

Psychometric Validation of the BODY-Q in Danish Patients Undergoing Weight Loss and Body Contouring Surgery

Lotte Poulsen, MD*

Anne Klassen, DPhil, PhD†

Michael Rose, MD‡

Kirsten K. Roessler, DPhil, PhD§

Claus Bogh Juhl, MD, PhD¶

René Klinkby Støvring, MD, PhD||

Stefan Cano, MSc, PhD**

Jens Ahm Sørensen, MD, PhD††

Background: A well-developed patient-reported outcome instrument is needed for use in Danish bariatric and body contouring patients. The BODY-Q is designed to measure changes in important patient outcomes over the entire patient journey, from obesity to post-body contouring surgery. The current study aims to psychometrically validate the BODY-Q for use in Danish patients.

Methods: The process consisted of 3 stages: translation and linguistic validation, field-test, and data analysis. The translation was performed in accordance with the International Society for Pharmacoeconomics and Outcomes Research and World Health Organization guidelines, and field-test data were collected in 4 departments in 2 different hospitals. Field-test data were analyzed using Rasch Measurement Theory.

Results: A total of 495 patients completed the Danish BODY-Q field-test 1–4 times, leading to a total of 681 assessments with an overall response rate at 76%. Cronbach α values were ≥ 0.90 , and person separation index values were in general high. The Rasch Measurement Theory analysis provided broad support for the reliability and validity of the Danish version of the BODY-Q scales. Item fit was outside the criteria for 34 of 138 items, and of these, 21 had a significant chi-square P value after Bonferroni adjustment. Most items (128 of 138) had ordered thresholds, indicating that response options worked as intended.

Conclusion: The Danish version of the BODY-Q is a reliable and valid patient-reported outcome instrument for use in Danish bariatric and body contouring patients. (*Plast Reconstr Surg Glob Open* 2017;5:e1529; doi: 10.1097/GOX.0000000000001529; Published online 20 October 2017.)

INTRODUCTION

In Denmark, between 2003 and 2015, there were approximately 16,400 bariatric procedures performed.^{1,2} Many patients who have bariatric surgery develop excess

skin as a consequence of massive weight loss. Body contouring with removal of excess skin is often needed to fully restore a patient's body image and health-related quality of life (HR-QOL).³

To be able to measure change in HR-QOL along the patient weight loss journey, a well-developed patient-reported outcome (PRO) instrument is needed. Previous literature concerning PRO instruments for bariatric and/or body-contouring surgery has identified the need for a new and comprehensive PRO instrument.^{4–9} In a recent review, Gilmartin et al.¹⁰ also call for a new PRO instrument designed for the patient group and point to the newly published BODY-Q as a possible candidate tool.¹¹

From the *Department of Plastic Surgery, Odense University Hospital, Odense, Denmark; †Department of Plastic Surgery, OPEN, Odense Patient data Explorative Network, Odense University Hospital, Odense, Denmark; ‡Department of Paediatrics, McMaster University, Hamilton, Canada; §Department of Plastic Surgery, Hospital of Southwest Jutland, Esbjerg, Denmark; ¶Department of Psychology, University of Southern Denmark, Odense, Denmark; ||Department of Endocrinology, Hospital of Southwest Jutland, Esbjerg, Denmark; **Center for Eating Disorders and Department of Endocrinology, Odense University Hospital; ††The Psychiatric Services in Southern Denmark, University of Southern Denmark, Denmark; and ‡‡Stefan Cano, Modus Outcomes, Lechworth Garden City, United Kingdom. Received for publication April 1, 2017; accepted August 23, 2017.

Copyright © 2017 The Authors. Published by Wolters Kluwer Health, Inc. on behalf of The American Society of Plastic Surgeons. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal. DOI: 10.1097/GOX.0000000000001529

Disclosure: The study was funded by a fund received from The Region of Southern Denmark and The Internationalization Fund at Odense University Hospital. The authors have no financial interest to declare in relation to the content of this article. The Article Processing Charge was paid for by the Department of Plastic Surgery, Odense University Hospital, Odense, Denmark.

Supplemental digital content is available for this article. Clickable URL citations appear in the text.

The Moorehead-Ardelt Quality of Life Questionnaire¹² is the PRO instrument that is included in the Danish Bariatric Surgery Database to document change in HR-QOL for Danish patients having bariatric and body-contouring surgery.² We have previously described the limitations of the Moorehead-Ardelt Quality of Life Questionnaire in the Danish system and the need for a more appropriate PRO instrument for use in monitoring outcomes following bariatric and body contouring surgery.^{1,13}

The BODY-Q is a new PRO instrument designed for weight loss and body-contouring treatments. The BODY-Q was specifically designed to measure changes in important patient outcomes over the entire patient journey, from obesity to post-body contouring surgery.^{7,11} The BODY-Q consists of 18 independently functioning scales grouped into 3 main domains: appearance, HR-QOL, and experience of health-care. There is also an obesity-specific symptom checklist; however, it is important to notice that the symptom checklist is not scored as a scale, but rather as a set of independent items measuring symptoms. Internationally recommended guidelines for item generation, item reduction, and psychometric evaluation were followed in the development of the BODY-Q.¹⁴⁻¹⁹ A strength of the BODY-Q is the use of Rasch Measurement Theory (RMT), which contributes to developing interval level measurement scales.^{20,21}

The BODY-Q was developed and field-tested in the United States, Canada, and United Kingdom, which limit its immediate generalization to other non-English-speaking countries. When adapting a PRO instrument for use in another country and/or culture, it is essential to start with an appropriate translation and linguistic validation study, followed by psychometric validation.^{22,23} Due to the lack of a comprehensive and robust PRO instrument for patients who undergo bariatric and body-contouring treatments in Denmark, we translated and evaluated the BODY-Q in a large sample of Danish patients having bariatric and/or body contouring. In this article, we describe the findings from the Danish field-test and psychometric validation study.

METHODS AND MATERIALS

The study was approved by the Danish Data Protection Agency before commencement. Ethics approval was applied for at the Regional Scientific Ethical Committee for Southern Denmark, who found no reason for approval, as our study was an interview and questionnaire survey.

The process consisted of 3 stages described further underneath: translation and linguistic validation, field-test, and data analysis.

Stage I: Translation, Cultural Adaption and Linguistic Validation

The translation and linguistic validation was performed according to recommendations of the International Society for Pharmacoeconomics and Outcomes Research²² and the World Health Organization²³ and is described in detail elsewhere.¹³ Briefly, 2 independent forward translations were produced by a professional trans-

lator and a clinician, both of whom had Danish as their mother tongue and were fluent in English. The harmonized forward translation was provided to a professional translator who had English as their mother tongue and was fluent in Danish who performed a back translation. The back translation was then compared with the original English version. Discrepancies were discussed with the instrument's developers and revised in an iterative manner until a final version was produced. An expert panel meeting was held to ensure the BODY-Q included all clinically relevant issues from the perspective of Danish clinicians. We then conducted 2 rounds of cognitive interviews with a total of 22 patients. These interviews were used to determine if any aspect of the BODY-Q was poorly worded or asked about issues not relevant to patients in Denmark. Feedback from patients and proof reading by 2 clinicians was used to revise and finalize the Danish version of the BODY-Q.

Stage II: Field-Test

The BODY-Q and relevant demographic and clinical questions were developed into an online REDCap (i.e., Research Electronic Data Capture) survey (<http://project-redcap.org>). Researchers at Odense University Hospital are granted access to REDCap through the Odense Patient data Explorative Network.²⁴ In the period June 2015 to June 2016, 4 groups of patients were recruited from Odense University Hospital (Department of Endocrinology and Department of Plastic Surgery) and Hospital of Southwest Jutland (Bariatric Center/Department of Endocrinology and Department of Plastic Surgery):

1. Prebariatric surgery: This group included patients who were referred for obesity treatment and/or bariatric surgery.
2. Postbariatric surgery: This group included patients who had bariatric surgery. Participants were asked to complete the BODY-Q survey at the following intervals: 4-5, 12, and 24 months postsurgery.
3. Prebody contouring: Patients referred for body-contouring surgery following massive weight loss.
4. Postbody contouring: Patients who had undergone body contouring following massive weight loss. Participants had undergone at least one of the following procedures: abdominoplasty, upper arm lift, thigh lift, buttocks lift, and/or breast lift. Patients were asked to complete the BODY-Q survey 3 and 12 months after body contouring.

Patients scheduled for an outpatient appointment were sent an information letter about the study, with a link to access and complete the BODY-Q survey in REDCap. At the Hospital of Southwest Jutland, patients were sent mobile text reminders before their scheduled outpatient clinic appointment. At both hospitals, patients attending an outpatient appointment who had not completed the BODY-Q survey before their appointment were invited to do so using a tablet in the outpatient clinic. Patients were, according to agreement with the Danish Data Protection Agency, asked electronically to provide informed consent

for their data being used for research. Exclusion criteria were inability to speak or understand Danish.

Stage III: Data Analysis

Descriptive data were analyzed using SPSS software (IBM SPSS Statistics, version 22.0, IBM Corp). The psychometric properties of the Danish BODY-Q were analyzed using RMT analysis to determine if the scales meet the requirements of the Rasch model. Analysis was performed using RUMM 2030 software (RUMM version 2030, 1998-14 RUMM Laboratory Pty Ltd.; <http://www.rummlab.com.au>).

RMT Analysis

RMT is a psychometric method that can provide insights into strengths and limitations of a PRO scale.²⁰ We performed the same set of analyses as reported in the BODY-Q development and psychometric validation article,¹¹ to compare the findings of the Danish and original versions. The psychometric parameters that were evaluated included the following:

1. Category threshold order: For each BODY-Q scales, we examined thresholds between item response options (e.g., definitely disagree, somewhat disagree, somewhat agree, definitely agree). Disordered thresholds can occur as a consequence of unclear definitions, the number of response options, and underutilization.²⁵ The measurement probabilities of a PRO scale can be compromised if response options are not used in an orderly fashion.²⁶
2. Measurement precision: For each scale, we estimated the measurement precision by the person separation index (PSI). PSI measures the error, that is, the higher PSI, the higher reliability.²⁷ PSI can be compared with the Cronbach α value.²⁸
3. Item fit statistics: We examined 3 indicators of item fit to the Rasch model: (1) Log residuals (item-person interaction); (2) χ^2 values (item-trait interaction); and (3) Item characteristic curves. Results were interpreted fit statistics together and in relation to the clinical importance of each item. Fit residuals should be between -2.5 and $+2.5$, and χ^2 values should be nonsignificant after Bonferroni adjustment.
4. Dependency: We examined the residual correlations between items. Correlation between pairs of items should be lower than 0.30. High residual correlations can artificially inflate scale reliability. If values were ≥ 0.30 , we performed substest analysis to investigate the impact they had on the PSI.²⁹
5. Targeting: Targeting examines the extent to which the items that form a scale measure the construct that is experienced by the study population. For targeting, we examined the person and item locations.
6. Differential Item Functioning: Differential item functioning (DIF) refers to the stability of an instrument, that is, whether an item is responded to differently by subgroups within a population. To examine DIF, the BODY-Q dataset from the original study sample (Canada, the United States, and the United Kingdom) was

included alongside the Danish dataset and analyzed in RUMM2030. We examined DIF by country to determine if the Danish sample differed from the original study sample. Items with potential DIF were located by significant chi-square values after Bonferroni adjustment. If DIF was located, the sample was split by DIF and then looked at to see if it made any difference.

7. Correlation to original scoring: To examine whether the original BODY-Q scoring key can be used, we correlated the logit scores for each scale's set of items with for the Danish study sample and original study sample (Canada, the United States, and the United Kingdom).

RESULTS

A total of 495 patients completed the Danish BODY-Q field test between 1 and 4 times. Including data for patients who filled out the experience of care scales twice (i.e., once in relation to their experience of care in weight loss clinic and once for their experience in relation to their experience of care in the bariatric surgery clinic), there were a total of 681 assessments. Characteristics for each patient group are outlined in Table 1. The overall response rate was 76%. For the bariatric group, the response rate was 84%, and for the body-contouring group, it was 72%. Sixteen percentage of responses were prebariatric surgery, 33% were postbariatric surgery, 34% were prebody contouring, and 17% were postbody contouring. The percentage of missing responses at item level was 1% in total.

Overall, the RMT analysis provided broad support for the reliability and validity of the Danish version of the BODY-Q scales. Appendix 1 provides the detailed RMT statistics for the 138 BODY-Q items, organized by scale, and within scale, by the item order based on the Rasch model in the original study (**Appendix, Supplemental Digital Content 1**, <http://links.lww.com/PRSGO/A572>). For item fit statistics, item fit was outside the criteria of -2.5 to $+2.5$ for 34 of 138 items, and of these, 21 had a significant chi-square P value after Bonferroni adjustment.

Most items (128 of 138) had ordered thresholds. Of the 10 disordered items, all except one item in the Physical Function scale [Bending over (e.g., to tie your shoes)] was in a patient experience scale, including the Doctor scale (item 100, 101, 102, 103, 105, and 106) and Information scale (item 129, 130, and 132).

Table 2 provides the scale performance findings. In terms of scale reliability, Cronbach α values were ≥ 0.90 for all scales. Appearance scales and QOL scales evidence high reliability, with all PSI values (with and without extremes) ≥ 0.74 [exception Hips/Outer Thighs = 0.37 (without extremes)]. The PSI values for the patient experience scales were lower and ranged from 0.32 to 0.70 (with extremes) and 0.69–0.86 (without extremes).

Item residual correlations were above 0.30 (range, 0.31–0.52) for 10 pairs of items within 8 scales. In substest analysis, the correlated items were found to have minor influence on scale reliability with a difference in PSI value ≤ 0.73 .

Table 1. Characteristics of the Field-Test Sample, Divided by Classification of Time of Journey

Characteristics	Prebariatric Surgery (N = 110)	Postbariatric Surgery (N = 226)	Prebody Contouring (N = 230)	Postbody Contouring (N = 114)
Age (N = 679)				
Mean	43	42	42	45
Range	(21–63)	(23–65)	(17–66)	(23–75)
Gender (N = 680) (%)				
Female	74 (67)	164 (73)	193 (84)	95 (83)
Male	36 (33)	62 (27)	37 (16)	19 (17)
BMI (N = 670)				
Mean	45	32	28	27
Range	(32–68)	(20–63)	(21–67)	(20–36)
Comorbidity (N = 674) (%)				
Yes	77 (71)	100 (45)	67 (29)	35 (31)
No	32 (29)	124 (55)	161 (71)	78 (69)
Marital status (N = 673) (%)				
Married	57 (52)	122 (54)	88 (39)	46 (41)
Living together	16 (15)	40 (18)	63 (28)	33 (29)
Widow	1 (1)	0 (0)	4 (2)	3 (3)
Separated	1 (1)	2 (1)	2 (1)	2 (2)
Divorced	10 (9)	16 (7)	26 (11)	11 (10)
Single, never married	25 (23)	44 (20)	44 (19)	17 (15)
Education (N = 679) (%)				
None	1 (1)	8 (4)	6 (3)	2 (2)
Primary school	25 (23)	44 (19)	37 (16)	18 (16)
Youth, vocational or secondary education	30 (27)	56 (25)	52 (23)	23 (20)
Short higher education	17 (15)	46 (20)	58 (25)	24 (21)
Medium higher education	32 (29)	62 (27)	65 (28)	42 (37)
Long higher education	5 (5)	8 (4)	12 (5)	4 (4)
PhD or master	0 (0)	2 (1)	0 (0)	0 (0)

Table 2. Overview of Scale Performance Statistics

Scale	Sample Size	PSI with Extremes	PSI No Extremes	Chi-Square	DF	P	F/C (%)	Crohnbachs Alpha
Body	568	0.92	0.92	181.23	90	0.000	7/1	0.95
Arms	565	0.91	0.90	149.49	56	0.000	14/3	0.95
Abdomen	567	0.92	0.93	103.36	42	0.000	33/5	0.98
Back	558	0.88	0.83	76.34	20	0.000	15/12	0.96
Buttocks	562	0.90	0.83	34.43	30	0.264	17/10	0.95
Hips and outer thighs	563	0.92	0.37	26.20	30	0.665	20/2	0.97
Inner thighs	561	0.79	0.80	48.18	20	0.000	43/3	0.96
Skin	481	0.85	0.86	48.55	35	0.064	29/3	0.95
Scars	112	0.86	0.88	46.21	20	0.001	1/26	0.95
Body image	566	0.90	0.90	67.63	49	0.040	23/1	0.96
Physical	450	0.80	0.82	57.57	35	0.010	2/24	0.92
Psychological	567	0.92	0.92	157.90	90	0.000	3/7	0.95
Sexual	559	0.80	0.74	70.35	40	0.002	16/5	0.90
Social	567	0.90	0.90	158.97	90	0.000	1/10	0.94
Doctor	342	0.32	0.69	18.41	20	0.561	0/56	0.95
Medical team	342	0.70	0.86	29.27	20	0.083	0/49	0.96
Office staff	329	0.47	0.86	21.18	20	0.387	1/70	0.97
Information	227	0.68	0.74	30.71	20	0.059	1/33	0.92

DF, degree of freedom; P, probability.

Figure 1 shows examples of the person-item threshold distribution for the social (1a) and psychological (1b) function scales. Person locations are shown in the upper half, whereas item locations are located in the lower half. The figures confirm that the scales are able to provide information for all levels of the constructs measured (Figs. 1, 2).

DIF was examined by comparing the original study population (Canada, the United States, and the United Kingdom) with the Danish study population. The original study sample consisted of 965 assessments provided by 734 participants. We adjusted the sample size to 500 for the analysis, and DIF was significant for 27 of 138 items. How-

ever, when these items were split on the variable (country) with DIF, and the new person locations were correlated with the original person locations, DIF was found to have a negligible impact. For 17 of 18 scales, the Pearson correlation was ≥ 0.99 for the calculated score. The correlation for the remaining scale (office staff) was 0.97. Our findings confirm that the original scoring key can be used in the Danish version.

The physical symptom checklist was completed 572 times. Table 3 shows frequencies for each symptom. The most frequent symptom reported to be there all the time was joint pain followed by short of breath with mild exercise and back pain.

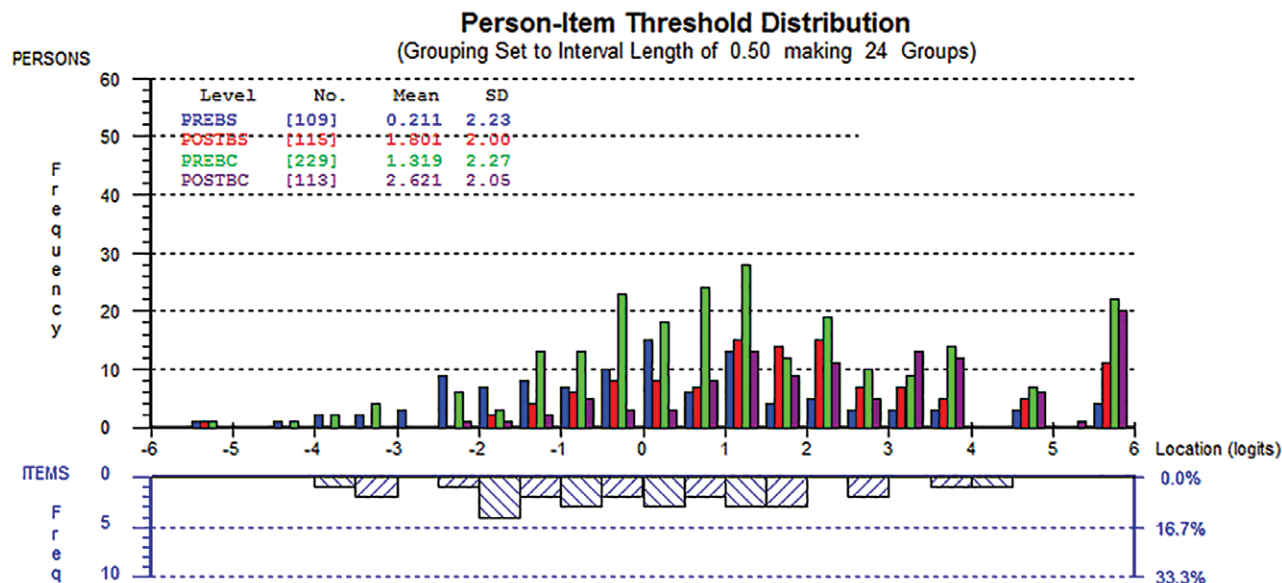


Fig. 1. Person-item threshold distribution for the social wellbeing scale.

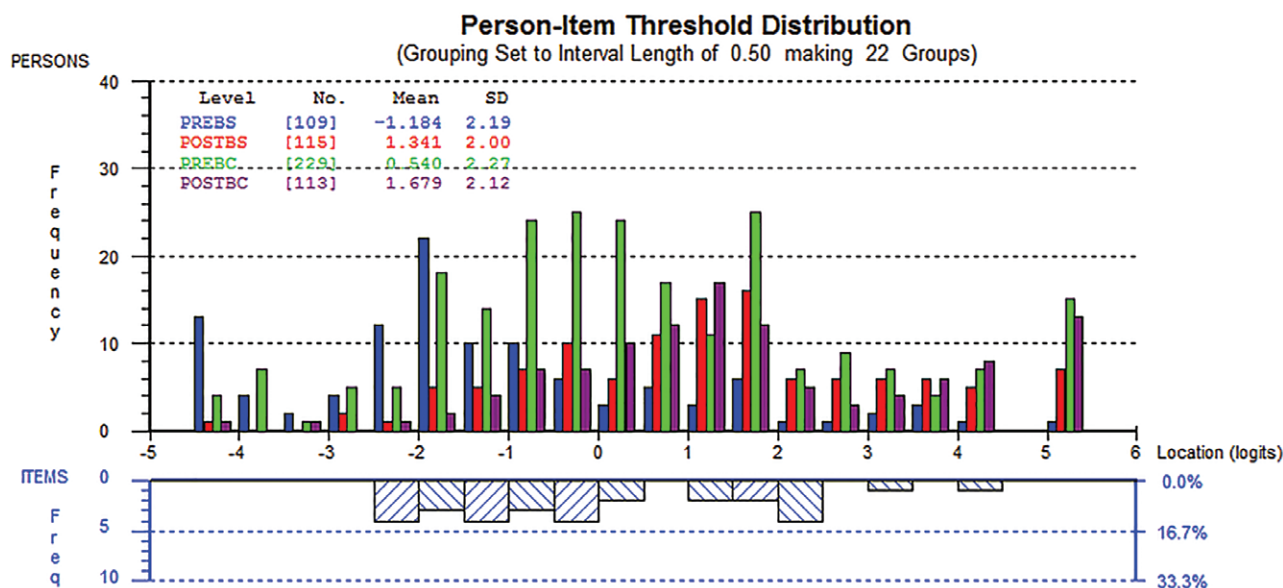


Fig. 2. Person-item threshold distribution for the psychological wellbeing scale.

Table 3. Physical Symptom Frequency, Total Number (%) of Assessments

Symptom	All the Time, N (%)	Often, N (%)	Sometimes, N (%)	Never, N (%)
Feeling tired during the day	58 (10)	140 (25)	215 (38)	39 (7)
Back pain	65 (11)	102 (18)	183 (32)	102 (18)
Joint pain	76 (13)	103 (18)	159 (28)	114 (20)
Leg pain or discomfort	46 (8)	102 (18)	156 (27)	147 (26)
Feeling of balance	27 (5)	57 (10)	194 (34)	174 (30)
Feeling weak	32 (6)	70 (12)	184 (32)	166 (29)
Short of breath with mild exercise	66 (12)	57 (10)	144 (25)	184 (32)
Swollen feet	40 (7)	58 (10)	152 (27)	202 (35)
Skin rash or infection	28 (5)	55 (10)	133 (23)	236 (41)
Too much perspiration	34 (6)	59 (10)	134 (23)	223 (39)

DISCUSSION

The findings from our study provide broad support that the BODY-Q scales were acceptable, reliable, and valid in a Danish sample of bariatric and body-contouring patients. Based on the high response rate obtained for the online questionnaire survey³⁰ and the low amount of missing data at item level (despite the large number of items tested), the BODY-Q was judged to be acceptable to patients.³¹ Patients completed all relevant scales, and a subset of patients even completed the experience of care scales twice. Elsewhere, we report that patients in the cognitive interview phase of the translation process found the BODY-Q to be acceptable and that they were happy to fill it out and even relieved to have the opportunity to express how they felt.¹³

Within the RMT analysis, we observed some degree of misfit that we need to explore. Item fit was outside the criteria for 34 of 138 items, and of these, 21 had a significant chi-square *P* value after Bonferroni adjustment. This only indicates that the observed responses to these items did not fit the Rasch model perfectly and that the misfit needs to be explored further. Previously a way to adjust for this was to exclude items; however, newer literature suggests that reasons should be explored and explained instead.³² As we found all items clinically relevant and important, we did not wish to exclude any items, and working through the affected items, we found good explanations to the statistical results. For example, 1 item was about the size of the body, which can be interpreted as both weight and subjective size. Another example was items asking about how toned specific parts of the body were. In Danish toned is not the most commonly used word, and as cognitive interviews suggested, it was translated into well trained instead, which might have influenced results.¹³ For thresholds, 128 of 138 items were ordered, indicating that response options worked as intended. Most disordered items were in the Doctor and Information scale, and when comparing the item order of these scales with the original BODY-Q item order, it was in broad agreement, which provides evidence of validity.

In terms of reliability, Cronbach α values were ≥ 0.90 for all scales, and PSI values for all scales measuring appearance and quality of life concerns were high. Lower reliability was noted for the experience of care scales, which may warrant further examination. This finding is likely due to the finding that a high proportion of patients in the Danish sample reported scores at the ceiling on these 4 scales, particularly according to the doctor scale. Ceiling effects are important since they limit the ability to measure change and thereby measurement accuracy. The observed ceiling effects in the experience scales might be explained by cultural differences, a too small sample size—or simply reflection of patients' satisfaction with treatment. These findings again could relate to the fact that the original sample is more heterogeneous, as participants were included from a range of weight loss and plastic surgery healthcare providers from 4 countries (England, Scotland, the United States, and Canada). Saiga et al.^{33,34} recently published a Japanese translation and psychometric analysis of the BREAST-Q and also found ceiling effect

in the experience of care scales. They suggest that the ceiling effect could be associated with study design, which could also be the situation in our design, where patients were asked to fill out the BODY-Q as preparation for their follow-up visit in the outpatient clinic.

In item residual correlations, a cutoff point at 0.3 is often used, though it needs to be used carefully while the number of items potentially can influence. When looking at the scales and items, some amount of dependency must be expected; however, it has the potential to falsely inflate reliability. To investigate this further, we did a subtest analysis showing a difference in PSI value ≤ 0.73 . If the reliability decreases in the subtest, it should be viewed as a more valid description of data.³⁵ Based on this, the correlated items that we found should only have minor influence on scale reliability.

There are several limitations to our study. First, the sample was composed mainly of women, which was similar to the development sample and reflects the fact that more women come forward for bariatric and body-contouring surgery.³⁶ Second, our sample does not include patients from all regions of Denmark. However, our study included 4 different departments in 2 hospitals, and as Denmark is a small country with a total population of approximately 5.7 million people,³⁷ our sample must be considered representative. Finally, we did not examine test–retest reliability or responsiveness. These psychometric aspects could be the focus of future research.

CONCLUSIONS

Our study confirms that the Danish version of the BODY-Q is a valid and reliable PRO instrument for use in Danish bariatric and body-contouring patients. Furthermore, our study and analysis underlines the importance of doing a thorough translation and linguistic validation followed by field-testing and psychometric evaluation. Finally, we have demonstrated that the Danish version of the BODY-Q fulfills the need of a valid and reliable PRO instrument for the bariatric and body-contouring population in Denmark.

Lotte Poulsen, MD

Department of Plastic Surgery
Odense University Hospital
Sdr. Boulevard 29
5000 Odense
Denmark

E-mail: Lotte.poulsen2@rsyd.dk; lotte_poulsen@hotmail.com

REFERENCES

1. Poulsen L, Roessler KK, Rose M, et al. Quality of life of bariatric and body contouring. *Ugeskr Laeger*. 2015;177(20):pii: V12140668.
2. Danish Database for Bariatric Surgery, The Danish National Board of Health. Available at <https://www.sundhed.dk/sundhedsfaglig/kvalitet/kliniske-kvalitetsdatabaser/planlagt-kirurgi/dansk-fedmekirurgi/>. Accessed November 7, 2016.
3. de Zwaan M, Georgiadou E, Stroh CE, et al. Body image and quality of life in patients with and without body contouring surgery following bariatric surgery: a comparison of pre- and post-surgery groups. *Front Psychol*. 2014;5:1310.
4. Tayyem R, Ali A, Atkinson J, et al. Analysis of health-related quality-of-life instruments measuring the impact of bariatric surgery:

- systematic review of the instruments used and their content validity. *Patient*. 2011;4:73–87.
5. Reavey PL, Klassen AF, Cano SJ, et al. Measuring quality of life and patient satisfaction after body contouring: a systematic review of patient-reported outcome measures. *Aesthet Surg J*. 2011;31:807–813.
 6. Jabir S. Assessing improvement in quality of life and patient satisfaction following body contouring surgery in patients with massive weight loss: a critical review of outcome measures employed. *Plast Surg Int*. 2013;2013:515737.
 7. Klassen AF, Cano SJ, Scott A, et al. Assessing outcomes in body contouring. *Clin Plast Surg*. 2014;41:645–654.
 8. Morley D, Jenkinson C, Fitzpatrick R. A structured review of patient-reported outcome measures used in cosmetic surgical procedures. Oxford, United Kingdom: Health Services Research Unit Department of Public Health University of Oxford; 2013.
 9. Danilla S, Cuevas P, Aedo S, et al. Introducing the Body-QoL®: a new patient-reported outcome instrument for measuring body satisfaction-related quality of life in aesthetic and post-bariatric body contouring patients. *Aesthetic Plast Surg*. 2016;40:19–29.
 10. Gilmartin J, Bath-Hextall F, Maclean J, et al. Quality of life among adults following bariatric and body contouring surgery: a systematic review. *JBI Database System Rev Implement Rep*. 2016;14:240–270.
 11. Klassen A, 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;13:4.
 12. Moorehead MK, Ardel-Gattinger E, Lechner H, et al. The validation of the Moorehead-Ardelt Quality of Life Questionnaire II. *Obes Surg*. 2003;13:684–692.
 13. Poulsen L, Rose M, Klassen A, et al. Danish translation and linguistic validation of the BODY-Q: a description of the process. *Eur J Plast Surg*. 2017;40:29–38.
 14. Lasch KE, Marquis P, Vigneux M, et al. PRO development: rigorous qualitative research as the crucial foundation. *Qual Life Res*. 2010;19:1087–1096.
 15. Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1—eliciting concepts for a new PRO instrument. *Value Health*. 2011;14:967–977.
 16. Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality of life instruments: attributes and review criteria. *Qual Life Res*. 2002;11:193–205.
 17. Cosmin checklist. Available at http://www.cosmin.nl/the_cosmin_checklist.html Accessed November 7, 2016.
 18. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63:737–745.
 19. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19:539–549.
 20. Rasch analysis. Available at <http://www.rasch-analysis.com/rasch-analysis.htm> Accessed November 7, 2016.
 21. Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. *Health Technol Assess*. 2009;13:iii, ix–ixix, 1.
 22. Wild D, Grove A, Martin M, et al.; ISPOR Task Force for Translation and Cultural Adaptation. 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 Health*. 2005;8:94–104.
 23. Process of translation and adaptation of instruments, World Health Organization. Available at http://www.who.int/substance_abuse/research_tools/translation/en/ Accessed November 7, 2016.
 24. Odense Patient data Explorative Network, University of Southern Denmark. Available at http://www.sdu.dk/en/om_sdu/institutter_centre/klinisk_institut/forskning/forskningsenheder/open Accessed November 7, 2016.
 25. Pallant JF, Miller RL, Tennant A. Evaluation of the Edinburgh Post Natal Depression Scale using Rasch analysis. *BMC Psychiatry*. 2006;6:28.
 26. 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.
 27. Bond TG, Fox CM. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. 2nd ed. London, United Kingdom: Lawrence Erlbaum Associates; 2007.
 28. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16:297–334.
 29. Wright BD, Masters GN. *Rating Scale Analysis: Rasch Measurement*. Chicago, Ill.: MESA Press; 1982.
 30. Shih TH, Xitao F. Comparing response rates from web and mail surveys: a meta-analysis. *Field Methods*. 2008;20:249–271.
 31. Edwards P, Roberts I, Sandercock P, et al. Follow-up by mail in clinical trials: does questionnaire length matter? *Control Clin Trials*. 2004;25:31–52.
 32. Quality of life measurement in neurodegenerative and related conditions. Jenkinson C, Peters M, Bromberg MB, eds. Cambridge, United Kingdom: Cambridge University Press; 2011.
 33. Pusic AL, Klassen AF, Scott AM, et al. Development of a new patient-reported outcome measure for breast surgery: the BREAST-Q. *Plast Reconstr Surg*. 2009;124:345–353.
 34. Saiga M, Taira N, Kimata Y, et al. Development of a Japanese version of the BREAST-Q and the traditional psychometric test of the mastectomy module for the assessment of HRQOL and patient satisfaction following breast surgery. *Breast Cancer*. 2017;24(2):288–298.
 35. Rasch models in health. Christensen KB, Kreiner S, Mesbah M, eds. Hoboken, New Jersey: John Wiley & Sons; 2013.
 36. Froylich D, Corcelles R, Daigle CR, et al. Weight loss is higher among patients who undergo body contouring procedures after bariatric surgery. *Surg Obes Relat Dis*. 2016;12(9):1731–1736.
 37. Statistics Denmark. Available at <https://www.dst.dk/da/Statistik/emner/befolkning-og-befolkningsfremskrivning/folketal> Accessed November 7, 2016.