




Development and Validation of New BODY-Q Scales Measuring Expectations, Eating Behavior, Distress, Symptoms, and Work Life in 4004 Adults From 4 Countries

Claire E. E. de Vries^{1,2}  · Danny Mou¹ · Lotte Poulsen³ · Trisia Breitkopf⁴ · Dennis J. S. Makarawung⁵ · Marinus J. Wiezer⁵ · Ruben N. van Veen² · Maarten M. Hoogbergen⁶ · Jens A. Sorensen³ · Ronald S. L. Liem^{7,8} · Simon W. Nienhuijs⁹ · Ali Tavakkoli¹ · Andrea L. Pusic¹ · Anne F. Klassen⁴

Received: 5 August 2020 / Revised: 28 April 2021 / Accepted: 5 May 2021 / Published online: 27 May 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Purpose The BODY-Q is a rigorously developed patient-reported outcome measure (PROM) for patients seeking treatment for obesity and body contouring surgery. A limitation of the uptake of the BODY-Q in weight management treatments is the absence of scales designed to measure eating-specific concerns. We aimed to develop and validate 5 new BODY-Q scales measuring weight loss expectations, eating behaviors, distress, symptoms, and work life.

Material and Methods In phase 1 (qualitative), patient and expert input was used to develop and refine the new BODY-Q scales. In phase 2 (quantitative), the scales were field-tested in bariatric and weight management clinics in the United States (US), The Netherlands, and Denmark between June 2019 and January 2020. Data were also collected in the US and Canada in September 2019 through a crowdsourcing platform. Rasch measurement theory (RMT) analysis was used for item reduction and to examine reliability and validity.

Results The new BODY-Q scales were refined through qualitative input from 17 patients and 20 experts (phase 1) and field-tested in 4004 participants (phase 2). All items showed ordered thresholds and good fit to the Rasch model. The RMT analysis provided evidence of reliability, with PSI values ≥ 0.72 , Cronbach alpha values ≥ 0.83 , and test-retest values ≥ 0.79 . Better scores on 4 scales (exception expectations scale) correlated with lower BMI, with the strongest correlation between the eating-related distress scale scores and BMI ($r = -0.249$, $P < 0.001$).

Conclusion The new BODY-Q scales can be used in research and clinical practice to assess weight loss treatments from the patient perspective.

✉ Claire E. E. de Vries
devries.cee@gmail.com

¹ Department of Surgery, Brigham and Women's Hospital, 75 Francis St, Boston, MA 02115, USA

² Department of Surgery, OLVG, Amsterdam, The Netherlands

³ Department of Plastic Surgery, Odense University Hospital, Odense, Denmark

⁴ Department of Pediatrics, McMaster University, Hamilton, Ontario, Canada

⁵ Department of Surgery, St. Antonius Hospital, Nieuwegein, The Netherlands

⁶ Department of Plastic and Reconstructive Surgery, Catharina Ziekenhuis, Eindhoven, The Netherlands

⁷ Department of Surgery, Groene Hart Hospital, Gouda, The Netherlands

⁸ Dutch Obesity Clinic, The Hague, The Netherlands

⁹ Department of Surgery, Catharina Ziekenhuis, Eindhoven, The Netherlands

Keywords Bariatric surgery · Weight management · Patient-reported outcome measures · Patient-centered care · Health-related quality of life

Introduction

Obesity is a growing public health concern linked to a range of health problems, including diabetes, cardiovascular disease, depression, and some forms of cancer [1–3]. Since 1975, worldwide obesity has more than tripled, with over 650 million people living with obesity today [4]. Given the rising obesity trends, there has been an increasing demand for obesity treatments [5, 6]. Interventions to lose weight include lifestyle, medical, or surgical weight loss management, each with different risks and benefits.

To facilitate the delivery of evidence-based healthcare, a comprehensive understanding of the full range of outcomes associated with weight loss treatments is essential. Although the clinical benefits of various weight loss treatments are well established [7–13], the outcomes from the patient's perspective have been poorly described [14, 15]. Weight loss intervention, particularly bariatric surgery, can be a life-changing event that significantly impacts patients' health and well-being, but outcomes from the patient perspective are not systematically measured in a rigorous way. The collection of patient-reported outcomes (PRO) data provides a means to measure outcomes most relevant to patients undergoing weight loss treatment. The patient perspective is best measured by means of PRO measures (PROMs). In comparative effectiveness research, the use of PROMs could support an evidence-approach to selecting optimal weight loss treatments [16].

A recent systematic review of PROMs in bariatric surgery highlighted that the BODY-Q was the most rigorously developed and validated PROM for this population [17]. The BODY-Q was developed in a mixed-methods study that involved a literature review, qualitative and cognitive interviews with patients, and input from experts [18]. The BODY-Q measures outcomes related to health-related quality of life (HRQOL), appearance, and experience of healthcare [19]. The BODY-Q is comprised of a set of independently functioning scales, which provides the flexibility of adding new scales to assess additional concepts of interest as they are identified [20, 21]. A gap in the BODY-Q is the lack of scales for specific eating-related concepts that are particularly relevant to weight loss treatment. In this study, we developed and field-tested 5 new BODY-Q scales that measure the following PROs not adequately covered by existing PROMs: weight loss expectations, eating behaviors, eating-related distress, eating-related symptoms, and work life.

Methods

The new BODY-Q scales were developed using a multiphase iterative mixed-methods approach that has been described in detail elsewhere [22]. This paper describes phases I (qualitative research) and II (quantitative research).

Phase I: Qualitative Research

To develop the new BODY-Q scales, general and specific codes related to eating from the original set of 63 patient interview transcripts [23] were reexamined. A set of items covering the following concepts were created using the lowest possible grade reading level: expectations, eating behavior, eating-related distress, eating-related symptoms, and work life. For each scale, instructions, a time frame for reporting, and response options were developed.

Local research ethics board approval at the coordinating center in Canada (Hamilton Integrated Research Ethics Board) was obtained prior to starting the study. Participants provided verbal consent to participate in the study at the beginning of recorded phone interviews.

To refine the scales and to establish content validity for people with obesity, we recruited participants who indicated they were willing to continue to be involved in BODY-Q research from the original BODY-Q field-test study [18]. These participants were sent an email which included an invitation to participate in a cognitive interview to review the new BODY-Q scales. The first 10 to respond to the invitation were interviewed in September and October 2018. In addition, 7 new participants were recruited from St Joseph's Healthcare bariatric program in Hamilton (Canada) between November 2018 and January 2019.

Cognitive interviews were conducted by an experienced qualitative interviewer. Participants were asked to provide feedback on the instructions, items, and response options, and to identify any items that were confusing or not relevant. Participants were also asked to identify any missing content. Interviews were audio-recorded, transcribed, and coded. All data related to the items, response options, and instructions was transferred to a Microsoft Excel worksheet for analysis. The scales were revised iteratively throughout the process.

Input by experts in the area of bariatric surgery and weight loss management was obtained. The experts were chosen across the 4 countries based on their extensive experience in bariatric surgery, weight loss management, bariatric nutrition, and/or PROMs development. The experts were sent the new BODY-Q scales via emailed or shown the scales in person. Expert feedback was provided on the scales' instructions, items, and response options. Experts were also asked to provide feedback to determine if all clinically important issues from the perspective of experts were included in the scales.

After the patient and expert input phase, the new BODY-Q scales were translated into Dutch and Danish in accordance with recommended guidelines [24, 25]. The translation involved two independent forward translations, one backward translation, expert feedback, and cognitive debriefing interviews with patients undergoing weight loss treatment. Feedback from experts and patients was used to revise and finalize the Dutch and Danish translation of the scales.

Phase II: Quantitative Research

The field-test study of the new BODY-Q scales took place between June 2019 and January 2020. This study was conducted in accordance with the Handbook for Good Clinical Research Practice of the World Health Organization and the

Declaration of Helsinki principles and was approved by the regional and local institutional review boards or data protection agency (Brigham and Women's Hospital, United States (US); Medical Research Ethics Committees United, The Netherlands; Danish Data Protection Agency).

The new BODY-Q scales were included in three cohort studies with bariatric and weight loss management patients aged 18 years and older in the US (Brigham and Women's Hospital), The Netherlands (OLVG, Amsterdam; St. Antonius Ziekenhuis, Nieuwegein; Catharina Hospital, Eindhoven), and Denmark (Odense University Hospital and Hospital of Southwest Jutland), and in a web-based survey using the online crowdworking platform Prolific Academic (www.prolific.co) that included participants from the US and Canada. After participants gave (online) consent, data were collected either face-to-face using iPads in the outpatient clinic, with data entered into a secure web-based application (Research Electronic Data Capture (REDCap);US) [26], or by email including a URL that linked directly into a secure web-based application (REDCap or Castor EDC; The Netherlands, Denmark, and Prolific) [26, 27]. The surveys included demographic questions and the 5 new BODY-Q scales. Participants were sent up to 3 email reminders, spaced by 7 days. At the end of the questionnaire, participants from the cohort studies in the US and The Netherlands were invited to participate in a test-retest study. Those who agreed were sent an email 1 week after their appointment with a URL link to the questionnaire, and up to two email reminders were sent spaced by 7 days. Branching logic was used to ensure that only participants who worked at a job with co-workers in the past 3 months completed the work life scale and that only participants who were pre-bariatric surgery, or had only their first appointment in the weight management clinic, completed the expectation scale.

Prolific pays participants at minimum \$6.50 per hour to complete a survey. Participants took between 20 and 30 min to complete the survey and were rewarded \$3.50 for a completed survey. Participants were informed about the payment before they agreed to participate in the study.

Data Analysis

Patient characteristics were described as the mean \pm standard deviation (SD) or by percentages. Descriptive characteristics were analyzed using SPSS software (IBM SPSS Statistics, version 26.0, IBM Corp). The psychometric analysis involved Rasch measurement theory (RMT) analysis [28] conducted within RUMM 2030 software (RUMM version 2030, RUMM Lab.). The following set of statistical and graphical tests were performed to identify the best subset of items to retain in each scale: threshold maps, item fit statistics, dependency, targeting, Differential Item Functioning (DIF), and Person Separation Index

(PSI) (Supplementary information, Table 1). For the item fit analysis, we amended the sample to 500 to adjust the P-values given the large sample size. The following subgroups were examined for the DIF analysis: sample (Prolific versus clinical), age (18 to 29, 30 to 39, 40 to 49, 50 to 59, 60 plus years), body mass index (BMI) classification (24.9 or less, 25.0 to 29.9, 30 to 34.9, 35.0 to 39.9, 40 plus), language (English, Danish, Dutch), sex (male, female), treatment (bariatric surgery, weight management), and bariatric surgery group (preoperative, first year, 1–2 years, 3 or more years after surgery). The Rasch logit score was used to convert the raw scores for each scale to a score ranging from 0 (worst) to 100 (best). For the test-retest (TRT) data, the intra-class correlation coefficients (ICC) were calculated using the two-way random effects model. Higher ICC values indicate greater test-retest reliability. The proportion of participants with scores at the floor and ceiling were computed. A high proportion of participants with scores at the floor and ceiling can be an indication that a scale is not comprehensive.

For concurrent validity, we included the EQ-5D-5L which has 2 parts [29]. Part 1 is a 5-item questionnaire that measures mobility, self-care, usual activities, pain/discomfort, and anxiety/depression in terms of 5 levels, with level 1 representing no problems and level 5 representing unable to/extreme problems [29]. Part 2 is the EQ-VAS which asks participants to choose a number from 0 to 100 to indicate their current health status, with higher scores indicating better health. The EQ-5D-5L has been validated for use in bariatric surgery, with adequate convergent/divergent validity. In terms of concurrent validity, we expected that the new BODY-Q scales would correlate weakly to moderately with the EQ-5D-5L and the EQ-VAS (the scales measure dissimilar constructs). We also expected that correlations between the eating behavior and eating-related distress scales would be higher than correlations between these two scales and the expectations, work life, and symptoms scales. We also tested three hypotheses. First, we hypothesized that patient characteristics (age, sex, and race) would correlate weakly with the new BODY-Q scales. Second, we hypothesized that worse scores on the new BODY-Q scales would be associated with higher BMI. Third, we hypothesized that participants who were waiting for bariatric surgery compared to participants who had undergone bariatric surgery would report lower mean scores on the eating behavior, eating-related distress, eating-related symptoms, and work life scales.

Depending on normality of the data, Pearson or Spearman correlations were used to examine associations between continuous variables, and t-tests or one-way analysis of variance (ANOVA) or the non-parametric equivalents were used to examine differences between groups. P-values of <0.05 were considered statistically significant for these tests of construct validity.

Results

Qualitative Phase

The new BODY-Q scales were reviewed by 15 females and 2 male patients of whom 12 had undergone bariatric surgery. The mean age of participants was 48 years (range 32–62 years). The scales were also reviewed by 20 experts including 5 bariatric surgeons, 3 psychologists, 1 bariatric physician, 1 bariatric physician assistant, 1 bariatric dietician, 2 bariatric nurses, 2 plastic surgeons, 1 plastic surgery trainee, and 4 physician trainees (PhD and/or resident) from 4 countries (Canada, Denmark, The Netherlands, the US). The [Supplementary Information](#) shows sample quotes of patients during the cognitive interviews (Table 2) and the number of items reviewed by participants and experts in each round that were retained, revised, dropped, and added (Table 3). At the end of the process, the set of scales included 89 items.

Instructions and Response Options

Due to participant feedback and expert input, the instructions of the scales were slightly changed throughout the rounds. In terms of response options, the expectations scale measures how likely (very unlikely, somewhat unlikely, somewhat likely, very likely) participants thought a set of statements would apply to them; the work life scale measures agreement (definitely disagree, somewhat disagree, somewhat agree, definitely agree) to a set of statements. The remaining scales measure frequency (never, sometimes, often, always). Participants generally did not have any problems with the response options of the scales during each round. Some participants wanted to add additional response options such as a neutral option (e.g., neither agree nor disagree), but in RMT, the scoring and interpretation of neutral options are unclear [30]. Additionally, 4–5 response options are considered ideal [31].

Translation

To finalize the field-test version, two items that proved difficult to translate into Dutch and Danish were dropped. After these deletions, the field-test version of the new scales consisted of 87 items.

Quantitative Phase

For the clinical samples, the response rate varied by country, i.e., The Netherlands 62%, Denmark 59%, the US 94%. A total of 4123 participants started the survey, and 4004 (2057 Prolific

and 1947 clinical) completed at least 1 scale and were included in the analysis. Table 1 shows the sample characteristics. The sample included more females (N=2743, 69%), had a mean age of 45 (SD 13), and mean BMI of 31.1 (SD 8.7).

For the RMT analysis, data for the sample of 1038 underweight and normal weight participants were excluded.

The RMT analysis reduced the 87 items to 59 items across 5 scales. Items that were excluded from the scales in the item reduction phase of the RMT analysis are shown in the [Supplementary Information](#) Table 4). Items were mainly excluded due to misfit to the Rasch model. All items in 4 scales had ordered thresholds, demonstrating that respondents could appropriately discriminate amongst response options (see example in [Supplementary Information](#)), Fig. 1). In contrast, 7 of 15 items in the expectations scale evidenced disordered thresholds. The two response options (very unlikely and somewhat unlikely) were combined to one category (unlikely) to simplify the scoring. The RMT analysis used the following rescored items: unlikely, 0; somewhat likely, 1; and very likely, 2. After rescoring, all 15 items had ordered thresholds.

The item fit statistics provided evidence of validity for the 5 scales ([Supplementary Information](#), Table 5). For each scale except the work life scale, the observed data fit the Rasch model, with non-significant overall model fit (Table 2). For item fit statistics, all 59 items had non-significant chi-square P-values after Bonferroni adjustment, and item fit was inside the criteria of ± 2.5 for 22 of the 59 items ([Supplementary Information](#), Table 5). The item residual correlations were greater than 0.20 for 11 pairs of items within 4 scales, indicating some degree of dependency. However, in substest analyses, the correlated items were found to have marginal influence on the reliability of the scales (Table 2).

In terms of targeting, the percentage of participants who scored within the scales' range of measurement was from 76% for the eating-related distress scale to 99% for the eating behavior scale (Table 2). The [Supplementary Information](#) (Fig. 1) shows an example of targeting for the eating behavior scale. The distribution of person measurement and item locations showed that patients with obesity, undergoing medical weight loss, or weight loss surgery were evenly distributed to match all levels of the construct that is measured for each scale, which is reflected by the mirrored distribution of the persons (top half of the graph) and item locations (lower half of the graph) of the scales.

The [Supplementary Information](#) (Table 6) shows the items that showed DIF by the relevant patient characteristic in the adjusted and unadjusted analysis. When the items that evidenced DIF were split by the relevant patient characteristic, Pearson correlations between the original and split person locations were 0.995 or higher showing negligible impact on scoring. This indicates that the total scores on the 5 scales can be compared between subgroups of participants with different patient characteristics.

Table 1 Patient characteristics of the field-test sample by country of recruitment

	United States (N=2648)	Canada (N=72)	The Netherlands (N=377)	Denmark (N=907)	Total (N=4004)
Sample, N (%)					
Prolific	1985 (75)	72 (100)	0 (0)	0 (0)	2057 (51.4)
Clinical	663 (25)	0 (0)	377 (69.2)	908 (100)	1947 (48.6)
Age in years					
Mean (SD)	44 (14)	56 (7)	47 (11)	45 (10)	45 (13.2)
Range	18–83	36–73	20–69	19–71	18–83
Sex, N (%)					
Female	1684 (63.7)	51 (70.8)	288 (81.8)	720 (79.4)	2743 (69.0)
Male	949 (35.9)	21 (29.2)	64 (18.2)	186 (20.5)	1220 (30.7)
Other	12 (0.5)	-	-	1 (0.1)	13 (0.3)
Race, N (%)					
White	2039 (77.0)	67 (93.1)	328 (87.0)	-	2434 (78.6)
Other	609 (23)	5 (6.9)	49 (13.0)	-	663 (21.4)
BMI in kg/m²					
Mean (SD)	30.4 (8.8)	29.6 (7.2)	30.1 (5.6)	33.6 (8.8)	31.1 (8.7)
Range	16.2–82.5	19.1–57.0	19.6–54.4	17.1–80.0	16.2–82.5
BMI classification group, N (%)					
Underweight and normal weight	842 (31.8)	22 (30.6)	51 (14.5)	123 (13.6)	1038 (26.1)
Overweight	671 (25.4)	24 (33.3)	151 (43.0)	274 (30.2)	1120 (28.2)
Obese class I	450 (17.0)	15 (20.8)	88 (25.1)	174 (19.2)	727 (18.3)
Obese class II	317 (12.0)	2 (2.8)	42 (12.0)	132 (14.6)	493 (12.4)
Obese class III	365 (13.8)	9 (12.5)	19 (5.4)	203 (22.4)	596 (15.0)
Comorbidities, N (%)					
Diabetes	278 (10.5)	9 (12.5)	27 (7.2)	140 (15.4)	454 (11.3)
Hypertension	600 (22.7)	25 (34.7)	60 (15.9)	215 (23.7)	900 (22.5)
Hyperlipidemia	335 (12.7)	11 (15.3)	22 (5.8)	150 (16.5)	518 (12.9)
Obstructive sleep apnea	344 (13.0)	11 (15.3)	58 (15.4)	156 (17.2)	569 (14.2)
Osteoarthritic disease	340 (12.8)	16 (22.2)	73 (19.4)	185 (20.4)	614 (15.3)
Cardiovascular disease	74 (2.8)	-	16 (4.2)	42 (4.6)	132 (3.3)
Reflux disease	504 (19.0)	15 (20.8)	25 (6.6)	147 (16.2)	691 (17.3)
None	1362 (51.4)	26 (36.1)	209 (55.4)	373 (41.1)	1970 (49.2)
Type of patients, N (%)					
Bariatric surgery	606 (91.4)	-	368 (97.6)	811 (89.4)	1785 (91.7)
Weight loss management	57 (8.6)	-	9 (2.4)	96 (10.6)	162 (8.3)
Bariatric surgery type, N (%)					
Gastric banding	40 (8.9)	-	-	9 (1.3)	49 (3.3)
Gastric bypass	129 (28.7)	-	228 (67.5)	562 (81.6)	919 (21.3)
Sleeve gastrectomy	276 (61.5)	-	108 (32)	115 (16.7)	499 (11.6)
Other	4 (0.9)	-	2 (0.6)	3 (0.4)	9 (0.2)
Bariatric surgery group, N (%)					
Preoperative	178 (28.6)	-	19 (5.4)	117 (14.4)	314 (17.6)
First year after surgery	193 (31.0)	-	188 (53.7)	98 (12.1)	479 (26.9)
1–2 years after surgery	91 (14.6)	-	130 (37.1)	226 (27.9)	447 (25.1)
3 or more years after surgery	161 (25.8)	-	13 (3.7)	369 (45.6)	543 (30.5)
Employment, N (%)					
Employed, full-time	1257 (47.4)	22 (30.6)	129 (34.2)	389 (42.9)	1797 (44.9)
Employed, part-time	341 (12.9)	13 (18.1)	133 (35.3)	152 (16.8)	639 (16.0)
Homemaker	161 (6.1)	5 (6.9)	25 (6.6)	11 (1.2)	202 (5.1)
Unemployed	226 (8.5)	7 (9.7)	23 (6.2)	69 (7.6)	335 (8.0)

Table 1 (continued)

	United States (N=2648)	Canada (N=72)	The Netherlands (N=377)	Denmark (N=907)	Total (N=4004)
Sick leave/disabled	136 (5.1)	3 (4.2)	38 (10.1)	35 (3.9)	212 (5.3)
Student	177 (6.7)	-	6 (1.6)	49 (5.4)	232 (5.8)
Other	350 (13.2)	22 (30.6)	23 (6.1)	202 (22.3)	587 (14.7)
Assessments for each scale, N (%)					
Expectations	188 (7.1)	-	11 (2.9)	112 (12.3)	311 (7.8)
Eating behavior	2639 (99.7)	72 (100)	376 (99.7)	906 (99.9)	3993 (99.7)
Eating-related distress	2623 (99.1)	72 (100)	369 (97.9)	892 (98.3)	3956 (98.8)
Work life	1607 (60.7)	32 (44.4)	256 (67.9)	620 (68.4)	2515 (62.8)
Eating-related symptoms	2624 (99.1)	72 (100)	368 (97.6)	894 (98.6)	3958 (98.9)

The RMT analysis provided evidence of reliability for the 5 scales, with PSI values with and without extremes ≥ 0.70 and Cronbach alpha values ≥ 0.81 (Table 2). For the TRT study, the response rate was 70% in The Netherlands and 43% in the US. A total of 303 participants completed a TRT 1 week after the initial assessment. For TRT, the ICC values, shown in Table 2, were sufficient, with values ≥ 0.79 for all 5 scales.

After transforming the scores from 0 to 100, the percentage of participants who scored on the floor (lowest score)/ceiling (highest score) for each scale were as follows: expectations = 1.3/17.0, eating behavior = 0.1/1.3, eating-related distress = 0.2/26.4, eating-related symptoms = 11.1/1.3, and work life = 0.3/21.8.

The results for concurrent validity are shown in Table 3. The expectations, eating behavior, and eating-related symptoms scales correlated weakly with the EQ-5D-5L and the EQ-VAS. The eating-related distress scale correlated moderately with anxiety/depression and the EQ-VAS, and the work life scale correlated moderately with the EQ-5D-5L items (exception self-care) and the EQ-VAS. The scales measuring eating behavior and eating-related distress correlated more

strongly with each other than with scales with unrelated constructs (expectations, work life, and eating-related symptoms). The expectations scale only correlated with the eating behavior scale. The majority of the results were in concordance with the three hypotheses. Patient characteristics (age, sex, and race) correlated weakly with the new BODY-Q scales. Four of the new BODY-Q scales (exception expectations) correlated with BMI, with the strongest correlation between the eating-related distress scale and BMI ($r = -0.249$, $P < 0.001$) (Table 3). Mean scores differed significantly across BMI classification group for all scales but the expectations scale, with lower (worse) scores for participants with class 1 to 3 obesity ($P < 0.001$ on ANOVA) (Fig. 1). Only the eating behavior scale correlated with time since surgery ($r = -0.141$, $P = 0.01$). Mean scores differed for the bariatric surgery group, with the lowest scores for preoperative patients ($P < 0.001$ on ANOVA). Only for the eating-related symptoms scale, the preoperative patients, who have not had bariatric surgery, scored better.

Table 2 Scale level results

Scale	% scored	# for RMT	# in RMT	Chi-square	DF	P-value	PSI +extr	PSI -extr	Cronbach α +extr	Cronbach α -extr	ICC	PerC <5%
Expectations	80.1	311	249	43.35	30	0.05	0.83	0.82	0.92	0.87	-	3.64
Eating behavior	98.5	2925	2882	84.25	81	0.380	0.84	0.83	0.85	0.84	0.84	5.51
Eating-related distress	75.9	2199	2899	63.04	60	0.370	0.86	0.89	0.95	0.93	0.88	3.46
Work life	79.2	1863	1476	109.68	80	0.016	0.80	0.83	0.91	0.88	0.79	5.36
Eating-related symptoms	92.9	1674	1555	123.64	135	0.749	0.72	0.70	0.83	0.81	0.86	1.68

RMT Rasch measurement theory, DF degrees of freedom, PSI person separation index, ICC intra-class correlation coefficient, *extr* extremes
PerC <5% top bottom indicates unidimensionality of the scale

There was not enough data in this study to calculate the ICC (TRT value) for the expectations scale

Table 3 Correlations between BODY-Q scales, EQ-5D-5L, and patient characteristics

Scale	Expectations	Eating behavior	Eating-related distress	Work life	Eating-related symptoms
Expectations	1.00				
Eating behavior	0.17**	1.00			
Eating-related distress	0.01	0.53**	1.00		
Work life	0.07	0.41**	0.48**	1.00	
Eating-related symptoms	0.04	0.15**	0.36**	0.19**	1.00
Mobility	-0.02	-0.12**	-0.24**	-0.36**	-0.16**
Self-care	-0.03	-0.14**	-0.21**	-0.24**	-0.09**
Usual activities	-0.02	-0.18**	-0.28**	-0.37**	-0.20**
Pain/discomfort	-0.08	-0.16**	-0.29**	-0.37**	-0.29**
Anxiety/depression	-0.03	-0.28**	-0.43**	-0.49**	-0.24**
EQ-VAS	0.03	0.29**	0.41**	0.55**	0.20**
BMI	0.02	-0.13**	-0.25**	-0.23**	-0.10**
Age	-0.29**	0.11**	0.12**	0.14**	0.002
Sex	-0.07	-0.11**	0.07**	-0.05*	0.16**
Race	-0.19**	-0.05**	-0.02	-0.02	0.01
Time since surgery		-0.14*	0.05	0.03	0.07

Patient characteristics: BMI, age and time since surgery = continuous variables; sex and race (White vs other) = dichotomous variables. Pearson correlations except for EQ-VAS and time since surgery used Spearman’s rho due to skewed data

*Correlation is significant at the 0.05 level (two-tailed)

**Correlation is significant at the 0.01 level (two-tailed)

Discussion

The new BODY-Q scales further augment the range of concepts important to patients covered by the BODY-Q. These new scales were developed using existing qualitative interview data from the original BODY-Q and with new input from patients and experts to ensure that the content of the scales was comprehensive and relevant to patients seeking or undergoing weight loss interventions. A modern psychometric approach, RMT analysis, was used to refine the content of the new BODY-Q scales while staying grounded in the

experience of patients and experts. The detailed RMT analysis provided evidence of validity and reliability of the new BODY-Q scales in a large international, multilingual sample. Different types of samples were included as well as different languages to increase generalizability of the new scales for international uptake. Importantly, we found that the scales worked the same across patients who differed by sample type, age, sex, and language. The scales were appropriately targeted to patients with obesity or undergoing weight loss treatment. The new BODY-Q scales can be used in national and international comparative effectiveness studies of various weight loss interventions.

Understanding patients’ expectations of bariatric surgery is critically important to ensuring effective preoperative informed consent [32]. Failing to establish realistic expectations can result in inappropriate patient selection and significant postoperative distress. The expectations scale of the BODY-Q directly assesses this important metric. Realistic patient expectations have been found to be important for better postoperative satisfaction and HRQOL outcomes [33, 34]. An improved understanding of the expectations of patients may help clinicians better educate and select patients who are most appropriate for weight loss treatment [35]. Future research is needed to examine how patients’ expectations of weight loss treatment are related to outcomes after treatment.

Eating plays a critical role in weight loss treatment. One of the main goals of weight loss treatment is to achieve better

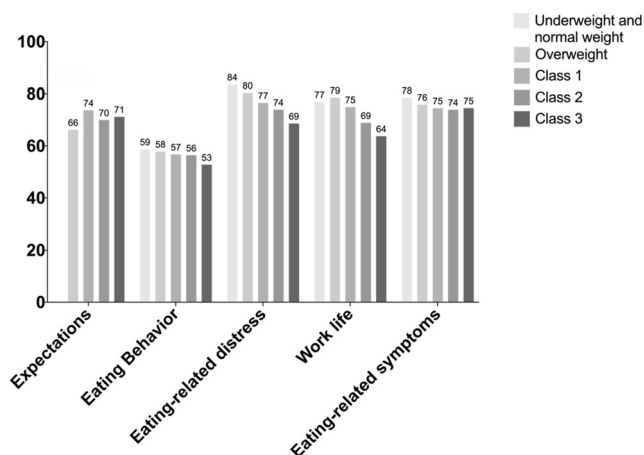


Fig. 1 Mean scores for the 5 new BODY-Q scales by BMI classification

eating habits [36–39]. For example, bariatric surgery alters the patient's anatomy to encourage smaller portion meals. Poor postoperative eating habits place patients at risk for weight regain [38, 40, 41]. Thus, identifying predictors for poor eating behaviors may be valuable in the preoperative setting to identify adequate bariatric surgery candidates. Additionally, patients often have strong emotions attached to eating. Achieving better control of eating may lead to positive feelings, while failing to change eating habits may lead to negative feelings, such as frustration [39, 40]. Other constructs associated with eating are also challenged after bariatric surgery. Some patients can experience unpleasant side effects following bariatric surgery, including nausea, vomiting, and diarrhea [36]. The new eating scales of the BODY-Q provide a means of measuring a range of eating behaviors, eating-related distress, as well as symptoms associated with eating. These scales are the first independently functioning scales specifically targeted to patients living with obesity and to patients undergoing weight loss treatment. The scales allow for quantitative measurements of important eating-related outcomes and can be used in research and individual clinical care where the collection and use of PRO data for weight loss treatment would be informative. Further application of these scales may give rise to better understanding of differences in eating behavior, eating-related distress, and eating-related problems following different surgical procedures, which can be used to inform decision-making.

Living with obesity can negatively impact one's function at work, which can compromise one's overall HRQOL. Patients living with obesity may feel stigmatized at work and have fewer opportunities than their colleagues, and improving work opportunities is an important factor to seek weight loss treatment [36, 42]. Patients who underwent bariatric surgery endorsed improvement of work opportunities, effectiveness at work, and increased recognition by their colleagues [43–47]. The BODY-Q work life scale is a rigorously developed PROM that assesses these obesity-associated work life experiences that are closely related with health-related quality of life.

There were some limitations to this study. First, we used an online platform (Prolific) as part of our research study to recruit participants. While the advantage of Prolific is that a large sample of participants can be recruited in a short period of time, a disadvantage is that the sample included more Caucasian participants than the US clinical sample, and the number of Canadian participants was limited. In addition, although we collected BMI information for the Prolific sample, questions about bariatric surgery or other weight loss treatments and waist circumference or body composition were not included in the survey. Furthermore, we did not collect data on medication that could have affected for example appetite or mood. Second, given that smaller numbers of

participants were pre-bariatric, we did not have enough data to examine DIF and test-retest reliability for the expectations scale. Third, the study data were cross-sectional, which is appropriate for PROM development. However, longitudinal studies using the new scales are necessary to assess ability to measure change and to determine minimal important difference.

Conclusion

The new BODY-Q scales enable rigorous assessment of outcomes that are particularly relevant to patients who live with obesity or undergo weight loss treatment. These scales equip clinicians and researchers with measurement instruments that can promote high-quality, patient-centered care.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11695-021-05462-2>.

Conflict of Interest The authors have no competing interests to report. The BODY-Q is jointly owned by Memorial Sloan-Kettering Cancer Center, Brigham and Women's Hospital, and McMaster University. Drs. Pusic and Klassen are co-developers of the BODY-Q and, as such, receive a share of any license revenues based on the inventor sharing policies of these three institutions.

Funding The phase I qualitative phase of the BODY-Q study was funded by a research grant received from the Plastic Surgery Foundation (120593).

References

1. Afshin A, Forouzanfar MH, Reitsma MB, et al. Health effects of overweight and obesity in 195 countries over 25 years. *N Engl J Med*. Massachusetts Medical Society. 2017;377:13–27.
2. Must A, McKeown NM. The disease burden associated with overweight and obesity. *Endotext*. 2000
3. Wang YC, McPherson K, Marsh T, et al. Health and economic burden of the projected obesity trends in the USA and the UK. *Lancet*. Lancet Publishing Group; 2011. p. 815–25.
4. World Health Organization Obesity and overweight, 2018.
5. Dietz WH, Baur LA, Hall K, et al. Management of obesity: improvement of health-care training and systems for prevention and care. *Lancet*. Lancet Publishing Group; 2015. p. 2521–33.
6. Heymsfield SB, Wadden TA. Mechanisms, pathophysiology, and management of obesity. *N. Engl. J. Med*. Massachusetts Medical Society; 2017. p. 254–66.
7. Colquitt JL, Pickett K, Loveman E, et al. Surgery for weight loss in adults (Cochrane Review). *Cochrane Collab*. 2014;244
8. Schauer PR, Bhatt DL, Kirwan JP, et al. Bariatric surgery versus intensive medical therapy for diabetes - 5-year outcomes. *N Engl J Med*. 2017;376:641–51.
9. Yanovski SZ, Yanovski JA. Long-term drug treatment for obesity: a systematic and clinical review. *JAMA*. 2014. p. 74–86.

10. Wadden TA, Volger S, Sarwer DB, et al. A two-year randomized trial of obesity treatment in primary care practice. *N Engl J Med. Massachusetts Medical Society.* 2011;365:1969–79.
11. Wadden TA, Berkowitz RI, Womble LG, et al. Randomized trial of lifestyle modification and pharmacotherapy for obesity. *N Engl J Med.* 2005;353:2111–20.
12. Appel LJ, Clark JM, Yeh HC, et al. Comparative effectiveness of weight-loss interventions in clinical practice. *N Engl J Med. Massachusetts Medical Society.* 2011;365:1959–68.
13. Khera R, Murad MH, Chandar AK, et al. Association of pharmacological treatments for obesity with weight loss and adverse events a systematic review and meta-analysis. *JAMA.* 2016;315:2424–34.
14. Coulman KD, Abdelrahman T, Owen-Smith A, et al. Patient-reported outcomes in bariatric surgery: a systematic review of standards of reporting. *Obes Rev.* 2013;14:707–20.
15. Lindekilde N, Gladstone BP, Lübeck M, et al. The impact of bariatric surgery on quality of life: a systematic review and meta-analysis. *Obes Rev.* 2015;16:639–51.
16. Patrick DL, Burke LB, Powers JH, et al. Patient-reported outcomes to support medical product labeling claims: FDA perspective. *Value Heal. Blackwell Publishing Inc.* 2007. p. S125–37.
17. de Vries CEE, Kalf MC, Prinsen CAC, et al. Recommendations on the most suitable quality-of-life measurement instruments for bariatric and body contouring surgery: a systematic review. *Obes Rev.* 2018;1395–411.
18. Klassen AF, Cano SJ, Alderman A, et al. The BODY-Q: a patient-reported outcome instrument for weight loss and body contouring treatments. *Plast Reconstr Surg - Glob Open.* 2016;4:e679.
19. Klassen AF, Cano SJ, Scott A, et al. Assessing outcomes in body contouring. *Clin Plast Surg.* 2014;645–54.
20. Poulsen L, Pusic A, Robson S, et al. The BODY-Q Stretch Marks Scale: a development and validation study. *Aesthetic Surg J. Oxford Academic.* 2018;38:990–7.
21. Klassen AF, Kaur M, Poulsen L, et al. Development of the BODY-Q chest module evaluating outcomes following chest contouring surgery. *Plast Reconstr Surg.* 2018;142:1600–8.
22. Riff K W Y W, Tsangaris E, Goodacre T, et al. International multi-phase mixed methods study protocol to develop a cross-cultural patient-reported outcome instrument for children and young adults with cleft lip and/or palate (CLEFT-Q). *BMJ Open.* 2017;7:1–9.
23. Klassen AF, Cano SJ, Scott A, et al. Satisfaction and quality-of-life issues in body contouring surgery patients: a qualitative study. *Obes Surg.* 2012;22:1527–34.
24. Wild D, Grove A, Martin M, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Heal. Blackwell Publishing Inc.* 2005;8:94–104.
25. Organization WHO. Process of translation and adaptation of instruments.
26. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)-a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform. NIH Public Access.* 2009;42:377–81.
27. Castor EDC. 2019. Available from: <https://castoredc.com>
28. Rasch G. *Studies in mathematical psychology: 1. Probabilistic models for some intelligence and attainment tests.* Danmarks Pædagogiske Institut: Copenhagen; 1960.
29. Rabin R, De Charro F. EQ-5D: A measure of health status from the EuroQol group. *Ann Med. Royal Society of Medicine Press Ltd;* 2001. p. 337–43.
30. Bradley KD, Peabody MR, Akers KS, et al. Rating scales in survey research: using the Rasch model to illustrate the middle category measurement flaw. *Surv Pract.* 2015;8:1–12.
31. Khadka J, Gothwal VK, McAlinden C, et al. The importance of rating scales in measuring patient-reported outcomes. *Health Qual Life Outcomes.* 2012;10:80.
32. Cohn I, Raman J, Sui Z. Patient motivations and expectations prior to bariatric surgery: a qualitative systematic review. *Obes Rev.* 2019;1608–18.
33. Poulsen L, Klassen A, Jhanwar S, et al. Patient expectations of bariatric and body contouring surgery. *Plast Reconstr Surg - Glob Open.* 2016;4:e694.
34. Parretti HM, Hughes CA, Jones LL. “The rollercoaster of follow-up care” after bariatric surgery: a rapid review and qualitative synthesis. *Obes Rev.* 2019;20:88–107.
35. Mou D, Sisodia RC, Castillo-Angeles M, et al. The surgeon’s perceived value of patient-reported outcome measures (PROMs). *Ann Surg.* 2020:1–6.
36. Coulman KD, MacKichan F, Blazeby JM, Owen-Smith A. Patient experiences of outcomes of bariatric surgery: a systematic review and qualitative synthesis. *Obes Rev.* 2017;547–59.
37. Natvik E, Gjengedal E, Råheim M. Totally changed, yet still the same: patients’ lived experiences 5 years beyond bariatric surgery. *Qual Health Res.* 2013;23:1202–14.
38. Engström M, Forsberg A. Wishing for deburdening through a sustainable control after bariatric surgery. *Int J Qual Stud Health Well-being.* 2011;6
39. Coulman KD, MacKichan F, Blazeby JM, et al. Patients’ experiences of life after bariatric surgery and follow-up care: a qualitative study. *BMJ Open. BMJ Publishing Group.* 2020;10:e035013.
40. Ogden J, Avenell S, Ellis G. Negotiating control: patients’ experiences of unsuccessful weight-loss surgery. *Psychol Health.* 2011;26:949–64.
41. Warholm C, Øien AM, Raheim M. The ambivalence of losing weight after bariatric surgery. *Int J Qual Stud Health Well-being.* 2014;9:22876.
42. Wee CC, Davis RB, Jones DB, et al. Sex, race, and the quality of life factors most important to patients’ well-being among those seeking bariatric surgery. *Obes Surg.* 2016;26:1308–16.
43. Mancini A, Borel AL, Coumes S, et al. Bariatric surgery improves the employment rate in people with obesity: 2-year analysis. *Surg Obes Relat Dis.* 2018;14:1700–4.
44. Kantarovich K, Wnuk S, Cassin S, et al. Employment outcomes 2 years after bariatric surgery: relationship to quality of life and psychosocial predictors. *Obes Surg.* 2019;29:2854–61.
45. Sockalingam S, Wnuk S, Kantarovich K, et al. Employment outcomes one year after bariatric surgery: the role of patient and psychosocial factors. *Obes Surg. Springer New York LLC*2015;25: 514–22.
46. Courtney MJ, Mahawar K, Burnell P, et al. Occupational outcomes of obesity surgery—do the employed return to work, and do the unemployed find work? *Obes Surg.* 2018;28:963–9.
47. Sharples AJ, Cheruvu CVN. Systematic review and meta-analysis of occupational outcomes after bariatric surgery. *Obes Surg.* 2017;27:774–81.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.