



The challenge of measuring quality of life in children with Hirschsprung's disease or anorectal malformation[☆]

Pauline Clermidi^{a,b,*}, Guillaume Podevin^{c,d}, Célia Crétolle^{d,e},
Sabine Sarnacki^{d,e}, Jean-Benoit Hardouin^b

^aDepartment of Pediatric surgery, University Hospital, Nantes, France

^bEA 4275 "Biostatistics, Pharmacoepidemiology and Subjective Measures in Health Sciences", Faculties of Medicine and Pharmaceutical Sciences, University of Nantes, Nantes, France

^cDepartment of Pediatric surgery, University Hospital, Angers, France

^dFrench Reference Center for Anorectal and Pelvic Malformations, Necker-Enfants Malades Hospital, AP-HP, France

^eDepartment of Pediatric Surgery, Necker-Enfants Malades Hospital, AP-HP, France

Received 25 October 2012; revised 20 February 2013; accepted 2 March 2013

Key words:

Quality of life;
Disease-specific quality of life questionnaire;
HAQL;
Anorectal malformation;
Hirschsprung's disease

Abstract

Purpose: The aim of the present study was to assess, after adaptation to French, the only specific quality of life (QoL) instrument for children with Hirschsprung's disease or anorectal malformation, the Hirschsprung's disease/Anorectal Malformation Quality of Life questionnaire (HAQL), in order to get a standardized QoL evaluation instrument that could further be used to help health care improvement.

Methods: The study was conducted in three teaching hospitals, including the French reference center for anorectal and pelvic malformations. After adaptation to French, QoL questionnaires were sent to the children and proxies. The questionnaire was mailed to 280 families. Psychometrics properties of the questionnaires (validity and reliability) were analysed from 120 proxy and 96 child questionnaires.

Results: The HAQL with the original structure was not acceptable. Exploratory steps led to a clinically pertinent structure that had acceptable fit and good validity and reliability properties. The final structure pools physical symptoms (continence, discomfort) and psychosocial dimensions (general well-being, social and emotional functioning) of QoL.

Conclusion: The final structure, despite the disadvantage of being a new structure, allows assessment of QoL in this population and has the advantage of being shorter and validated on the clinical postoperative questionnaire from the Krickenbeck international consensus.

© 2013 Elsevier Inc. All rights reserved.

[☆] This study received financial support from the Pfizer foundation (former Wyeth foundation for the health of children and adolescents), public call research project 2009. We also acknowledge Ms Hartman and Hanneman and colleagues (from the Division of Medical Psychology, University of Amsterdam, the Netherlands) for providing us the HAQL questionnaire and scoring supply and for their help in the use of their questionnaire.

* Corresponding author. Service de Chirurgie Infantile, Hôpital Mère et Enfant, 38 boulevard Jean-Monnet, 44093 Nantes Cedex 1, France. Tel.: +33 2 40 08 35 85; fax: +33 2 40 08 35 46.

E-mail address: pclermidi@gmail.com (P. Clermidi).

Hirschsprung's disease (HD) and anorectal malformations (ARM) are two congenital anomalies involving the digestive tract, occurring in about 1/5000 births. ARM results from developmental anomalies of the caudal pole. Malformations range from minor defect to defect involving the urinary and genital tracts. HD is characterized by congenital aganglionosis of a various length of the digestive tract most frequently involving rectum and sigmoid, rarely the whole colon and small bowel, resulting in a neonatal obstruction. The main concerns for the medical team after the surgical correction are bowel control, urinary control and sexual function. Fecal and urinary control may not be perfect even with an excellent anatomic repair, due to short length of the remaining colon and sphincter control in HD or associated malformations and deficient nerve supply in ARM. The most frequent functional symptoms encountered after treatment are constipation, soiling, diarrhea and more rarely urinary incontinence. For these patients, an effective bowel management program, including enema and dietary restrictions, has been devised to improve their quality of life (QoL) [1]. Parents carry out such treatment at home for young children and thus a significant burden of daily care has been taken up by families.

QoL evaluation tools are found useful in many fields of health evaluation. Measures of QoL that would allow distinguishing between different treatments may influence comprehensive decisions about treatment options [2,3]. To assess the efficacy of an intervention to improve QoL among children with the same diseases, comparison within those children may be needed. In these cases, disease-specific rather than generic QoL measures may be appropriate. Disease-specific QoL measures provide greater sensitivity and specificity. They may have greater salience for clinicians by better focus on functional areas of particular concern and greater responsiveness to disease-specific interventions [4,5].

In 2001, Hanneman and colleagues designed the Hirschsprung's disease Anorectal malformation QoL questionnaire (HAQL), a disease-specific instrument to assess QoL in patients suffering from HD or ARM. Their study provided questionnaires for patients from 6 years to adulthood:

- for 6 and 7 years old: a proxy form
- for 8 to 11 and 12 to 16 years old: a child and a proxy form
- for 17 years and older: an adult form.

The HAQL has been used in the Netherlands (original language of the questionnaire) and Italy. The aim of the present study was to translate and adapt for French the HAQL₈₋₁₁ (questionnaire for 8 to 11 year-old children) and the HAQL_{proxy} (questionnaire for proxy of 6- to 11 year-old children) and evaluate their measurement properties in order to get a standardized instrument of QoL evaluation for patients with ARM or HD that could further be used to help health care evaluation and improvement. If the reliability and validity could not be assessed, we would propose a new valid

scale structure for the French version of the HAQL₈₋₁₁ and HAQL_{proxy}.

1. Methods

1.1. Original HAQL questionnaire (Appendix)

HAQL questionnaires were built from 38 items grouped into dimensions that cover physical, emotional and social functions, and disease-related symptoms [6]. The version of the child questionnaire HAQL₈₋₁₁ recommended by authors was composed of 5 dimensions (27 items) [6,7]: presence of diarrhea (2 items), fecal continence (8 items), physical symptoms (9 items), emotional functioning (6 items), and body image (2 items). The version of the proxy questionnaire HAQL_{proxy} recommended by authors was composed of 9 dimensions (38 items): the same 5 dimensions as in the HAQL₈₋₁₁ and the following four: use of laxative diet (2 items), use of constipating diet (2 items), urinary continence (4 items), and social functioning (3 items). The respondent indicates how often a specific problem occurred in the past 7 days (never, sometimes, often, very often). The responses are scored for each item from 1 to 4 and linearly transformed to a 0 to 100 scale, so that higher scores indicate better quality of life. Dimensions scores are computed as the sum of item scores divided by the number of items answered if more than 50% of items are answered.

1.2. Design

The research protocol was approved by the local ethics committee.

The study was conducted in 3 teaching hospitals (Angers, Nantes and Paris-Necker), including the French reference center for anorectal and pelvic malformations. By accordance with the exclusion criteria used by Hanneman and colleagues, we excluded patients who had a cloaca, mental retardation or lack of basic proficiency in French.

The HAQL questionnaires were subjected to translation and adaptation to French [8–11].

The generic health-related QoL questionnaire used was the Vécu et Santé Perçue de l'enfant et de l'Adolescent (VSPA): VSPAe for children and VSPAproxy for proxy [12,13].

The clinical status questionnaire was based on Krickenbeck consensus for post-operative clinical results evaluation [14,15]. Voluntary bowel movements, soiling, constipation, and autonomy status were evaluated by proxies.

HAQL and VSPA questionnaires appropriate for child age and proxies were sent. In addition, face validity and clinical status questionnaires were specifically addressed to proxies. A letter on aims and use of the present study was addressed to children and proxies. A reminder letter was sent if no response was received after 3 weeks (test step). Each family who completed the test step was sent the same

questionnaires a second time 4 weeks after the first response (retest step).

1.3. Assessment of validity and reliability of the French version of HAQL

The SAS® (version 9.1, SAS Institute Inc.), STATA® (version 11.2, Stata Corp), and LISREL® (version 8.80, SSI Inc.) software programs were used for statistical analysis. Data are expressed by mean \pm standard deviation. *P* value significance level was fixed at 0.05.

1.4. Properties of the original structure [2,10].

1.4.1. Validity

Validity of the questionnaires was studied through analysis of: face validity, criterion validity and known-groups validity.

Face validity verifies a good understanding of the questions (items) by patients or their proxies. It was assessed by a specific questionnaire for the proxies about presentation and relevance of each item.

Construct validity verifies that the assumed structure of the questionnaire (relationship between items and dimensions) is valid. It was studied first through multitrait correlation matrix. It is a table representing the correlation coefficient between each item and each score computed from each dimension. Pearson's correlation coefficients (ρ) were used. When ρ was higher than 0.40 between one item and its dimension, converged validity was assessed. For one item, when its correlation was higher with its own dimension than the others, divergent validity was assessed. The second method used to analyse construct validity was the confirmatory factor analysis. It determines fit indices of the data to the assumed structure of the questionnaire [16]. Several fit indices were used: Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean square Residual (SRMR), Comparative Fit Index (CFI), and Non-Normed Fit Index (NNFI). Acceptance levels for a good fit were: RMSEA ≤ 0.06 , SRMR ≤ 0.08 , CFI ≥ 0.96 , NNFI ≥ 0.95 .

Criterion validity verifies that the scores computed with the studied questionnaire are correlated to the scores of another questionnaire measuring close concepts. It was studied by the correlations between scores to HAQL and VSPA questionnaires. The VSPA questionnaire was used as the criterion measure. VSPA is the pediatric QoL questionnaire constructed for and validated on the French population. Criterion validity was good between two close dimensions if $\rho \geq 0.40$.

Known-groups validity is the ability to discriminate between groups of patients differing in clinical status. It verifies that patients having different known clinical status will have different mean values of the scores. For the study, children were artificially divided into groups according to clinical data. We compared the mean of dimensions scores

for HAQL₈₋₁₁ and HAQL_{proxy} between groups using ANOVA [17].

Clinical groups were:

- disease: HD versus MAR
- anatomical severity at birth: 'mild' versus 'severe'

The two previous criteria are those studied by Hanneman and colleagues in the original validation article.

- present severity based on clinical status questionnaire (soiling, constipation, voluntary bowel movement) and autonomy status.

1.4.2. Reliability

If validity was assessed, the reliability of the questionnaires was evaluated through internal consistency and test-retest reliability.

Internal consistency, measured by Cronbach's α , assesses that a homogeneous concept is measured for each dimension. That is to say that the items are positively correlated and that the dimension is composed of enough items to obtain a reliable score. Internal consistency was considered correct if $\alpha_{\text{Cronbach}} > 0.70$ [18].

Test-retest reliability (stability) verifies that two replications of the questionnaire by the same patient in a short time interval and without notable clinical changes allow obtaining comparable scores. Its study was based on the comparison of the two sets of scores from the families that completed the questionnaires on two occasions with a 4 weeks interval. The test-retest agreement was evaluated by intraclass correlation (ICC) and considered good for $ICC \geq 0.60$ [17,19,20].

1.5. Exploratory methods to determine a new structure for the HAQL questionnaire

Exploratory analyses were carried out on childhood data. Results from statistical methods were confronted to clinical relevance at each step of analysis. Proxy data were secondly used to assess the structures proposed from child responses.

1.6. Validation of the new structure

The same process as described in part A was used. The choice of the final structure was made from clinical relevance and fit properties.

2. Results

2.1. Response rate (Fig. 1)

Two hundred twenty-four child questionnaires (8 to 11 years old) and 280 proxy questionnaires (proxies of 6- to 11-

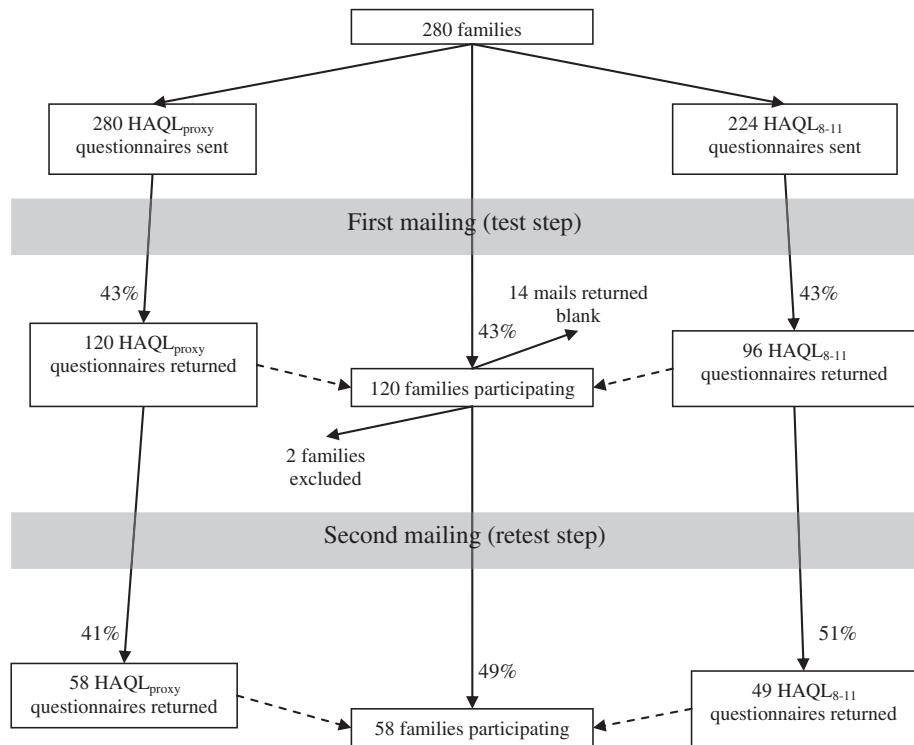


Fig. 1 Flow chart. Questionnaires appropriated for child age and proxies were sent. A reminder letter was sent if no response was received after 3 weeks. The first step response is called 'test step'. To each family who completed the test step was sent the same questionnaires a second time 4 weeks after the first response. This second step response is called 'retest step'. From the first mailing, 12 mails were returned for a not up-to-date address and 2 for death of the child since last report. Two families were excluded of the retest step because of current clinical evolution.

year-old children) were sent in March 2010. The response rate after one recall was 43% for the child and proxy groups.

2.2. Face validity

The percentage of overall missing item response was 3% for children and 16% for proxies.

No family complained on a specific item or the response structure; there was also a lower item response rate for some specific items, which might suggest face validity issues for those items.

2.3. Properties of the original structure after translation

2.3.1. Validity

Construct validity: the multitrait analysis of HAQL₈₋₁₁ revealed 13 over 27 items with a poor convergent or divergent validity. Dimensions 'body image' and 'physical symptoms' are the dimensions with the worst results, respectively, 2 over 2 and 8 over 9 items with poor convergent and/or divergent validity.

Confirmatory factor analysis could not assess a good fit of the structure ($\text{RMSEA} = 0.15$, $\text{SRMR} = 0.30$, $\text{CFI} = 0.84$, $\text{NNFI} = 0.83$).

As the original structures of HAQL₈₋₁₁ in French do not reach validity criteria, the structure of the questionnaire required improvement to be able to be used in French.

2.4. Proposition of a new structure: properties and characteristics of the proposed structure

Results of several statistical methods were pooled in order to propose structures with a strong psychometric coherence. The propositions were confronted to the judgments of clinicians, with the possibility to remove some items in order to propose a structure coherent from the psychometric and clinical points of views.

The proposed structure encompasses 25 items in 7 dimensions (Appendix): 'general well-being', 'diarrhea and feeling of urge', 'diurne fecal continence', 'fecal continence and social functioning', 'urinary continence', 'emotional functioning', 'abdominal discomfort'.

2.4.1. Validity

Construct validity: the multitrait matrix analysis of HAQL₈₋₁₁ revealed only one misclassified item in divergent validity (item 39) and no convergent validity issue. Confirmatory factor analysis of HAQL₈₋₁₁ confirmed adequacy of the proposed model to the data ($\text{RMSEA} = 0.06$, $\text{SRMR} = 0.07$, $\text{CFI} = 0.96$, $\text{NNFI} = 0.95$). HAQL_{proxy} did not reach adequacy criteria: $\text{RMSEA} = 0.13$, $\text{SRMR} = 0.10$, $\text{CFI} = 0.76$, $\text{NNFI} = 0.73$.

Criterion validity (Table 1): all dimensions of HAQL₈₋₁₁ are significantly correlated to at least one dimension of VSPAe. Pearson correlation coefficient was higher than 0.40

Table 1 Criterion validity: correlation between dimensions of HAQL₈₋₁₁ (proposed structure) and VSPAe questionnaires.

Dimensions	VSPAe						
	Family	Self-esteem	Vitality	Friends	General well-being	Leisure	School
HAQL ₈₋₁₁ Diurne fecal continence	*	*	*	*	**	*	
Diarrhea and feeling of urge	*	*	*	*	0.43		
Abdominal discomfort	*	*	*		*		
General well-being					**		
Fecal continence and social functioning		*			0.43		
Urinary continence	**	*			0.40		
Emotional functioning	0.43 **	0.43 **	*	*	0.49 **		
	0.61						

In gray: significant correlation between a HAQL₈₋₁₁ (proposed structure) dimension and a VSPAe dimension.

* $p < 0.05$; ** $p < 0.001$; Pearson correlation coefficients shown if ≥ 0.40 .

between close dimensions of HAQL₈₋₁₁ and VSPAe. Dimensions ‘school work’ and ‘family’ of VSPAe did not have close dimension in HAQL₈₋₁₁ and no Pearson correlation coefficient reached 0.40 (except between dimensions ‘abdominal discomfort’ of HAQL₈₋₁₁ and ‘family’ of VSPAe).

Known-groups validity (Table 2): every dimension scores of HAQL₈₋₁₁ and HAQL_{proxy} were able to discriminate between different soiling score, except ‘abdominal discomfort’ of HAQL_{proxy} ($P = 0.10$). On the contrary, the constipation score was not discriminated by any dimension of the HAQL₈₋₁₁, and only by ‘urinary continence’ and ‘emotional functioning’ of the HAQL_{proxy}. The voluntary bowel movements score was discriminated by the dimension ‘diarrhea and feeling of urge’ of both HAQL₈₋₁₁ and HAQL_{proxy}, never by the dimensions ‘urinary continence’, ‘emotional functioning’ or ‘abdominal discomfort’. Every dimension of HAQL₈₋₁₁ and HAQL_{proxy} discriminated autonomy status except ‘abdominal discomfort’ of HAQL_{proxy}.

Disease severity was discriminated by the overall score, dimensions ‘diarrhea and feeling of urge’, ‘abdominal discomfort’ and ‘general well-being’, but not by ‘urinary continence’, ‘fecal continence and social functioning’ with close between HAQL₈₋₁₁ and HAQL_{proxy}.

Disease status was discriminated by ‘diarrhea and feeling of urge’ and ‘abdominal discomfort’ of HAQL₈₋₁₁ and HAQL_{proxy} and by ‘urinary continence’ in addition for HAQL_{proxy}.

2.4.2. Reliability (Table 3)

Internal consistency and *reproducibility* were good for all dimensions of HAQL₈₋₁₁ except for ‘abdominal discomfort’ ($\alpha_{Cronbach} = 0.65$ and $ICC = 0.59$) and “emotional functioning” ($\alpha_{Cronbach} = 0.73$ but $ICC = 0.27$). In HAQL_{proxy}, internal consistency results for ‘diarrhea and feeling of urge’ and ‘abdominal discomfort’ were respectively $\alpha_{Cronbach} = 0.64$ and $\alpha_{Cronbach} = 0.58$. Test-retest on HAQL_{proxy} showed

five dimensions and the overall scale with good concordance, and two with a weak concordance: ‘abdominal discomfort’ and ‘emotional functioning’ (respectively $ICC = 0.33$ and $ICC = 0.27$).

3. Discussion

3.1. On methodology

Assessment of a patient’s QoL by a proxy or health-care professional is likely to differ from the patient’s own assessment. Proxies often underestimate QoL compared to the patients and surgeons used to tend to focus on physiological functioning as the major indicator of QoL.

Priority should be given to child QoL evaluation by a child questionnaire in now-a-day consensus. However the use of proxies is of particular relevance in pediatric practice where patients may be considered unreliable because of difficulties with reading, language and memory [21]. The use of both child and proxy evaluations may be relevant.

The original structure of HAQL questionnaire is not valid in French. Thus, the questionnaire should not be used as is in French. Making any change to the structure modifies the psychometric properties. A QoL questionnaire with good psychometric properties cannot be assumed to be transferable to other populations including a different age bracket or proxy. In determining the content of scales, different procedures could be adopted: either an expert approach or a patient-centered approach. We choose to favour the child measure over the proxy one and to mix both approaches, in order to construct a new questionnaire’s structure resulting from a compromise between good performance characteristics and clinical relevance [2,3,9,22].

Comparing results and QoL of patients based on anatomical classification at birth is problematic because

Table 2 Known-groups validity: capacity of the dimensions scores of the proposed scale's structure to discriminate groups of children (8–11 years old) differing in clinical status determined by anatomical classification or by the clinical status questionnaire.

Proposed structure: scores by dimension	Anatomical classification		Clinical status questionnaire			
	Disease status	Severity status	Autonomy status	Soiling	Constipation	Voluntary bowel movements
Overall score	553 ± 123 (0.12)	Mild: 578 ± 115 (0.0092) Severe: 509 ± 124	Usual A.: 602 ± 91 (<.0001) Lower A.: 471 ± 116	Grade 0: 648 ± 47 Grade 1: 606 ± 67 Grade 2: 495 ± 118 Grade 3: 419 ± 97 (<.0001)	553 ± 123 (0.50)	Grade 0: 595 ± 89 Grade 1: 571 ± 103 Grade 2: 524 ± 128 Grade 3: 490 ± 145 (0.0370)
Fecal continence and social functioning	92 ± 17 (0.86)	92 ± 17 (0.21)	Usual A.: 96 ± 14 (0.0062) Lower A.: 84 ± 22	Grades 0 or 1 or 2: 96 ± 11 Grade 3: 76 ± 31 (0.0001)	92 ± 17 (0.51)	95 ± 9 (0.12)
Diarrhea and feeling of urge	HD: 68 ± 32 (0.0040) ARM: 83 ± 22	Mild: 82 ± 21 (0.0069) Severe: 68 ± 34	Usual A.: 86 ± 23 (0.0002) Lower A.: 63 ± 27	Grades 0 or 1: 90 ± 13 Grade 2: 69 ± 33 Grade 3: 49 ± 29 (0.0012)	77 ± 27 (0.44)	Grade 0: 86 ± 19 Grades 1 or 2 or 3: 74 ± 28 (0.0426)
Abdominal discomfort	HD: 63 ± 29 (0.0002) ARM: 78 ± 19	Mild: 78 ± 22 (0.0001) Severe: 61 ± 26	Usual A.: 76 ± 25 (0.0149) Lower A.: 62 ± 20	Grades 0 or 1: 80 ± 21 Grades 2 or 3: 61 ± 31 (0.0163)	72 ± 25 (0.95)	72 ± 25 (0.90)
General well-being	72 ± 27 (0.35)	Mild: 77 ± 24 (0.0479) Severe: 64 ± 30	Usual A.: 81 ± 23 (<.0001) Lower A.: 55 ± 27	Grade 0: 91 ± 14 Grades 1 or 2: 73 ± 23 Grade 3: 47 ± 30 (<.0001)	72 ± 27 (0.25)	Grade 0: 80 ± 23 Grade 1: 77 ± 22 Grades 2 or 3: 61 ± 30 (0.0150)
Diurne fecal continence	64 ± 33 (0.49)	Mild: 70 ± 32 (0.0906) Severe: 55 ± 33	Usual A.: 76 ± 28 (<.0001) Lower A.: 44 ± 32	Grade 0: 98 ± 4 Grade 1: 75 ± 20 Grades 2 or 3: 32 ± 24 (<.0001)	64 ± 33 (0.66)	Grade 0: 75 ± 30 Grade 1: 65 ± 34 Grades 2 or 3: 51 ± 32 (0.0168)
Urinary continence	90 ± 19 (0.91)	90 ± 19 (0.96)	Usual A.: 96 ± 7 (0.0001) Lower A.: 83 ± 22	90 ± 19 (0.41)	90 ± 19 (0.97)	90 ± 19 (0.19)
Emotional functioning	88 ± 18 (0.23)	88 ± 18 (0.18)	Usual A.: 92 ± 14 (0.0409) Lower A.: 86 ± 13	Grades 0, 1 or 2: 95 ± 10 Grade 3: 81 ± 19 (0.0003)	88 ± 18 (0.91)	88 ± 18 (0.0196)

Usual A. = Usual autonomy; Lower A. = Lower autonomy.

P value in brackets. P value in bold if <0.05.

Clinical status is defined by the anatomical classification (disease: HD versus ARM; and disease severity 'mild' versus 'severe'). Clinical status questionnaire is based in the international classification for post-operative results from Krickenbeck consensus and on autonomy level defined subjectively by the proxies compare to the same age group (8–11 years old). Post-operative evaluation from Krickenbeck consensus evaluates soiling, constipation and voluntary bowel movement. Results scores range from grade 0 (no symptom) to grade 3 (symptom).

Dimensions scores range from 0 (lower quality of life) to 100 (higher quality of life). Overall score range from 0 (lower quality of life) to 700 (higher quality of life).

For significant discriminant score (level of signification: P < 0.05), the average score of each clinical status is indicated.

Table 3 Reliability of the HAQL proposed structure.

Questionnaires	Dimensions	Number of items	Internal consistency Cronbach's α coefficient	Test-retest concordance Intraclass correlation
HAQL ₈₋₁₁	Overall scale	25	0.93	0.83
	Fecal continence and social functioning	5	0.89	0.82
	Diarrhea and feeling of urge	3	0.78	0.80
	Abdominal discomfort	2	0.65	0.59
	General well-being	4	0.85	0.69
	Diurne fecal continence	4	0.89	0.70
	Urinary continence	4	0.89	0.69
HAQL _{proxy}	Emotional functioning	3	0.73	0.27
	Overall scale	25	0.91	0.56
	Fecal continence and social functioning	5	0.88	0.58
	Diarrhea and feeling of urge	3	0.64	0.72
	Abdominal discomfort	2	0.58	0.33
	General well-being	4	0.84	0.65
	Diurne fecal continence	4	0.89	0.73
HAQL _{proxy}	Urinary continence	4	0.88	0.84
	Emotional functioning	3	0.71	0.27

In bold: values for dimensions with good internal consistency (Cronbach's $\alpha > 0.70$) or good reproducibility (ICC ≥ 0.6).

HD and ARM are a spectrum of malformations. A classification for postoperative results was proposed by the Krickenbeck consensus to get a uniform international scoring system for comparable follow-ups. Three relevant parameters were chosen: voluntary bowel movements, soiling and constipation [15]. That classification has the advantage of coming from an international consensus, being short and easy to use in repeated medical evaluation, and should be integrated in the French follow-up recommendation for ARM. We choose to use that classification, besides the disease, disease severity, and autonomy status, as a reference of clinical status to compare scores for the study of known-group validity.

3.2. On results

Lower item response rate for specific items is caused either by an unadapted item to the situation of the child, by a formulation issue or conceptual understanding issues (items 30-31-32-33 and 35) or, for proxy response, by the fact that children did not always report symptoms or what occurs at school (items 27-28, 32-33-34 and 41) or how they feel to their parents (items 10-13-16).

Most dropped items in the construction process of the proposed structure were items found to have lower response rate suggesting non-adapted items. Some items found to have a medium response rate were kept in because we judged them clinically pertinent (items 10-16-27-28).

Reliability of the questionnaire with the proposed structure was good except for the dimension 'abdominal discomfort' which did not meet the requirement but was close to the cut off. Cronbach's alpha value for a dimension is affected by the size of that dimension. Cronbach's alpha value gets higher as a dimension is built with more items.

'Abdominal discomfort' is a small dimension. This may be a reason for our results in that dimension.

Low test-retest reliability may reflect actual change among some individuals instead of poor reliability. The four-week time interval was considered short enough to limit this phenomenon. Even so, two families declare significant change in the meantime and therefore were excluded of test-retest analysis. Shortening the time interval may expose a memory bias. Asymmetric item score distribution may lead to a ceiling effect as seen in bad results in some item reproducibility evaluations. It was a compromise accepted in expectance of favouring the measure of QoL changes.

Criterion validity: comparison with another QoL questionnaire is limited because no other specific questionnaire for children with ARM or HD is available. Information given by the VSPAe was generally correlated to information given by HAQL₈₋₁₁. However, the dimension 'school' was not significantly correlated to any dimension of HAQL₈₋₁₁. This could mean that the QoL aspect represented by school is not so much correlated to the severity of the disease or its impact on daily life, as if 'school' was a preserved dimension of child life. Correlation between dimensions 'abdominal discomfort' of HAQL₈₋₁₁ and 'family' of VSPAe was not expected but well understood if abdominal discomfort is seen as a symptom of anxiety. Criterion validity for HAQL₈₋₁₁ with the proposed structure was satisfactory even if comparison between the generic and specific QoL questionnaires requires caution.

Construct validity: with the proposed structure, dimensions with the best discriminating property were the ones pertaining to diarrhea, fecal continence and its consequences on daily life. They especially discriminate between disease severity status and soiling score. The voluntary bowel movements score was poorly discriminated by HAQL,

maybe because it is not a main concern for young children aged 8 to 11 and does not affect their daily QoL so much even if it is an important criterion for result evaluation by the medical team. Dimensions pertaining to urinary continence and emotional functioning had noteworthy better discriminating power on proxy than the child questionnaire. There may be a different analysis of these symptoms by children and proxies. Constipation score was not discriminated by any dimension of the HAQL₈₋₁₁. Hypothesis of a symptom not affecting the QoL may be done, but it may also be caused by the extremely asymmetrical response distribution.

3.3. On the proposed structure choice:

Exploratory methods may lead to different structures with their own advantages and statistical properties. Nevertheless, dimensions built by the different exploratory procedures were usually close and the same items were rejected.

Choices had been done to propose only one new structure. As in the original questionnaire, we chose to keep the dimension pertaining to urinary continence even if its discriminating validity was not assessed by the present study because it seemed to be an important symptom in the daily life of children and parents and in the evolution of the diseases. This symptom was rarely declared so we faced a ceiling effect. The disease status and the severity status were not important criteria for the emotional functioning score. The emotional functioning dimension discriminated only autonomy status of the clinical status questionnaire in the child questionnaire. However, it also seemed an important aspect of QoL.

As the HAQL was considered long (38 items), validating a shorter one (25 items) may be considered as an improvement. It facilitates use in the clinical setting and population-based health surveys. The original structure of the HAQL questionnaire, because of heterogeneity between child and proxy questionnaire in the number of items and dimensions, may lead to difficulty in comparing QoL scores. The published studies using the HAQL did not use the proxy questionnaire [7,23–25]. Recent recommendations advise the use of a concomitant measure of the child and proxy feeling [3]. We proposed the same structure for the child and the proxy questionnaires to favour homogeneity and ease of use, even if properties of the proxy questionnaire could be enhanced by modification in its structure.

The HAQL questionnaire may now be used in French speaking countries. The adaptation process should be continued to dispose of that QoL evaluation instrument in more countries.

4. Conclusion

The original versions of HAQL₈₋₁₁ and HAQL_{proxy} in French does not have satisfactory psychometric properties. We propose a new valid structure of the HAQL₈₋₁₁ and the HAQL_{proxy}, which can be used with the first advantage of a

shorter and simpler questionnaire and the second advantage of discriminant validity confirmed with the clinical status questionnaire constructed after the Krickenbeck consensus.

Appendix. HAQL₈₋₁₁ and HAQL_{proxy} questionnaires, original and proposed structures

Item numbers are in brackets.

1. HAQL₈₋₁₁ (27 items/5 dimensions) and HAQL_{proxy}, (38 items / 9 dimensions) original structures

Presence of diarrhea (HAQL₈₋₁₁ and HAQL_{proxy})

(24) Thin stools; (25) More than four times a day thin stools

Fecal continence (HAQL₈₋₁₁ and HAQL_{proxy})

(9) Important to be in the neighbourhood of a toilet; (36) Loss of feces before reaching the toilet; (37) Soiling at daytime; (38) Soiling at nighttime; (39) Loss of feces at nighttime; (40) Loss of feces during physical activity; (41) Loss of feces during emotional moments; (42) Loss of feces during coughing or sneezing

Physical symptoms (HAQL₈₋₁₁ and HAQL_{proxy})

(27) Feeling swollen (abdomen); (28) Feeling no urge while bowels are full of feces; (29) Going to the toilet while having no urge; (30) Difficulty to lose feces; (31) Difficulty to discriminate between flatulence or feces; (32) Being flatulent; (33) Difficult to let a flatus; (34) Having bowel movements; (35) Having abdominal pain

Emotional Functioning (HAQL₈₋₁₁ and HAQL_{proxy})

(10) Being ashamed of leaving the classroom to go to the toilet; (11) Feeling that you are teased more than other kids; (14) Feeling ashamed; (15) Feeling different; (16) Feeling less appreciated by others; (43) Feeling afraid that others will smell your feces

Body Image (HAQL₈₋₁₁ and HAQL_{proxy})

(12) Feeling less attractive; (13) Feeling dissatisfied with your body

Constipating Diet (HAQL_{proxy})

(2) Eating special food on purpose to get thick stools; (3) Avoid eating special food to maintain thick stools

Urinary Continence (HAQL_{proxy})

(5) Loss of urine before reaching the toilet; (6) Loss of urine during physical activity; (7) Loss of urine during emotional moments; (8) Loss of urine during coughing or sneezing

Laxative Diet (HAQL_{proxy})
 (1) Eating special food on purpose to get thin stools; (4)
 Avoid eating special food to maintain thin stools

Social Functioning (HAQL_{proxy})
 (44) Daily activities; (45) Staying the night elsewhere;
 (46) Social activities

Dropped item
 (26) Solid stools

2. HAQL₈₋₁₁ and HAQL_{proxy}, proposed structure (25 items/7 dimensions)

General well-being
 (9) Important to be in the neighbourhood of a toilet; (11)
 Feeling that you are teased more than other kids; (14) Feeling
 ashamed; (15) Feeling different

Diarrhea and feeling of urge
 (25) More than four times a day thin stools; (28) Feeling
 no urge while bowels are full of feces; (39) Loss of feces at
 nighttime

Diurne fecal continence
 (36) Loss of feces before reaching the toilet; (37) Soiling
 at daytime; (40) Loss of feces during physical activity; (43)
 Feeling afraid that others will smell your feces

Fecal continence and social functioning
 (41) Loss of feces during emotional moments; (42) Loss
 of feces during coughing or sneezing; (44) Daily activities;
 (45) Staying the night elsewhere; (46) Social activities

Urinary continence
 (5) Loss of urine before reaching the toilet; (6) Loss of urine
 during physical activity; (7) Loss of urine during emotional
 moments; (8) Loss of urine during coughing or sneezing

Emotional functioning
 (12) Feeling less attractive; (16) Feeling less appreciated
 by others; (10) Being ashamed of leaving the classroom to go
 to the toilet

Abdominal discomfort
 (27) Feeling swollen (abdomen); (34) Having bowel
 movements

Dropped items
 (2) Eating special food on purpose to get thick stools; (3)
 Avoid eating special food to maintain thick stools; (24) Thin
 stools; (38) Soiling at nighttime; (13) Feeling dissatisfied
 with your body; (29) Going to the toilet while having no urge;
 (30) Difficulty to lose feces; (31) Difficulty to discriminate

between flatulence or feces; (32) Being flatulent; (33)
 Difficult to let a flatus; (35) Having abdominal pain; (1)
 Eating special food on purpose to get thin stools; (4) Avoid
 eating special food to maintain thin stools; (26) Solid stools

References

- [1] Levitt MA, Peña A. Anorectal malformations. *Orphanet J Rare Dis* 2007;2:33-46.
- [2] Eiser C, Morse R. Quality-of-life measures in chronic diseases of childhood. *Review*. *Health Technol Assess* 2001;5(4):1-157.
- [3] Varni JW, Limbers CA, Burwinkle TM. Parent proxy-report of their children's health-related quality of life: an analysis of 13,878 parents' reliability and validity across age subgroups using the PedsQL™ 4.0 Generic Core Scales. *Health Qual Life Outcomes* 2007;5:2.
- [4] Guyatt GH, Deyo RA, Charlson M, et al. Responsiveness and validity in health status measurement: a clarification. *J Clin Epidemiol* 1989;42(5):403-8.
- [5] Deyo RA, Patrick DL. Barriers to the use of health status measures in clinical investigation, patient care, and policy research. *Med Care* 1989;27(3 Suppl):S254-68.
- [6] Hanneman MJG, Sprangers MAG, De Mik EL, et al. Quality of life in patients with anorectal malformation or Hirschsprung's disease: development of a disease-specific questionnaire. *Dis Colon Rectum* 2001;44(11):1650-60.
- [7] Hartman EE, Oort FJ, Aronson DC, et al. Explaining change in quality of life of children and adolescents with anorectal malformations or Hirschsprung disease. *Pediatrics* 2007;119:e374-83.
- [8] Amarenco G, Bayle B, Lagache D, et al. Construction et validation des échelles de qualité de vie. *Ann Réadapt Méd Phys* 2000;43:263-9.
- [9] Bullinger M, Alonso J, Apolone G, et al. Translating health status questionnaires and evaluating their quality: the IQOLA Project approach. *International Quality of Life Assessment*. *J Clin Epidemiol* 1998;51(11):913-23.
- [10] Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993;118(8):622-9.
- [11] Herdman M, Fox-Rushby J, Badia X. A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. *Qual Life Res* 1998;7(4):323-35.
- [12] Auquier P, Clément A, Sapin C, et al. Development, validation of HRQL measurement in children: VSP-Ae. *Qual Life Res* 2001;10(3):268.
- [13] Siméoni M-C, Sapin C, Antoniotti S, et al. Health-related quality of life reported by French adolescents: a predictive approach of health status? *J Adolesc Health* 2001;28:288-94.
- [14] Demirogullari B, Ozen IO, Karabulut R, et al. Colonic motility and functional assessment of the patients with anorectal malformations according to Krickenbeck consensus. *J Pediatr Surg* 2008;43: 1839-43.
- [15] Holschneider A, Hutson J, Peña A, et al. Preliminary report on the International Conference for the Development of Standards for the Treatment of Anorectal Malformations. *J Pediatr Surg* 2005;40: 1521-6.
- [16] Hu LT, Bentler P. Cutoff criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equation Model* 1999;6:1-55.
- [17] Fayers PM, Machin D. *Quality of life*. 2nd ed. Wiley; 2007.
- [18] Nunnally JC. *Psychometric theory*. New York: McGraw-Hill; 1978.
- [19] McGraw KO, Wong SO. Forming inferences about some intraclass correlation coefficients. *Psychol Methods* 1996;1(1):30-46.
- [20] Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;2:420-8.
- [21] Trajanovska M, Catto-Smith AG. Quality of life measures for fecal incontinence and their use in children. *J Gastroenterol Hepatol* 2005;20:919-28.

- [22] Varni JW, Limbers CA, Burwinkle TM. How young children can reliably and validly self-report their health-related quality of life? An analysis of 8,591 children across age subgroups with the PedsQL™ 4.0 Generic Core Scales. *Health Qual Life Outcomes* 2007;5:1.
- [23] Grano C, Aminoff D, Lucidi F, et al. Self-efficacy, postoperative care satisfaction, body image and sexual functioning in ARM patients. *Pediatr Surg Int* 2008;24:1201-5.
- [24] Hartman EE, Oort FJ, Sprangers MAG, et al. Factors affecting quality of life of children and adolescents with anorectal malformations or Hirschsprung disease. *J Pediatr Gastroenterol Nutr* 2008;47(4): 463-547.
- [25] Grano C, Aminoff D, Lucidi F, et al. Disease-specific quality of life in children and adults with anorectal malformations. *Pediatr Surg Int* 2010;26:151-5.