Validation of a New Measure of Quality of Life in Obesity Trials: IWQOL-Lite Clinical Trials Version

Ronette L Kolotkin,[1] Valerie SL Williams,[2] Claire M Ervin,[2] Henrik Hjorth Meincke,[3] Shanshan Qin,[2] Lisa von Huth Smith,[3] Sheri E Fehnel[2]

[1]Quality of Life Consulting, Durham, NC, United States; Department of Community and Family Medicine, Duke University School of Medicine, Durham, NC, United States; Western Norway University of Applied Sciences, Førde, Norway; Centre of Health Research, Førde Hospital Trust, Førde, Norway; Morbid Obesity Centre, Vestfold Hospital Trust, Tønsberg, Norway; [2]RTI Health Solutions, Research Triangle Park, NC, United States; [3]Novo Nordisk A/S, Søborg, Denmark

(qrs.ly/fc7s13d)
Scan for more Novo Nordisk scientific materials

Background

- Health-related quality of life (HRQOL) and patient functioning are important secondary outcomes in clinical trials of new interventions for weight loss.
- The Impact of Weight on Quality of Life–Lite (IWQOL-Lite) is widely used to assess HRQOL and functioning in evaluations of diverse weight loss interventions, including pharmaceutical trials.[1-3]
- While the IWQOL-Lite was developed using rigorous qualitative research methods and has demonstrated strong psychometric properties, the content was initially based on the input of individuals undergoing intensive residential treatment for obesity and related comorbid conditions.
  - Accordingly, the IWQOL-Lite may be missing some concepts that are relevant to clinical trial populations or may include concepts that are not relevant to these populations.
- As a result, the United States (US) Food and Drug Administration (FDA) has not allowed clinical trial results based on this measure to be described in product labeling.
- Thus, an alternative version of the IWQOL-Lite questionnaire optimized for use in patient populations typically targeted for weight loss clinical trials, the IWQOL-Lite Clinical Trials Version (IWQOL-Lite-CT), has been developed.[4]

Objective

- The objective of this study was to validate the IWQOL-Lite-CT in accordance with the FDA patient-reported outcomes (PRO) guidance.[5]

Methods

Study Population

- Psychometric analyses of the IWQOL-Lite-CT were conducted in the context of two randomized trials.
  - Study NCT02453711 (Study 1) was a multinational, randomized, double-blind, placebo-controlled phase 2 trial of treatment with subcutaneous semaglutide for 52 weeks conducted among individuals with obesity (body mass index [BMI] ≥ 30.0 kg/m^2) and without diabetes. Psychometric analyses were conducted using the subset of US patients who completed patient-reported measures at baseline (n = 329).
  - Study NCT02906930 (Study 2) was a multinational, randomized, double-blind, placebo-controlled phase 3a trial of treatment with oral semaglutide for 26 weeks in patients with type 2 diabetes mellitus treated with diet and exercise only. Psychometric analyses were conducted using the subset of English-speaking US patients with BMI ≥ 27.0 kg/m^2 who completed the patient-reported measures at baseline (n = 145).

Measures

- The psychometric evaluation included the following measures:
  - IWQOL-Lite-CT (Study 1 [23 items] and Study 2 [22 items])
  - Short Form Health Survey–36 (SF-36) subscale scores and physical and mental component summary scores (PCS and MCS) (Studies 1 and 2)
  - Patient Global Impression of Change (PGI-C) for Physical Functioning and Mental Health (Studies 1 and 2)
  - PGI-C for Quality of Life (Study 1 only)
  - Patient Global Impression of Status (PGI-S) for Physical Functioning and Mental Health (Study 2 only)

Analytic Method

- Figure 1 describes the psychometric analysis methods used to evaluate the IWQOL-Lite-CT.

Figure 1: Psychometric Methods Used to Evaluate the IWQOL-Lite-CT

Descriptive characteristics and response distributions
- Standard descriptive statistics were computed to characterize the sample
- Item-level response frequency distributions were examined for floor and ceiling effects[a]

Structure and scoring
- Inter-item correlations were computed
- Exploratory factor analyses (Study 1) and confirmatory factor analyses (Study 2) were conducted

Reliability
- To evaluate internal consistency, Cronbach's coefficient alpha was computed[6]
- To evaluate test-retest reliability, intraclass correlation coefficients were computed for IWQOL-Lite-CT composite scores (Physical, Physical Function, and Psychosocial)
  - "Test" data were Week 48 composite scores in Study 1 and Week 0 composite scores in Study 2; "retest" data were Week 52 composite scores in Study 1[b] and Week 8 composite scores in Study 2

Validity

Construct validity
- Correlations were computed between IWQOL-Lite-CT scores and scores on SF-36 subscales and components, PGI-C items, and PGI-S items, at individual time points and for change scores
  - In Study 1, at least moderate (|r| ≥ 0.30) correlations were hypothesized for the following:
    - IWQOL-Lite-CT Physical and Physical Function scores and the SF-36 PCS and the SF-36 Physical Functioning scores
    - IWQOL-Lite-CT Psychosocial score and the SF-36 MCS and the SF-36 Mental Health and Social Functioning subscales
  - In Study 2, at least moderate (positive) correlations (|r| ≥ 0.30) were hypothesized for the following:
    - IWQOL-Lite-CT Physical and Physical Function scores and the PGI-S physical functioning item; the SF-36 PCS; and the SF-36 Physical Functioning, Role-Physical, and Vitality subscale scores
    - IWQOL-Lite-CT Psychosocial score and the PGI-S mental health item, the SF-36 MCS, and the SF-36 Mental Health and Vitality scores

Discriminating ability
- Known-groups analyses of variance examined mean differences in IWQOL-Lite-CT composite scores between patients classified into subgroups (BMI and % weight loss)

Responsiveness
- To evaluate responsiveness, effect sizes, standardized response means, and Cohen's *d* statistics were computed[7]

[a] A floor or ceiling effect would require that more than 40% of the patients select the worst or best (modal) response category, respectively.

[b] Only patients with less than 5% change in body weight from Week 48 to Week 52 (in Study 1) and from Week 0 to Week 8 (in Study 2) were used in the test-retest reliability analyses.

Results

Participant Characteristics

Table 1: Participant Characteristics at Baseline

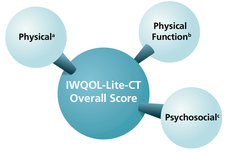
Characteristic	Study 1 (N = 329)	Study 2 (N = 145)
Age (years), mean (SD), median	47.9 (11.9), 49.0	55.1 (11.5), 56.0
Min, Max	19.0, 76.0	28.0-79.0
Sex, n (%)		
Male	116 (35.3)	77 (53.1)
Female	213 (64.7)	68 (46.9)
Height (meters), mean (SD), median	1.69 (0.1), 1.7	1.67 (0.1), 1.7
Min, Max	1.4, 2.0	1.5-2.0
Weight (pounds), mean (SD), median	250.8 (53.6), 242.1	215.9 (50.0), 203.9
Min, Max	164.9, 476.8	138.0, 465.0
BMI, mean (SD), median	39.9 (7.6), 38.1	34.8 (5.8), 33.2
Min, Max	29.8, 77.1	27.1-63.1
Race, n (%)		
Asian	2 (0.6)	4 (2.8)
Black or African American	57 (17.3)	29 (20.0)
White	261 (79.3)	105 (72.4)
Native Hawaiian/Pacific Islander	1 (0.3)	1 (0.7)
American Indian/Alaska Native	4 (1.2)	3 (2.1)
Other	4 (1.2)	3 (2.1)
Hispanic or Latino	40 (12.2)	49 (33.8)

SD = standard deviation.

Response Distributions

- The frequency distributions in both studies generally supported the appropriateness of the response categories.
  - In Study 1, three IWQOL-Lite-CT items exhibited ceiling effects (i.e., reporting Never or Not at all true): Unable to stand comfortably, Self-conscious eating in social settings, and Avoid social gatherings.
  - In Study 2, nine items displayed substantial ceiling effects at baseline and Week 26: Self-conscious eating in social settings, Feel judged by others, Less important/worthy of respect, Feel down or depressed, Avoid social gatherings, Less productive, Limited self-esteem, Self-conscious about weight, and Frustrated or upset about weight.
  - On average, participants' BMIs at baseline were lower in Study 2 than in Study 1.

Figure 2: IWQOL-Lite-CT Structure



[a] Comprises Items 1-5, 16, and 17 (see Table 3). [b] Comprises Items 1-3, 16, and 17, with Item 4 (Uncomfortable in small seats) and Item 5 (Bodily pain) excluded. [c] Comprises Items 6-15 and 18-20.

Composites are mean scores on a 0-100 scale.

Structure and Scoring

- Exploratory (Study 1) and confirmatory (Study 2) factor analyses supported the structure of the IWQOL-Lite-CT shown in Figure 2.
  - Inter-item correlations identified several redundancies that informed item reduction and led to the final 20-item IWQOL-Lite-CT.

Reliability

- Internal consistency and test-retest reliability were supported in both studies.
  - Cronbach's alpha was ≥ 0.77 for all composite scores at each time point across both studies.
  - Intraclass correlation coefficients were ≥ 0.80 for all composite scores at each time point across both studies.

Validity

- Cross-sectional correlational analyses provided support for the construct validity of the IWQOL-Lite-CT composite scores.
  - In the Study 1 analyses:
    - Physical and Physical Function scores correlated strongly with the SF-36 PCS score and the SF-36 Physical Functioning subscale.
    - Psychosocial scores correlated moderately with the SF-36 MCS and Mental Health and Social Functioning scores.
  - In the Study 2 analyses:
    - Physical and Physical Function scores correlated strongly with the PGI-S physical functioning item and the SF-36 PCS and Physical Functioning, Role-Physical, and Vitality subscale scores.
    - Psychosocial scores correlated moderately with the SF-36 MCS and correlated strongly with the Mental Health and Vitality subscales.
- Longitudinal analyses provided additional support for the construct validity of the composite scores.
- Known-groups analyses of variance supplied evidence of the discriminating ability of the IWQOL-Lite-CT.
  - At Weeks 28 and 52 in Study 1:
    - All mean differences were in the correct direction, but only the Physical Function score obtained statistically significant differences between groups of patients with ≥ 5% weight loss vs. patients with weight gain at Week 52.
    - All composite scores demonstrated statistically significant group differences in the expected direction among patients with BMI < 30 vs. patients with BMI > 42.
  - In Study 2:
    - All composite scores were in the correct direction but were not significantly different for patients with ≥ 5% weight loss vs. patients with weight gain at Week 26.
    - All composite scores demonstrated statistically significant group differences in the expected direction among patients with BMI < 30 vs. patients with BMI > 42 at Weeks 8 and 26.

Responsiveness

- In Study 1:
  - All IWQOL-Lite-CT composite scores achieved large effect-size estimates and SRMs at Week 52.
  - Responsiveness statistics for most of the items were mostly moderate to large at Weeks 28 and 52, indicating that the IWQOL-Lite-CT items were easily capable of detecting change.
  - The IWQOL-Lite-CT items that exhibited ceiling effects had somewhat smaller responsiveness statistics, as expected.
- In Study 2, responsiveness statistics were smaller than in Study 1, which was unsurprising given that only minor changes in BMI were observed from baseline to the end of the study.
- Table 2 details the results of the composite-level responsiveness analyses.

Table 2: Ability to Detect Change—Effect Size Estimates, Standardized Response Means, and Cohen's *d* Statistics

Scores	Effect Size Estimate	SRM	Cohen's *d* – PGI-S: Improved vs. Unchanged	Cohen's *d* – PGI-S: Worsened vs. Unchanged	Cohen's *d* – PGI-S: Improved vs. Worsened	Cohen's *d* – PGI-C: Improved vs. Unchanged	Cohen's *d* – PGI-C: Worsened vs. Unchanged	Cohen's *d* – PGI-C: Improved vs. Worsened	Cohen's *d* – Change in Weight: Improved vs. Unchanged	Cohen's *d* – Change in Weight: Worsened vs. Unchanged	Cohen's *d* – Change in Weight: Improved vs. Worsened
Study 1											
Baseline to Week 28											
IWQOL-Lite-CT Total[a]	0.9	1.1	—	—	—	0.2	–2.3	1.1	0.5	—	—
Physical[b]	0.8	1.0	—	—	—	0.3	–0.8	0.6	0.6	—	—
Physical Function[b]	0.8	1.0	—	—	—	0.4	–0.7	0.5	0.6	—	—
Psychosocial[c]	0.8	1.1	—	—	—	0.4	–0.9	0.8	0.4	—	—
Baseline to Week 52											
IWQOL-Lite-CT Total[a]	1.1	1.2	—	—	—	0.6	–0.6	1.3	0.7	–0.8	2.5
Physical[b]	1.0	1.0	—	—	—	0.4	–0.6	1.3	0.6	–1.0	3.2
Physical Function[b]	1.0	1.0	—	—	—	0.5	–0.6	1.2	0.7	–1.2	3.5
Psychosocial[c]	1.0	1.1	—	—	—	0.7	–1.0	1.1	0.7	–0.6	1.8
Study 2											
IWQOL-Lite-CT Total[d]	0.1	0.2	0.4	–0.5	0.9	0.7	—	—	0.2	0.2	–0.0
Physical[e]	0.1	0.1	0.5	–0.2	0.6	0.4	—	—	0.0	–0.1	0.1
Physical Function[e]	0.1	0.1	0.5	–0.3	0.6	0.4	—	—	–0.0	–0.1	0.1
Psychosocial[f]	0.2	0.2	0.4	–0.4	0.8	0.4	–0.1	0.4	0.3	0.4	–0.1

PGI-C PF = PGI-C Physical Functioning; PGI-C PS = PGI-C Psychosocial; PGI-C QoL = PGI-C Quality of Life; SRM = standardized response mean; — = sample size of fewer than 5 patients.

Note: PGI-C improved = 1 (Much better), 2 (Moderately better), or 3 (A little better); unchanged = 4 (No difference); and worsened = 5 (A little worse), 6 (Moderately worse), or 7 (Much worse).

Body weight improved = 5% or more weight loss, unchanged = weight change (gain or loss) less than 5%, worsened = 5% or more gain in weight.

[a] Using PGI-C QoL. [b] Using PGI-C PF. [c] Using PGI-C PS. [d] Using both PGI-S PF and PGI-S MH or PGI-C PF and PGI-C MH. [e] Using PGI-S PF or PGI-C PF. f Using PGI-S MH or PGI-C MH.

Table 3: IWQOL-Lite-CT Concepts

Item Number	Concept
Item 1	Trouble bending over
Item 2	Tired or winded
Item 3	Unable to stand comfortably
Item 4	Uncomfortable in small seats
Item 5	Bodily pain
Item 6	Self-conscious eating in social settings
Item 7	Less confident
Item 8	Feel judged by others
Item 9	Frustrated shopping for clothes
Item 10	Feel bad or upset about pictures of self
Item 11	Feel down or depressed
Item 12	Less interested in sexual activity
Item 13	Avoid social gatherings
Item 14	Less productive
Item 15	Lack energy
Item 16	Not physically active
Item 17	Unable to walk far/quickly
Item 18	Worried about health
Item 19	Self-conscious
Item 20	Frustrated or upset about weight

Note: The Physical composite score comprises Items 1-5, 16, and 17. The Physical Function composite comprises Items 1-3, 16, and 17, with Item 4 (Uncomfortable in small seats) and Item 5 (Bodily pain) excluded. The Psychosocial composite comprises Items 6-15 and 18-20.

Final IWQOL-Lite-CT Measure

- Items that were essentially redundant with others were removed from the scale to reduce patient burden without sacrificing content validity.
- The final 20-item IWQOL-Lite-CT includes two primary domains: Physical (7 items) and Psychosocial (13 items) (Table 3).
  - Based on feedback from the FDA and to facilitate labeling in the US, a 5-item subset of the Physical domain—the Physical Function composite—was also developed and supported.
- In addition, the IWQOL-Lite-CT was evaluated with Spanish speakers in Study 2 (data not shown), and the results were generally satisfactory for the small sample.

Conclusion

- The IWQOL-Lite-CT is a reliable, valid, and responsive measure of weight-related functioning in the populations commonly targeted for clinical trials of new weight loss medications.
- The Physical Function scale may be particularly appropriate to support product labeling based on the proximal nature of changes in the underlying construct to changes in patients' weight.

Contact Information

Ronette L Kolotkin, PhD
Quality of Life Consulting
rkolotkin@qualityoflifeconsulting.com

These trials were sponsored by Novo Nordisk and are registered with ClinicalTrials.gov (NCT02453711 and NCT02906930).
The authors acknowledge the medical writing assistance of Kate Lothman.
Presented at Obesity Week, November 11-15, 2018, Nashville, TN, United States.

References: (**1**) Kolotkin RL, Fujioka K, Wolden ML, Brett J, Bjørner JB. Improvements in health-related quality of life with liraglutide compared with placebo in weight management. Clin Obes. 2016;6(4):233-42. (**2**) Kolotkin RL, Gadde KM, Peterson CA, Crosby RD. Health-related quality of life in two randomized controlled trials of phentermine/topiramate for obesity: what mediates improvement? Qual Life Res. 2016;25(5):1237-44. (**3**) Kolotkin RL, Chen S, Klassen P, Gilder K, Greenway F. Patient-reported quality of life in a randomized, placebo-controlled trial of naltrexone/bupropion for obesity. Clin Obes. 2015(5);5:237-44. (**4**) Kolotkin RL, Ervin CM, Meincke HH, Højbjerre L, Fehnel SE. Development of a clinical trials version of the Impact of Weight on Quality of Life-Lite questionnaire (IWQOL-Lite Clinical Trials Version): results from two qualitative studies. Clin Obes. 2017 Oct;7(5):290-9. (**5**) Food and Drug Administration. Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims. December 2009. Available at: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf. Accessed August 22, 2018. (**6**) Cronbach L. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;16:294-334. (**7**) Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale NJ: Lawrence Erlbaum Associates; 1988.