.88	44	0.81	58	0.65	39	0.85	130	0.81	83
	Emotional	0.80	71	0.84	47	0.72	59	0.63	39	0.76	130	0.77	86
	Social	0.81	73	0.81	47	0.68	59	0.48	39	0.78	132	0.71	86
	School	0.74	73	0.80	44	0.71	60	0.77	40	0.72	133	0.77	84
	Total Score[a]	0.91	70	0.94	41	0.86	57	0.83	39	0.89	127	0.90	80

Note Sample sizes can be lower than presented in Table 3 because of incompletion of the instruments. Mean scale scores were computed for each scale when 80% of the data were available

α = Cronbach's Alpha

[a]Sum of all scales

[b]Only assessed in parent report

[c]Sum of the scales physical, social, emotional

Psychometric properties of the APLES

Mean values of the APLES instrument indicated that in both countries parents and children report a relatively high HRQOL, which was also seen in the PedsQL and the QoLISSY. Feasibility of the instrument was good with no observed floor or ceiling effects. Cronbach's alpha scores were very satisfactory in the German sample ($\alpha > 0.7$), while scores were lower in the Spanish sample and unsatisfactory in the *recreation* and *self-perception* scales among children and in the scale *physical function* in parent report ($\alpha < 0.50$) (Table 4).

Except for a significant correlation between the scale *friends* with BMI in child report for Germany ($r = 0.32$, $p \leq 0.05$), correlations between the APLES dimensions and clinical variables were not significant in both countries. Results of the MANCOVA yielded a significant main effect of body proportionality groups on APLES dimensions for child report, Pillai's trace = 0.259, $F_{(10, 148)} = 2.204$, $p = 0.02$, $\eta_p 2 = 0.13$. However, the subsequent univariate analyses showed no significant effects and should therefore be interpreted with caution. Nevertheless, the mean values of children with a severely disproportionate body are tending to be higher than in children with a less disproportionate body in the *friends* and *physical* scale, which would surprisingly mean that children with a more disproportionate body have a higher HRQOL in these domains (Table 5). There were no significant main effects in parent report.

Discussion

Using a process of appraisal and elimination, the following instruments were identified via literature review and interviews with clinical experts: KIDSCREEN, PedsQL, EQ-5D-Y were selected as generic tools, the PODCI as a musculoskeletal-specific instrument and the QoLISSY, APLES as condition-specific instruments.

Despite the small sample size in the pilot test, the results were similar among age groups and countries. The height-specific QoLISSY questionnaire and the condition-specific APLES questionnaire were the tools most favored by children and their parents because of the clear focus on domains related to height (QoLISSY) and height as well

Table 5 Univariate analysis of variance of HRQOL for the APLES instrument between three different groups based on the body proportionality index

APLES	Group 1 Little disproportionate BPI: < 1.57 M (SD)	Group 2 Very disproportionate BPI: 1.57–1.94 M (SD)	Group 3 Severely disproportionate BPI: > 1.94 M (SD)	*F*	*p*	η_p^2
Patient-reports						
Self-perception	73.61 (18.98)	69.64 (25.55)	66.11 (30.61)	.439	.64	.01
Friends	79.93 (22.75)	84.72 (13.89)	82.77 (15.57)	.590	.55	.01
Recreation	77.34 (19.40)	64.88 (21.51)	70.41 (25.49)	2.531	.08	.06
School	77.08 (26.72)	72.22 (22.88)	71.11 (22.01)	.403	.66	.01
Physical function	55.78 (30.10)	70.11 (21.14)	72.16 (22.45)	2.487	.08	.06
Total score[a]	72.75 (13.64)	72.31 (13.49)	72.36 (14.03)	.0004	.99	.00
Parents-reports						
Self-perception	51.78 (21.19)	53.18 (24.93)	55.12 (26.14)	.127	.88	.002
Friends	73.36 (23.49)	67.46 (20.30)	61.63 (22.93)	1.992	.14	.03
Recreation	56.99 (23.85)	48.69 (23.07)	52.88 (26.93)	1.258	.28	.02
School	73.80 (24.29)	65.50 (23.95)	67.94 (21.04)	1.261	.28	.02
Physical function	53.30 (22.13)	55.13 (20.22)	48.84 (23.12)	1.107	.33	.02
Total score[a]	61.85 (14.47)	57.99 (13.53)	57.58 (14.86)	1.048	.35	.02

Note BPI body proportionality index, *M* mean (0–100 scale with 100 representing the highest HRQOL), *SD* standard deviation, *F F*-value, *p p* value, η_p^2 effect size eta square. Univariate effects were not significant ($p \geq 0.05$)

[a]Sum of all scales

as proportionality (APLES) which are relevant domains of achondroplasia. The generic PedsQL questionnaire was favored by one group and showed satisfactory results in the pilot test and was therefore included in the subsequent field-test, as well.

Results of the field-test demonstrated excellent reliability in the total scores of the instruments and adequate to satisfactory on scale levels, with the exception of low internal consistency in the APLES sub-domains *self-perception*, *recreation*, and *physical function* in the Spanish sample. However, when interpreting the results for each country separately, the small sample size needs to be considered and additional work is needed on the newly developed APLES instrument. Furthermore, the field-test results indicate that the reliability in the physical domain tended to be strongest for PedsQL in the parent-report. This might suggest that a combination of psychometrically validated tools that complement each other may be needed to best quantify the impact of children with achondroplasia: a well performing generic instrument in combination with one that specifically addresses issues related to height.

Although expected, our results did not show that anthropometric measures, such as degree of short stature, body disproportionality, and BMI, are clearly associated with impaired HRQOL in all the instrument sub-scales. The PedsQL showed no association with clinical measures, while the QoLISSY is associated with all clinical variables included in the correlation analysis. However, results of the MANOVA yielded no significant main effect of body proportionality groups on QoLISSY dimensions. But since QoLISSY was originally designed for children and adolescents with proportionate short stature who do not suffer from a disproportionate body, this result was not unexpected. Contrary to our assumption, the APLES as a condition-specific instrument showed no association with the proportionality index. Only a significant but weak association between the subscale *friends* and BMI in German children was detected. And, although a significant main effect of body proportionality groups on APLES dimensions was detected in child report, univariate analyses were not significant and should be interpreted carefully. This suggests that further refinement of the instrument is necessary to be included in further clinical trials.

Although we controlled for country in the analysis, the unexpected results regarding body proportionality might be due to differences in the recruitment procedures, data collection (clinical measurement vs. self-report), samples composition and size, as well as the lack of comparable clinical documentation across countries. Specifically, body proportionality was assessed with different methods in Germany and in Spain, and although measures were harmonized they only concerned sitting height, which is associated with the lower limb length, not including upper extremities. However, especially during infancy shortening of the limbs is present in the upper limbs, while the lower limbs are not statistically different in length compared to normal controls [41]. Hence, the upper limb length is an important indicator

when describing body proportions in achondroplasia; however, this was not included in the body proportionality index used in this study. Furthermore, a relatively high proportion of achondroplasia patients had undergone limb-lengthening surgery in Spain, but the proportion was unknown in Germany, which may have influenced the results and may reflect the significant body height difference between German and Spanish participants aged 13–14 years in the field-test.

Finally, cutoff lines and in-depth international reference data of body proportions are still lacking for disproportionate growth disorders. Hence, the proportionality groups devised for this study might not reflect the possible clinical differences and thus negatively affects the sensitivity of the instruments.

Limitations

There are several limitations of the study. First, the meaningfulness of all analyses, including the proportionality ratio, is limited, because the data were collected using different methods. Furthermore, the sample composition and small sample limit the generalizability of study results, especially in parametric methods, such as ANOVA. Participants of the focus groups and pilot-test were recruited through patient organizations in both countries and may not be generalizable to either those countries or more broadly. In the field-test, German patients were recruited through a patient organization, while Spanish patients were recruited in a clinical center in Spain. Furthermore, in Germany, anthropometric measures were self-reported by parents, which likely increased variability in the measurements and may have compromised the accuracy and reliability of the anthropometric measurements. In Spain, clinical data were collected from the centers but the questionnaires were completed under supervision of psychologists, which may have introduced a social desirability bias. However, in light of the fact that achondroplasia is a rare disease, and finding a large population is difficult, the different recruitment procedures were a pragmatic way to conduct the research study.

Conclusion

PRO measures are increasingly used in clinical trials or medical practice to understand disease burden and effects of treatment. Valid and reliable instruments are required which should be tested in the target population, as well as in a cross-cultural context. This study identified three PRO instruments as most appropriate for use in children with achondroplasia. While the short stature-specific QoLISSY and the generic PedsQL questionnaire have both been psychometrically validated and might be considered for use in research with children with achondroplasia, the condition-specific APLES requires additional work to better understand the performance of the sub-scales and to psychometrically validate the measure in different countries and cultures.

Despite poor sensitivity to clinical differences, psychometric characteristics of the instruments, especially reliability, were satisfactory. However, Cronbach's alpha values of the APLES were unacceptable in some scales. Future research with larger sample sizes and a more heterogonous cross-cultural sample is needed to have a robust understanding of the performance of these instruments, to describe the impact of achondroplasia on HRQOL, and to unravel determinants of HRQOL in order to improve the well-being and functioning of patients and their families.

Acknowledgements We would like to thank all participating families who contributed to this study, as well as the patient organizations BKMF in Germany and the ALPE Foundation in Spain. Also, we thank the team from BioMarin Pharmadeutical who sponsored the study and Laser Analytica who helped to conduct the study. Furthermore, we thank Dr. Belén Pimentel from FIMABIS for her valuable scientific support.

Funding This study was sponsored by BioMarin Europe Ltd. PR was supported through an Incorporation of PhD contracts from the University of Málaga and José Castillejo Program (CAS18/00263) from the Spanish Ministry of Science, Innovation and Universities. All authors received research funding for the conduct of the study. The authors received no financial support for the development of the manuscript.

Compliance with ethical standards

Conflict of interests All authors received research funding for the conduct of the study. The authors received no financial support for the development of the manuscript.

Ethical approval All procedures performed in the studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Ethical approval was obtained from the local ethics committees in Germany and Spain before start of the study.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Horton, W., Hall, J., & Hecht, J. (2007). Achondroplasia. *Lancet, 370,* 162–172.
2. Mohnike, K., Klingebiel, K., Vaupel, N., & Zabel, B. (2013). *Achondroplasie und Hypochondroplasie* (2nd ed.). Berlin: ABW Wissenschaftsverlag.
3. Baujat, G., Legeai-Mallet, L., Finidori, G., Cormier-Daire, V., & Merrer, M. (2008). Achondroplasia. *Best Practice & Research Clinical Rheumatology, 22*(1), 3–18.
4. Richette, P., Bardin, T., & Stheneur, C. (2008). Achondroplasia: From genotype to phenotype. *Joint Bone Spine, 75,* 125–130.

5. Ireland, P. J., McGill, J., Zankl, A., Ware, R. S., Pacey, V., Ault, J., et al. (2011). Functional performance in young Australian children with achondoplasia. *Developmental Medicine and Child Neurology, 53,* 944–950.
6. Hoover-Fong, J., Schulze, K. J., McGready, J., Barnes, H., & Schott, C. (2008). Age-appropriate body mass index in children with achondroplasia: interpretation in relation to indexes of height. *The American Journal of Clinical Nutrition, 88,* 364–371.
7. Gollust, S., Thompson, R., Gooding, H., & Biesecker, B. (2003). Living with achondroplasia in an average-sized world: An assessment of quality of life. *American Journal of Medical Genetics, 120(A),* 447–458.
8. Hunter, A. (1998). Some psychosical aspects of nonlethal chondrodysplasias: II. Depression and anxiety. *American Journal of Medical Genetics, 78,* 9–12.
9. Bullinger, M. (2002). Assessing health related quality of life in medicine. An overview over concepts, methods and applications in international research. *Restorative Neurology and Neuroscience, 20*(3–4), 93–101.
10. Patrick, D., Burke, L., Powers, J., Scott, J., Rock, E., Dawisha, S., et al. (2007). Patient-reported outcomes to support medical product labeling claims: FDA perspective. *Value in Health, 10*(2), 125–137.
11. Cohen, J., & Biesecker, B. (2010). Quality of life in rare genetic conditions: A systematic review of the literature. *American Journal of Medical Genetics, 152A,* 1136–1156.
12. Brütt, A. L., Sandberg, D. E., Chaplin, J., Wollmann, H., Noeker, M., Koltowska-Haggstrom, M., et al. (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*(2), 65–73.
13. Kim, S. J., Balce, G. C., Agashe, M. V., Song, S. H., & Song, H. R. (2012). Is bilateral lower limb lengthening appropriate for achondroplasia?: Midterm analysis of the complications and quality of life. *Clinical Orthopaedics and Related Research, 470*(2), 616–621.
14. Savarirayan, R., & Rimoin, D. L. (2002). The skeletal dysplasias. *Best Practice & Research Clinical Endocrinology & Metabolism, 16,* 547–560.
15. Kim, S. J., Agashe, M. V., Song, S. H., Choi, H. J., Lee, H., & Song, H. R. (2012). Comparison between upper and lower limb lengthening in patients with achondroplasia: A retrospective study. *Journal of Bone and Joint Surgery—Series B, 94*(B(1)), 128–133.
16. Ornitz, D. M., & Legeai-Mallet, L. (2017). Achondroplasia: Development, pathogenesis, and therapy. *Developmental Dynamics, 246*(4), 291–309.
17. Miccoli, M., Bertelloni, S., & Massart, F. (2016). Height outcome of recombinant human grwoth hormone treatment in achondroplasia children: A meta-analysis. *Hormone Research in Paediatrics, 86*(1), 27–34.
18. Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., et al. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research, 19*(4), 539–549.
19. Brod, M., Tesler, L. E., & Christensen, T. L. (2009). Qualitative research and content validity: Developing best practices basen on science and experience. *Quality of Life Research, 18,* 1263–1278.
20. Patrick, D. L., Burke, L. B., Gwaltney, C. J., Leidy, N. K., Martin, M. L., Molsen, E., et al. (2011). Content validity—Establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: Part 2—assessing respondent understanding. *Value Health, 14*(8), 978–988.
21. Cheng, K. C., & Clark, A. (2017). Qualitative methods and patient-reported outcomes: Measures development and adaption. *International Journal of Qualitative Methods, 16,* 1–3.
22. IBM Corp. (2012). *IBM SPSS statistics for windows (version 21.0)*. Armonk, NY: IBM Corp.
23. Terwee, C. B., Bot, S. D. M., Boear, M. R., Van der Windt, D., Knol, D. L., Dekker, J., et al. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology, 60,* 34–42.
24. MAXQDA. (1989-2013). Software for qualitative data analysis. Berlin: VERBI Software - Consult - Sozialforschung GmbH.
25. Petrillo, J., Cano, S. J., McLeod, L., & Coon, C. D. (2015). Using classical test theory, item response theory, and Rasch measurement theory to evaluate patient-reported outcome measures: A comparison of worked examples. *Value in Health, 18*(1), 25–34.
26. Organization W. H. (2007). WHO Growth Reference Program.
27. Turan, S., Bereket, A., Omar, A., Berber, M., Ozen, A., & Bekiroglu, N. (2005). Upper segment/lower segement ratio and armspan-height differences in healthy Turkish children. *Acta Paediatrica, 94,* 407–413.
28. Frederiks, A., van Buuren, S., van Heel, W., Dijkman-Neerincx, R., Verloove-Vanhorick, S., & Wit, J. (2005). Nationwide age references for sitting height, leg length, and sitting height/height ratio, and their diagnostic values for disproportionate growth disorders. *Archives of Disease in Childhood, 90,* 807–812.
29. Del Pino, M., Mejia, R., & Fano, V. (2018). Leg length, sitting height, and body proportions references for achondroplasia: New tools for monitoring growth. *American Journal of Medical Genetics, 176,* 896–906.
30. Ottenbacher, K. J., Msall, M., Lyon, N., Duffy, L., Ziviani, J., Granger, C., et al. (2000). The WeeFIM instrument: Its utility in detecting change in children with developmental disabilities. *Archives of Physical Medicine and Rehabilitation, 81*(10), 1317–1326.
31. Haley, S. M., Coster, W. J., Dumas, H., Fragala-Pinkham, M., Kramer, J., Ni, P., et al. (2011). Accuracy and precision of the Pediatric Evaluation of Disability Inventory computer-adaptive tests (PEDI-CAT). *Developmental Medicine and Child Neurology, 53*(12), 1100–1106.
32. Ravens-Sieberer, U., Herdman, M., Devine, J., Otto, C., Bullinger, M., Rose, M., et al. (2013). The European KIDSCREEN approach to measure quality of life and well-being in children: development, current application, and future advances. *Quality of Life Research, 1,* 1. https://doi.org/10.1007/s11136-013-0428-3.
33. Varni, J. W., Seid, M., & Rode, C. A. (1999). The PedsQL™: Measurement model for the Pediatric Quality of Life Inventory. *Medical Care, 37,* 126–139.
34. Wille, N., Badia, X., Bonsel, G., Burstrom, K., Cavrini, G., Devlin, N., et al. (2010). Development of the EQ-5D-Y: A child-friendly version of the EQ-5D. *Quality of Life Research, 19*(6), 875–886.
35. Daltroy, L., Liang, M., Fossel, A., & Goldberg, M. J. (1998). The POSNA pediatric musculoskeletal functional health questionnaire: Report on reliability, validity, and sensitivity to change. Pediatric Outcomes Instrument Development Group. Pediatric Orthopaedic Society of North America. *Journal of Pediatric Orthopaedics, 18*(5), 561–571.
36. Allen, D., Gortons, G., Oeffinger, D., Tylkowski, C., Tucker, C., & Haley, S. (2008). Analysis of the pediatric outcomes data collection instrument (PODCI) in ambulatory children with cerebral palsy using confirmatory factor analysis and item response theory methods. *Journal of Pediatric Orthopaedics, 28*(2), 192–198.
37. Rohenkohl, A. C., Bullinger, M., & Quitmann, J. (2015). Quality of life in children, adolescents, and young adults with achondroplasia. *Orthopäde, 44*(3), 212–218.

38. The European QoLISSY Group. (2013). *Quality of life in short stature youth. The QoLISSY Questionnaire—User's manual*. Lengerich: Pabst Science Publishers.
39. Sommer, R., Blömeke, J., Dabs, M., Witt, S., Bullinger, M., & Quitmann, J. (2016). An ICF-CY-based approach to assessing self- and observer-reported functioning in young persons with achondroplasia—Development of the pilot version of the Achondroplasia Personal Life Experience Scale (APLES). *Disability and Rehabilitation, 16,* 1–5.
40. Bloemeke, J., Sommer, R., Witt, S., Dabs, M., Badia, F. J., Bullinger, M., et al. (2018). Piloting and psychometric properties of a patient-reported outcome instrument for young people with achondroplasia based on the International Classification of Functioning Disability and Health: The Achondroplasia Personal Life Experience Scale (APLES). *Disability and Rehabilitation, 8,* 1–11.
41. Shelmerdine, S. C., Brittain, H., Arthurs, O., & Calder, A. (2016). Achondroplasia: Really rhizomelic? *American Journal of Medical Genetics, 170*(8), 2039–2043.