

A review of patient-reported outcomes for children and adolescents with obesity

Bani Ahuja · Anne F. Klassen · Randi Satz ·
Nikita Malhotra · Elena Tsangaris ·
Matthew Ventresca · Nora Fayed

Accepted: 11 June 2013 / Published online: 26 June 2013
© Springer Science+Business Media Dordrecht 2013

Abstract

Introduction Obesity is a chronic condition that can impact the physical, emotional, mental and social elements that encompass a child's life. The objectives of this study were to identify which generic and obesity-specific patient-reported outcome (PRO) instruments are used in obesity literature, as well as review their conceptual approach, health and health-related content, ethical content and psychometric properties.

Method PubMed, CINAHL, EMBASE and PsycINFO were searched from the inception of each database to May 2012 to identify all studies using multi-dimensional PRO instruments with children who are overweight or obese. The most common generic and all obesity-specific instruments were analyzed according to the study objectives.

Results From 4,226 articles identified by our search, 70 articles used 6 generic and 4 obesity-specific PRO instruments. While the most commonly used PRO instrument was the generic PedsQL 4.0 (used in 53 studies), many health domains were found in the obesity-specific instruments that are not measured by the PedsQL 4.0. Summary of the development and psychometric properties of the generic and obesity PROs identified that no one instrument meets all the guideline criteria for instrument development and validation, e.g., only one instrument included qualitative input from children with obesity in the content development phase.

Discussion This comprehensive review provides information to aid in selecting multi-dimensional PRO instruments in children with obesity according to various aspects of content as well as psychometric properties. The conceptual analysis shows that the reviewed PRO instruments contain inconsistencies in their conceptual approaches. Also, certain relevant health domains to children and youth with obesity were not included in the most commonly used generic instrument. The obesity-specific instruments require further validation before they can be used in intervention studies.

B. Ahuja (✉)
Collage Pediatric Therapy, 1100 Central Pkwy W, Mississauga,
ON L5C 4E5, Canada
e-mail: bani.ahuja@gmail.com

A. F. Klassen
Department of Pediatrics, McMaster University, HSC 3A, 1280
Main Street West, Hamilton, ON L8S 4K1, Canada

R. Satz
Infant and Child Development Services Peel, Trillium Health
Partners, 6660 Kennedy Road, Suite 200, Mississauga, ON L5T
2M9, Canada

N. Malhotra
University of Ottawa, Ottawa, ON, Canada

E. Tsangaris · M. Ventresca
Department of Clinical Epidemiology and Biostatistics,
McMaster University, Hamilton, ON L8N 3Z5, Canada

M. Ventresca
Faculty of Applied Health Sciences, Brock University, St.
Catharines, ON L2S 3A1, Canada

N. Fayed
Division of Healthcare and Outcomes Research, Toronto
Western Research Institute, University Health Network, Toronto,
ON, Canada

Keywords Quality of life (QOL) · Health-related quality of life (HRQOL) · Child · Obesity · Patient-reported outcome · International classification of functioning disability and health (ICF)

Abbreviations

CHQ	Child Health Questionnaire
FDH	Functioning disability and health
HRQOL	Health-related quality of life
HUI	Health Utilities Index
KINDL	KINDL (registered) Quality of Life Questionnaire
ICF-CY	International Classification of Functioning, Disability, and Health Children and Youth Version
IWQOL-Kids	Impact of Weight on Quality of Life-Kids
PEDSQL	Pediatric Quality of Life Inventory
PRO	Patient-reported outcome
QOL	Quality of life
WHO	World Health Organization
YQOL-S	Youth Quality of Life-Surveillance Version
YQOL-W	Youth Quality of Life-Weight

Introduction

Overweight or obesity in children has become a prominent health challenge in the twenty-first century. Currently, an estimated 42 million children are obese or overweight globally [1]. This chronic condition can impact upon many aspects of the lives of children and adolescents including the physical, emotional, and social domains of life. These domains should be addressed when choosing an outcome measure and can be measured with functioning disability and health (FDH), health-related quality of life (HRQOL) and quality of life (QOL) patient-reported outcome (PRO) instruments.

In child health services and obesity research, the terms FDH, HRQOL and QOL are often used interchangeably, which poses problems with the interpretation of content of instruments designed to assess these concepts [2, 3]. World Health Organization (WHO) definitions can serve as a basis for delineating the conceptual approach that is measured regardless of the PRO instrument's title. A biopsychosocial view of health as conceptualized through *functioning* is found in the International Classification of Functioning, Disability, and Health Children and Youth version (ICF-CY) as ratified by the WHO [4]. The FDH approach focuses on impairments, capacity, performance and barriers and facilitators of health. In contrast, the WHO-QOL group defines QOL as “a (child's) perception of their position in life... in relation to their goals, expectations, standards and concerns.” In the WHO-QOL approach, mention of perceptual or subjective elements of life must be explicit in order to measure a child's QOL [5]. When HRQOL is subsumed under the WHO-QOL's definition, it

includes the child's *perception* of the health and health-related states (Fig. 1).

Even if the overall approach within a PRO instrument is consistent within a FDH, HRQOL or QOL concept, the specific health and health-related content within an instrument can vary considerably. The ICF-CY classifies specific health and health-related components of body functions and structures and the activities that constitute a child's participation in life roles, all of which occur in the context of a child's social and physical environment as well as the personal factors that make each child unique. These components and categories can delineate the specific health content found in FDH, HRQOL and QOL PRO instruments. For example, one HRQOL instrument can assess a child's emotional and physical domains, while another can assess cognition, peer relationships and family support. Coding the content of PRO instruments using the ICF can highlight the gaps and overlap between different instruments.

The specific health and health-related content of PRO instruments should be relevant to children who are obese as well as the purpose to which the PRO instruments will be applied. The extent to which the content of different generic and obesity-specific instruments reflects health issues that are relevant to children with obesity has to date been under explored. While generic instruments can provide a valuable means for comparing various groups of children (e.g., those with and without obesity), disease-specific instruments are hypothesized to have more relevant content covering the health concerns of the target group (i.e., obesity). Thus, obesity-specific PRO instruments are considered more appropriate for measuring change related to obesity-related health improvements, though this hypothesis requires empirical verification.

There is a paucity of information that directly compares the health content of the available obesity-specific instruments with generic PRO instruments. Such information is needed for researchers and clinicians to weigh their options in the selection of instruments and in the interpretation of differences in scores between groups or over the course of an intervention. The ICF-CY classification has been useful for making the health content of such instruments explicit in childhood epilepsy, cancer, cerebral palsy and other conditions [2, 6, 7].

The ethics of having children and adolescents complete FDH, HRQOL and QOL instruments has been underexplored in reviews of PRO instruments. The nature of these instruments is such that they are typically designed to collect information directly from participants and/or their proxies without cueing, interpretation or debriefing from third parties such as clinicians or other care providers [8]. The items are potential indicators that can alert children to aspects of their health for which they were not aware, having an impact on the way children view themselves

Fig. 1 Conceptual approaches of FDH, HRQOL and QOL applied to PRO instruments

Concept to Measure	Definition Used	Graphic Representation	Item Example
Functioning, Disability & Health (FDH)	The interaction or the individual components of body functions, activities and participation that occur in the context of the environment or personal factors		<p>During the past month, tell us how much you found it hard to swing, climb, skip, bounce a ball, or jump rope because of your size?*</p> <p>*Item 6 from Sizing Me Up</p>
Health-related quality of life (HRQOL) a sub-component of QOL	A child's goals expectations, standards or concerns <i>about</i> their overall health and health-related domains		I worry that my thoughts of food distract me from other things
Quality of life (QOL)	A child's perception of their position in life in the context of the culture and value systems in which they live in relation to their goals, expectations, standards and concerns		My life is as good as I ever expected

[9, 10]. Thus, the extent to which negative content and phrasing is present in PROs should also be reviewed to confidently apply such instruments in an ethical manner.

Finally, in addition to reviewing the conceptual approach, the gaps and overlaps of health domains, and the ethical amount of negative content of PRO instruments, their selection and application require a review of psychometric properties. Many generic measures have been described and reviewed [11–20], but detailed information about the psychometric properties of the obesity-specific outcomes is yet to be reviewed.

The objectives of this review are as follows: (1) to identify the most commonly used PRO instruments and characteristics of their use in the literature; (2) to describe, compare and contrast the conceptual approach to measurement with WHO and related definitions (i.e., FDH, HRQOL or QOL), health and health-related domains (e.g., emotional, physical and social domains using the ICF), as well as the ethical (negative) content; and (3) to summarize the development and psychometric properties of PRO instruments used with pediatric patients who are overweight or obese.

Methods

Literature search to identify commonly used PRO instruments, their characteristics and use

We conducted a systematic review of English language articles in PubMed, CINAHL, EMBASE and PsychINFO

from the inception of each database until May 2012. A comprehensive list of keywords was used to identify articles about being overweight or obesity and QOL and/or PRO instruments. Abstract and title screening were performed by one reviewer who removed abstracts, theses, review articles and articles that were not about children, obesity and/or FDH/HRQOL/QOL. The full text of all remaining papers was obtained and examined independently by two reviewers. Articles were retained if the following criteria were met: published in English; participants were aged up to 21 years; participants had a BMI percentile >85; and participant or a proxy completed a generic and/or obesity-specific PRO instrument. Studies were excluded if the PRO instrument used met the following criteria: measured only one health domain (e.g., fatigue, pain); was ad hoc (i.e., one without published evidence of a development or validation process); was not intended for children or youth; or was a modified version of adult PRO instrument. Citations of the included articles were examined in order to identify any studies that might have been missed in the search.

Overview of content analysis procedure

The included PRO instruments were coded by two independent reviewers on an item by item basis; if there was disagreement in coding, a third reviewer was consulted to reconcile discrepancies. Only the final consensus list of conceptual approach (FDH, HRQOL, QOL), health and related content (ICF codes and categories), and ethical (negative) content is reported.

Analysis of conceptual approach (FDH, HRQOL, QOL) using WHO definitions

Included instruments were assessed using a method specifically intended to measure the conceptual approach (e.g., FDH vs. HRQOL vs. QOL) as shown in Fig. 1 and validated in content analyses of cancer-specific and generic child PRO instruments. The WHO definition of functioning in the context of environment and personal factors was used to code if an FDH approach to measurement was used; the WHO-QOL definition of QOL that emphasizes a child's subjective perceptions about life was used to code if a QOL approach was used. Items were coded as HRQOL if they explicitly mentioned a child's personal perception of a health or health-related domain irrespective of the type of health domain (emotional, physical, social) mentioned. Examples of this coding can be found in Fig. 1.

Analysis of specific health and health-related content using the ICF-CY

The health and health-related content of PRO instruments for children was analyzed using a method that codes each item using the ICF-CY. This method has been validated in reviews of childhood cancer, epilepsy, cerebral palsy [2, 6, 7] and across conditions and is reported in detail in previous publications, which have been cited.

Analysis of ethical content

Finally, both the content and phrasing of each item were coded as negative or neutral/positive using the method described by Fayed et al. [2]. All steps were performed by two independent trained content analysts on an item by item basis.

Summary of psychometric properties

Eligible instruments identified by the search were appraised for adherence to published guidelines and criteria for the development and validation of PRO instruments [2, 8]. When information about instrument development and validation was lacking from an article, an attempt was made to contact the corresponding author for further information. Two reviewers extracted findings about the psychometric properties of each qualifying PRO instrument and compared findings to ensure consensus. For each instrument, we examined whether recommended procedures for item generation, item reduction and psychometric evaluation were used in the development process, and examined whether minimum standards for internal reliability [21] and reproducibility [22] were achieved (i.e., reliability coefficients of at least 0.70 for group level

comparisons) and findings for construct validity and responsiveness.

Results

Commonly used instruments, their characteristics and use

Figure 2 displays the search strategy implemented that retrieved a total of 70 publications used in this study. These articles came from 12 different countries; most research was conducted in the USA (42 publications).

Six generic and four obesity-specific pediatric measures were identified from the 70 publications (see Figs. 3, 4 for frequency of use). Fifteen publications used an obesity-specific questionnaire (of which five used only an obesity-specific). Sixty-five publications used a generic instrument (of which 55 used only a generic). The most commonly used PRO instrument, the generic Pediatric Quality of Life (PedsQL 4.0), was used in 53 publications compared with eight publications for the most commonly used obesity-specific measure, Impact of Weight on Quality of Life (IWQOL-Kids). Since the PedsQL 4.0 dominated a disproportionate level of generic instrument use and the

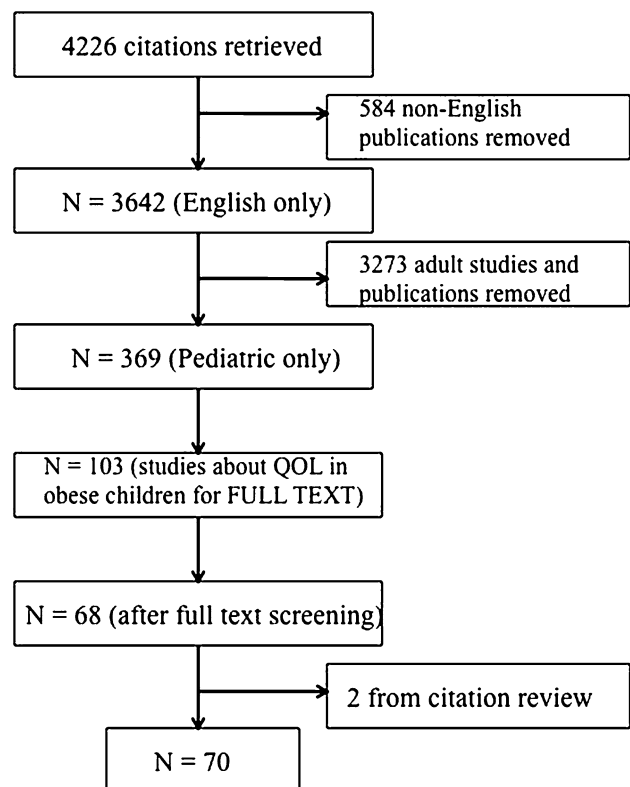


Fig. 2 Application of the inclusion and exclusion in the literature search

remaining generic measures were used with negligible frequency, only the PedsQL 4.0 was retained for the subsequent analyses. The four obesity-specific PRO instruments included in this review were as follows: Impact of Weight on Quality of Life (IWQOL-Kids); KINDL Quality of Life Questionnaire: Obesity Module; Sizing Me Up (self-report)/Sizing Them Up (parent proxy); and Youth Quality of Life–Weight module (YQOL-W).

Content analysis

Conceptual approach (FDH, HRQOL, QOL) using WHO definitions

The dominant approach found in each PRO instrument was FDH and not HRQOL or QOL (see Table 1). The YQOL-W was the only instrument to include a QOL/HRQOL approach for a substantial proportion of its items (33.3 % measured QOL/HRQOL).

Health and health-related content using the ICF-CY

The ICF components emphasized among all the PRO instruments varied (see Table 2). *Activities and participation* was the dominant ICF component in the PedsQL 4.0, and one of the principal components in one of the obesity-specific instruments, i.e., IWQOL. The KINDL and Sizing Me Up/Sizing Them Up emphasized the *body functions* component. The focus of the IWQOL was on contextual factors of environment, whereas YQOL-W focused on personal factors components of the ICF (Table 2).

No specific ICF-CY (health and health-related) categories were used across all five instruments (see Table 3). The ICF-CY category *b152 emotional functions* were the most common category included within the *body functions* health component across instruments. The category *b530 weight maintenance functions* were the second most common category across instruments, although it was not included in the PedsQL 4.0. While recurring health and related domains found across the different PRO instruments in the *activity and participation* component were not evident, *d activity and participation* and *d550 eating* were

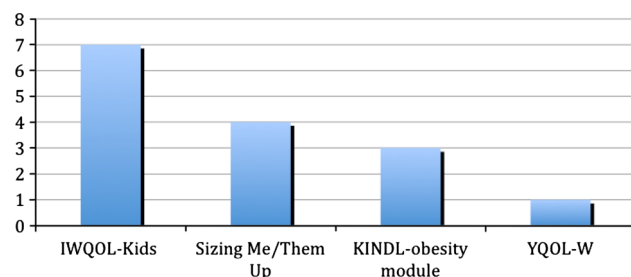


Fig. 3 Number of publications for each obesity-specific instrument

the more common categories each appearing in three PRO instruments. In terms of environment, the *e4 attitudes* category was represented in all the obesity-specific measures and *e425 Individual attitudes of acquaintances, peers, colleagues, neighbours and community members* was represented in the PedsQL 4.0. Categories representing social supports for children with obesity were inconsistently represented.

Analysis of ethical content

The analysis of negative content (see Table 4) showed that almost all PRO instruments cue respondents to answer questions by using negative phrasing such as difficulties or problems. The proportion of items classified as negative content in the instruments was moderate for all the PRO instruments with the exception of the YQOL-W, which included negative content for the majority of items.

Summary of psychometric properties

Table 5 shows the methods used to develop each instrument as well as psychometric validation.

Below we summarize key findings for the PedsQL 4.0 and each obesity-specific instrument.

PedsQLTM4.0

The PedsQL 4.0 Generic Core Scales are a generic PRO instrument designed to measure HRQOL in children aged 2–18 years. This instrument was developed from initial research with pediatric cancer patients [23–25], and over time was developed into the current generic instrument for use with any pediatric population. Although item generation and reduction for the early PCQL are described, the process that lead to the development of the PedsQL 4.0 is not as clearly delineated. The 4.0 version of the PedsQL has separate versions for child and parent report and different age-modules (i.e., 2–4, 5–7, 8–12, 13–18 years).

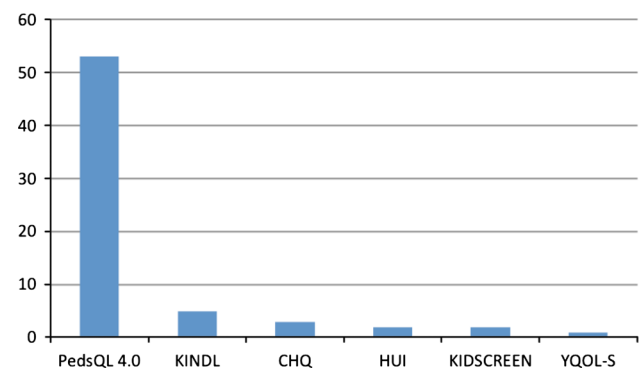


Fig. 4 Number of publications for each generic instrument

Table 1 Proportion of conceptual approaches (%) found in each measure

	PedsQL	IWQOL-Kids	KINDL obesity module	Sizing Me Up	YQOL-W
QOL/HRQOL	4.4	0.0	0.0	0.0	33.3
Functioning, disability and health (FDH)	95.6	100	100	100	66.7
Main conceptual approach	Measure of FDH with one QOL item	Measure of FDH	Measure of FDH	Measure of FDH	Measure of primarily FDH with an HRQOL subcomponent.

Table 2 Proportion of ICF-CY (%) components in quality of life measures

	Generic Measure		Obesity-Specific Measures		
	PedsQL	IWQOL-Kids	KINDL Obesity module	Sizing Me Up	YQOL-W
Total # of ICF codes	28	28	32	54	43
Body functions (<i>b</i>)	32.1	14.3	56.3	55.6	23.3
Activities & participation (<i>d</i>)	50.0	35.7	15.6	27.8	16.3
Environment (<i>e</i>)	10.7	42.9	12.5	9.3	11.6
Personal factors (<i>pf</i>)	3.6	7.1	15.6	5.6	48.8
Not-covered by the ICF (<i>nc</i>)	3.6	0	0	0	0
Covered but not defined by the ICF (<i>nd</i>)	0	0	0	1.9	0

Body structure component(s) is not included; none was found in included PROs

Content measures four domains (physical functioning, emotional functioning, social functioning and school functioning) with 23 items. An exception is that the module for completion by parents of children aged 2–4 has 21 items. Respondents are asked to rate how much of a problem each item has been in the past month on a scale ranging from 0 (never a problem) to 4 (almost always a problem). Scale scores are generated and range from 0 to 100, with higher scores reflective of better HRQOL. In addition, a total score and summary scores for physical and psychosocial HRQOL can be computed.

The PedsQL 4.0 has been studied extensively, and a listing of all PedsQL 4.0 publications is available; elsewhere, the review of which is beyond the scope of this paper [26]. Table 5 summarizes the results from two large-scale early psychometric studies [27, 28]. First, a study of 963 children and 1,629 parents recruited from a pediatric healthcare setting [27]. Feasibility was assessed by calculating the percentage of missing values, which was 2 % for both self- and proxy-report. In regards to internal consistency, for the total sample as well as each age-group, for both self- and proxy-report, the Cronbach α coefficient values for the total and summary scores exceeded the minimum 0.70. For the sample as a whole, one scale (school function self-report) was below 0.70. By age-group, values for at least one age-group were below 0.70 for social and school function for self-report, and emotional, social and school function for proxy-report. Using

ANOVA, the hypothesis that the PedsQL 4.0 could distinguish between healthy, acutely ill and chronically ill subgroups was supported for each scale as well as the summary scores and total score for both self- and proxy-report. Finally, factor analysis for self- and proxy-report provided evidence that was largely consistent with the a priori hypothesized five-factor structure [27]. Second, a large-scale study involved 10,241 families (response rate 51 %) who completed the PedsQL 4.0 in a statewide mail survey to evaluate enrollees in the State's Children's Health Insurance Program [28]. In this study, missing data were minimal and the majority of scales and total scores exceeded the minimum reliability standard of 0.70. The scales were also able to distinguish between health children and those with chronic health conditions, and scores were related in hypothesized directions with indicators of healthcare access, days missed from school, days sick in bed or too ill to play and days needing care.

Impact of weight on quality of life-Kids (IWQOL-Kids)

The IWQOL-Kids purport to measure weight-related QOL from the perspective of youth aged 11–19 years. The content of the questionnaire was modeled after the adult IWQOL [29, 30] and IWQOL-Lite [31, 32] and developed from literature on child/adolescent obesity and clinical expertise, but with no patient interviews. Exploratory factor analysis was used for item reduction resulting in 27

Table 3 Frequency of specific ICF-CY codes represented in included PROs

ICF Category	ICF Code	Generic measure PedsQL	Obesity-specific measures			
			IWQOL-Kids	KINDL	Sizing Me Up	YQOL-W
Body functions						
Confidence	b1266		1			6
Energy level	b1300	1				
<i>Sleep functions</i>	b134	1				
<i>Memory functions</i>	b144	1				
<i>Emotional functions</i>	b152	3		2	7	
Insight	b1644			1		
Experience of self	b1800		1			
Body image	b1801		1	1		2
<i>Sensation of pain</i>	b280	2				
Pain in lower limb	b28015		1			
Sensation of pain, other specified and unspecified	b289					1
<i>Respiratory muscle functions</i>	b445			1	1	
<i>Exercise tolerance functions</i>	b455				1	1
General physical endurance	b4550	1				
Fatiguability	b4552			1		
<i>Weight maintenance functions</i>	b530			12	19	
<i>Growth maintenance functions</i>	b560				1	
Activity and participation						
<i>Activities and participation</i>	d	2		1	1	
<i>Focusing attention</i>	d160	1		1		
Sitting	d4103		2		1	
Bending	d4105		1			
<i>Transferring oneself</i>	d420					
<i>Lifting and carrying oneself</i>	d430	1				
<i>Mobility</i>	d4			1	1	
<i>Hand and arm use</i>	d445				1	
Walking short distances	d4500	1				
<i>Moving around</i>	d455		1		1	
Climbing	d4551		1		1	
Running	d4552	1				
Jumping	d4553				1	
<i>Moving around using equipment</i>	d465				1	
<i>Washing oneself</i>	d510	2				
<i>Dressing</i>	d540					2
Taking off clothes	d5401		1			
Choosing appropriate clothing	d5404		1			
<i>Eating</i>	d550			1	1	1
Shopping	d6200		1			
<i>Doing housework</i>	d640	1				
Assisting others in interpersonal relations	d6603				1	
Forming relationships	d7200		1			
Relating with equals	d7402	1				
<i>Intimate relationships</i>	d770					1
<i>School education</i>	d820	3			2	
<i>Acquiring, keeping and terminating a job</i>	d845					1

Table 3 continued

ICF Category	ICF Code	Generic measure PedsQL	Obesity-specific measures			
			IWQOL-Kids	KINDL	Sizing Me Up	YQOL-W
<i>Recreation and leisure</i>	d920				3	1
Sports	d9201	1	1	1		
Socializing	d9205				1	1
Environment						
Products or substances for personal consumption	e1100			1		
General products and technology for personal use in daily living	e1150					1
<i>Design, construction and building products and technology of buildings for private use</i>	e150		1			
Design, construction and building products and technology for entering and exiting buildings for public use	e1500				1	
<i>Support and relationships</i>	e3					1
<i>Health professionals</i>	e355					
<i>Attitudes</i>	e4		4	2	1	3
Individual attitudes of immediate family members	e410		6	1		
Individual attitudes of acquaintances, peers, colleagues, neighbors and community members	e425	2			3	1
<i>Attitudes, unspecified</i>	e499		1			
Health services	e5800	1				

Body Structure component(s) is not included; none was found in included PROs

items measuring four constructs (physical comfort, body esteem, social life, family relations) plus a total score. Five response options are provided and range from “always true” to “never true.” Item scores are summed and transformed to a 0–100 scale, with higher scores representing better outcome. Psychometric properties were studied in a sample of 642 youth and included internal consistency reliability (Cronbach’s α ranged 0.88–0.96). Test–retest reliability was examined in a separate psychometric publication [33] with intraclass correlation coefficients that ranged from 0.75 to 0.88. Validity was examined through correlations and mean group differences for subgroups by zBMI and clinical versus community samples, with significant differences found in the hypothesized direction (higher BMI related with lower purported HRQOL). Convergent validity was demonstrated with higher correlations found between similar scales of the PedsQL 4.0 than between non-similar scales. Responsiveness was measured with 80 weight camp participants who reported significant improvements on all IWQOL-Kids scales after intervention. A separate publication mentions a parent-completed version of the IWQOL-Kids, though psychometric information is not reported [34].

Sizing Me Up/Sizing Them Up

Sizing Me Up [35] and Sizing Them Up [36] questionnaires (developed in parallel) are self- and proxy-report obesity-specific measures for children aged 5–18 years. The content of both instruments was developed from peer-reviewed literature of issues in child/adolescent obesity as well as expert advice from three pediatric obesity clinicians and researchers. Patient interviews were not conducted. Psychometric information came from studies of 220 obese youth and their parents (Sizing Them Up) and 141 obese children (Sizing Me Up). Exploratory factor analysis was used for item reduction resulting in two 22-item instruments that measure similar constructs with some overlapping content. Domains include physical function, emotional function, teasing and marginalization, Positive Social Attributes, social avoidance, school function and mealtime challenges. Four response options are provided to measure frequency from “never” to “always.” For each scale, the relevant items are summed and transforming to a 0–100 scale, with higher scores indicating better outcome. A total score can also be computed by adding together the 22 items. A separate 7-item parent-report scale was

Table 4 Percentage of negative phrasing and content in quality of life measures

Measure	Generic or obesity-specific	Negative phrasing	Negative content
Pediatric quality of life inventory (PedsQL 4.0)	Generic	87.0	47.8
Impact of weight on quality of life (IWQOL-Kids)	Obesity-specific	92.6	40.7
KINDL (obesity module)	Obesity-specific	75.0	50.0
Sizing Me Up	Obesity-specific	72.7	56.3
Youth quality of life–Weight quality of life module (YQOL-W)	Obesity-specific	100.0	95.2

Table 5 Summary of psychometric properties including instrument development and validation

	IWQOL-Kids	KINDL	Sizing Me Up/ Sizing Them Up	YQOL-W	PedsQL
Item generation					
Patient interviews	No	No	No	Yes	Yes
Expert opinion	Yes	No	Yes	Yes	No
Literature review	Yes	No	Yes	Yes	No
Item reduction					
Endorsement frequencies	No	No	No	Yes	No
Item redundancy	Yes	No	Yes	Yes	No
Expert opinion	No	No	No	Yes	No
Missing data	Yes	No	No	Yes	Yes
Factor analysis	Yes	No	Yes	Yes	Yes
Tests of scale assumptions	No	No	No	Yes	Yes
Psychometric analyses					
Acceptability	Yes	No	No	Yes	Yes
Internal consistency reliability	Yes	Yes	Yes	Yes	Yes
Item-total correlations	Yes	No	Yes	Yes	Yes
Inter-rater reliability	No	No	No	No	No
Test–retest reliability	Yes	No	Yes	Yes	No
Validity					
Within scale	Yes	No	Yes	Yes	Yes
Comparison with other measures	Yes	No	Yes	Yes	Yes
Hypothesis-testing (of construct validity)	Yes	No	Yes	Yes	Yes
Responsiveness	Yes	No	No/Yes	No	No

developed to measure adolescent developmental adaptation and includes items about dating, hobbies and extracurricular activities. Cronbach's α values were above 0.70 for all domains and the total scores, with the exception of the Positive Social Attributes scale (0.59 for Sizing Them Up; 0.68 for Sizing Me Up). Intraclass correlations were above 0.70 for all scales and the total score, with the following exceptions: Teasing/marginalization (0.67 Sizing Them Up; 0.58 Sizing Me Up), Positive Social Attributes (0.60 Sizing Them Up), Social Avoidance (0.53 Sizing Me Up) and Emotional Function (0.66 Sizing Me Up). Convergent validity was assessed through correlations of scale scores to the domains and total scores of the generic PedsQL 4.0

(physical, emotional, social, school) and the obesity-specific IWQOL-Kids (physical comfort, body esteem, social life, family relations), with small-/moderate-to-high correlations reported for similar domains. Validity was examined through correlations between scale scores and zBMI, with higher zBMI associated with poorer HRQOL for physical function (Sizing Them Up) and emotional function (Sizing Me Up). Responsiveness was examined in the Sizing Them Up study 6 months after bariatric surgery, with significant improvements in all scale scores except for Mealtime Challenges domain. Overall, this instrument has been carefully developed and appears to meet most psychometric standards.

KINDL quality of life questionnaire (KINDL)

The KINDL obesity module [37] is a clinical subscale of the generic KINDL [38] that captures specific experiences associated with pediatric overweight or obesity from either the child or parent perspective. This 12-item instrument captures content from six areas: physical well-being, emotional well-being, self-esteem, family, friends and everyday functioning. Five response options are provided for each item to measure frequency from “never” to “all of the time.” Items are summed and transformed to values between 0 and 100, with higher scores indicating better outcome. There is very limited information available for the development and psychometric evaluation of this instrument. In a multi-centered study of 1,916 overweight and obese patients aged 8–16 years seeking treatment in Germany, a Cronbach α of 0.77 is reported for the scale [39]. No other information about this scale was found.

Youth Quality of Life–Weight module (YQOL-W)

The YQOL-W is a weight-specific measure for completion by youth aged 11–18 years who are obese [40, 41]. The content of this instrument was developed in a manner that included in-depth interviews with 68 adolescents, input from an expert panel and an examination of existing instruments. The authors report that this 21-item instrument aimed to reflect the WHO definition of QOL [42], which is measured with three domains as follows: self, social and environment. Item reduction involved a range of methods (e.g., expert input, item redundancy, missing data and factor analysis). All items have an 11-point scale anchored by “not at all” to “very much”, and scale scores are transformed on a 0–100 scale with higher scores indicating better outcome. In a psychometric study involving 443 subjects [40], the developers reported corrected item–total correlations that ranged from 0.69 to 0.85, Cronbach’s α that ranged from 0.90 to 0.95 (0.97 for the total score) and test–retest reliability and intraclass correlation coefficients that ranged from 0.71 to 0.73 (0.77 for the total score). Validity was examined through correlations with scale scores and zBMI, with higher zBMI significantly associated with self, social, environment and the total score, indicating that as weight increased, weight-related QOL decreased. Construct validity was assessed through correlations with the Children’s Depression Inventory and the YQOL (generic measure of QOL developed by the same team). Worse YQOL-W total scores were related with higher depression scores and poorer generic YQOL. Overall, this instrument has been comprehensively developed and appears to meet all of psychometric standards. The responsiveness of the instrument was not reported.

Discussion

Our findings indicate that the most often used generic instrument and all four obesity-specific instruments included in this study reflected a conceptual approach most consistent with a biopsychosocial view of FDH as opposed to HRQOL or QOL. A conceptual understanding of the approach to measurement of any particular PRO instrument is needed so that researchers and clinicians can make conclusions about what aspects of life or health distinguish between groups of children with obesity. This conceptual awareness of what is being measured with a PRO instrument is also important to determining whether it is FDH, HRQOL or QOL that changes (or not), following interventions [43]. For example, a child or family might report that a counselling intervention made the child “better” but a lack of change in scores on a FDH-based instrument might still be observed. This lack of change in scores is more easily interpreted if one did not expect the child’s emotional, physical and social health (FDH) to change but their perception of their position in life (QOL) did change. Thus, the conceptual approach to measurement one chooses to employ in using a PRO instrument is not merely an academic exercise but has practical implications [3, 20]. This review highlights that the use of any of the obesity-specific measures, with the exception of the YQOL-W, means that the focus is a FDH approach to measurement, not an HRQOL or QOL approach. Reviews of the conceptual approach found in other generic measures can further highlight which PROs for children can be used to measure HRQOL or QOL if those approaches are needed [11, 44].

This review found important differences in the health and health-related content of the most popular generic PRO instrument used in childhood obesity, the PedsQL 4.0, compared with that of obesity-specific PRO instruments. Certain domains essential to assessing the health of children with obesity, such as *weight maintenance* functions and *eating*, are overlooked by the PedsQL 4.0 yet captured by the obesity-specific PROs [35, 38]. Also, the *attitudes* of people in the child’s social environment were emphasized in the obesity-specific instruments but not in the PedsQL 4.0 [25, 35, 41]. In practical terms, a generic measure might be helpful when the purpose is to compare the health of children with obesity to children without. However, one must consider what health domains are being overlooked in adopting only a generic strategy both in terms of comparing children with obesity to healthy children, or to evaluating obesity interventions. Our review showed that the PedsQL 4.0 was the most often used PRO instrument in the context of obesity and this finding was not limited to cross-sectional studies involving non-obesity groups. If only a generic PRO instrument is chosen, this approach raises the

question about whether it is acceptable and/or clinically meaningful to use an instrument with a patient group if that instrument overlooks the issues that matter the most to them. As stated in the 2009 US Food and Drug Administration guidelines, since the PRO measure aims to capture the patient's experience, it cannot be considered as a credible instrument if there is no evidence of its use from the target population [8].

The analysis of health content demonstrated remarkable differences in the domains that were measured among the PRO instruments. This was not an issue of the health domains included in the generic versus obesity-specific instruments, but rather an issue of consistency within the four included obesity-specific instruments. Although the 2009 US Food and Drug Administration guidelines emphasizes the importance of including qualitative input from the target population in the development of a PRO instrument that evaluates their care [8], this standard was met only by the YQOL-W. The variety of methods used to generate items among measures, including but not limited to literature review, expert input and adaptation from other PROs, might account for the variety in health domains expressed among the obesity-specific instruments. Regardless, this analysis highlights how difficult it will be to summarize the effectiveness of obesity interventions that use different PROs as the basis for evaluation. Such an issue can be resolved through the adoption of one rigorously created PRO instrument, the creation of a core set, or even data harmonization strategies, none of which have yet been demonstrated to date in our awareness of the literature.

From a psychometric standpoint, there were two important shortcomings observed in most of the obesity-specific PRO instruments reviewed here; the first being the aforementioned lack of qualitative patient input and the second the lack of responsiveness studies. Ultimately, the rationale behind selecting an obesity-specific PRO often relates to its potential to be sensitive to changes from childhood obesity interventions. If this is the case, studies of the responsiveness of the PRO instrument to detect changes will be important to their adoption in the evaluation of care.

Finally, PRO instruments are part of the realm of assessments available to researchers and clinicians for measuring FDH, HRQOL or QOL of children. Yet, exposing children to these instruments has ethical implications that should not be ignored. Qualitative studies of parents of children with health conditions have shown that merely answering many personal and negative questions found in PRO instruments can have a negative impact [45] and there is reason to expect this to be the case for children who respond for themselves. Thus, in selecting a PRO instrument, researchers and clinicians must be aware of the extent to which negative phrasing and content are present and compensate through debriefing or other strategies as necessary.

Conclusion

This extensive review of the conceptual, health and ethical content of PRO instruments for children with obesity as well as the psychometric properties of obesity-specific instruments demonstrates that no one instrument demonstrates all the characteristics assessed. Instruments that included qualitative input from children and youth with obesity had greater potential for demonstrating a holistic conceptual approach to PRO measurement yet the responsiveness of such tools need to be further validated.

Acknowledgments Anne Klassen is a recipient of the Canadian Institute of Health Research career award, and Nora Fayed is a recipient of the Canadian Institute of Health Research postdoctoral fellowship.

References

1. World Health Organization. (2013). Childhood overweight and obesity. Resource Document. Geneva: World Health Organization. <http://www.who.int/dietphysicalactivity/childhood/en/>. Accessed on January 2 2013.
2. Fayed, N., Schiariti, V., Bostan, C., Cieza, A., & Klassen, A. (2011). Health status and QOL instruments used in childhood cancer research: Deciphering conceptual content using World Health Organization definitions. *Quality of Life Research*, 20(8), 1247–1258.
3. Fayed, N., De Camargo, O. K., Kerr, E., Rosenbaum, P., et al. (2012). Generic patient-reported outcomes in child health research: A review of conceptual content using World Health Organization definitions. *Developmental Medicine and Child Neurology*, 54(12), 1085–1095.
4. Williams, J., Wake, M., Hesketh, K., Maher, E., & Waters, E. (2005). Health-related quality of life of overweight and obese children. *The Journal of the American Medical Association*, 293(1), 70–76.
5. World Health Organization. (1998). *WHOQOL user manual*. Geneva: World Health Organization.
6. Fayed, N., & Kerr, E. N. (2009). Comparing quality of life scales in childhood epilepsy: What's in the measures? *International Journal of Disability, Community & Rehabilitation*, 8(3), 1–13.
7. Schiariti, V., Fayed, N., Cieza, A., Klassen, A., et al. (2011). Content comparison of health-related quality of life measures for cerebral palsy based on the international classification of functioning. *Disability and Rehabilitation*, 33(15–16), 1330–1339.
8. U.S. Food and Drug Administration. (2009). Patient reported outcome measures: Use in medical product development to support labeling claims. Resource Document. Available at: <http://www.fda.gov/downloads/Drugs/GuidancecomplianceRegulatoryinformation/guidelines/UCM193282.pdf>. Accessed on January 2, 2013.
9. Heritage, J. (2002). The limits of questioning: Negative interrogatives and hostile question content. *Journal of Pragmatics*, 34(10–11), 1427–1446.
10. Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, 54(2), 93–105.
11. Ronen, G. M., Fayed, N., & Rosenbaum, P. L. (2011). Outcomes in pediatric neurology: A review of conceptual issues and recommendations The 2010 Ronnie Mac Keith lecture. *Developmental Medicine and Child Neurology*, 53(4), 305–312.

12. Davis, E. (2006). Paediatric quality of life instruments: A review of the impact of the conceptual framework on outcomes. *Developmental Medicine and Child Neurology*, 48(04), 311–318.
13. De Civita, M. (2005). Evaluating health-related quality-of-life studies in paediatric populations: Some conceptual, methodological and developmental considerations and recent applications. *PharmacoEconomics*, 23(7), 659–685.
14. Ravens-Sieberer, U., Erhart, M., Wille, N., Wetzel, R., Nickel, J., & Bullinger, M. (2006). Generic health-related quality-of-life assessment in children and adolescents: Methodological considerations. *PharmacoEconomics*, 24(12), 1199–1220.
15. Clarke, S., & Eiser, C. (2004). The measurement of health-related quality of life (QOL) in paediatric clinical trials: A systematic review. *Health and Quality of Life Outcomes*, 2, 66.
16. Eiser, C., & Morse, R. (2001). A review of measures of quality of life for children with chronic illness. *Archives of Disease in Childhood*, 84(3), 205–211.
17. Harding, L. (2001). Children's quality of life assessments: A review of generic and health related quality of life measures completed by children and adolescents. *Clinical Psychology & Psychotherapy*, 8(2), 79–96.
18. Matza, L. S., Swenson, A. R., Flood, E. M., Secnik, K., & Leidy, N. K. (2004). Assessment of health-related quality of life in children: A review of conceptual, methodological, and regulatory issues. *Value in health*, 7(1), 79–92.
19. Solans, M., Pane, S., Estrada, M. D., Serra-Sutton, V., Berra, S., Herdman, M., et al. (2008). Health-related quality of life measurement in children and adolescents: A systematic review of generic and disease-specific instruments. *Value in health*, 11(4), 742–764.
20. Rajmil, L., Herdman, M., Fernandez de Sanmamed, M. J., Detmar, S., Bruil, J., Ravens-Sieberer, U., et al. (2004). Generic health-related quality of life instruments in children and adolescents: A qualitative analysis of content. *Journal of Adolescent Health*, 34(1), 37–45.
21. Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
22. Giraudeau, B., & Mary, J. Y. (2001). Planning a reproducibility study: How many subjects and how many replicates per subject for an expected width of the 95 per cent confidence interval of the intraclass correlation coefficient. *Statistics in Medicine*, 20(21), 3205–3214.
23. Varni, J. W., Katz, E. R., Seid, M., Quiggins, D. J. L., & Friedman-Bender, A. (1998). The pediatric cancer quality of life inventory-32 (PCQL-32): I. Reliability and validity. *Cancer*, 82(6), 1184–1196.
24. Varni, J. W., Katz, E. R., Seid, M., Quiggins, D. J. L., Friedman-Bender, A., et al. (1998). The pediatric cancer quality of life inventory (PCQL): I. Instrument development, descriptive statistics, and cross-informant variance. *Journal of Behavioral Medicine*, 21(2), 179–204.
25. Varni, J. W., Rode, C. A., Seid, M., Katz, E. R., Friedman-Bender, A., & Quiggins, D. J. L. (1999). The pediatric cancer quality of life inventory-32 (PCQL-32): II. Feasibility and range of measurement. *Journal of Behavioral Medicine*, 22, 397–406.
26. Varni, J. W. (2013). *The PedsQL measurement model for the pediatric quality of life inventory*. <http://www.PedsQL.org/>. Accessed on May 3, 2013.
27. Varni, J. W., Seid, M., & Kurtin, P. S. (2001). PedsQL 4.0: Reliability and validity of the pediatric quality of life inventory version 4.0 generic core scales in healthy and patient populations. *Medical Care*, 39(8), 800–812.
28. Varni, J. W., Burwinkle, T. M., Seid, M., & Skarr, D. (2003). The PedsQL 4.0 as a pediatric population health measure: Feasibility, reliability, and validity. *Ambulatory Pediatrics*, 3(6), 329–341.
29. Kolotkin, R. L., Head, S., & Brookhart, A. (1997). Construct validity of the impact of weight on quality of life questionnaire. *Obesity Research*, 5(5), 434–441.
30. Kolotkin, R. L., Head, S., Hamilton, M., & Tse, C. K. (1995). Assessing impact of weight on quality of life. *Obesity Research*, 3(1), 49–56.
31. Kolotkin, R. L., & Crosby, R. D. (2002). Psychometric evaluation of the impact of weight on quality of life-lite questionnaire (IWQOL-lite) in a community sample. *Quality of Life Research*, 11(2), 157–171.
32. Kolotkin, R. L., Crosby, R. D., Kosloski, K. D., & Williams, G. R. (2001). Development of a brief measure to assess quality of life in obesity. *Obesity Research*, 9(20), 102–111.
33. Modi, A. C., & Zeller, M. H. (2011). The IWQOL-Kids(©): establishing minimal clinically important difference scores and test-retest reliability. *International Journal of Paediatric Obesity*, 6(2–2), 94–96.
34. Nadeau, K., Kolotkin, R. L., Boex, R., Witten, T., McFann, K. K., Zeitler, P., et al. (2011). Health-related quality of life in adolescents with comorbidities related to obesity. *Journal of Adolescent Health*, 49(1), 90–92.
35. Zeller, M. H., & Modi, A. C. (2009). Development and initial validation of an obesity-specific quality-of-life measure for children: Sizing me up. *Obesity*, 17(6), 1171–1177.
36. Modi, A. C., & Zeller, M. H. (2008). Validation of a parent-proxy, obesity-specific quality-of-life measure: Sizing them up. *Obesity*, 16(12), 2624–2633.
37. Ravens-Sieberer, U., Bullinger M. (2000). KINDLE^R questionnaire for measuring health-related quality of life in children and adolescents revised version manual.
38. Ravens-Sieberer, U., & Bullinger, M. (1998). Assessing health-related quality of life in chronically ill children with the German KINDL: First psychometric and content analytical results. *Quality of Life Research*, 7(5), 399–407.
39. Wille, N., Bullinger, M., Holl, R., Hoffmeister, U., Mann, R., Goldapp, C., et al. (2010). Health-related quality of life in overweight and obese youths: Results of a multicenter study. *Health and Quality of Life Outcomes*, 8, 36.
40. Morales, L. S., Edwards, T. C., Flores, Y., Barr, L., & Patrick, D. L. (2011). Measurement properties of a multicultural weight-specific quality-of-life instrument for children and adolescents. *Quality of Life Research*, 20(2), 215–224.
41. Patrick, D.L., Edwards, T.C., Skalicky, A. (2010). Youth quality of life instrument-weight module (YQOL-W). User manual and interpretation guide & scoring programs. (1st Ed.). Seattle, Washington.
42. WHOQOL Group. (1994). The development of the world health organization quality of life assessment instrument (WHOQOL). In W. Kuyken & J. Orley (Eds.), *Quality of life assessment: International perspectives* (pp. 41–57). Berlin: Springer.
43. Fayed, N. (2011). Content issues in child health status and quality of life instruments: Addressing the challenges with new methods (Doctor of Philosophy, McMaster University). Open Access Dissertations and Theses Paper, 6066.
44. Colver, A. (2012). Quality of life. In A. Majnemer (Ed.), *Measures for children with developmental disability framed by the ICF-CY* (1st ed.). Cambridge: Mac Keith Press.
45. Wright, V. et al. (under-review). Parent's thoughts on the PEDI, Developmental Medicine Child Neurology.