ORIGINAL RESEARCH ARTICLE

Clinical OBESITY   WILEY

The Norse Feedback in a population of patients undergoing bariatric surgery—Psychometric properties of a digital computer-adaptive questionnaire assessing mental health

Pål A. Hegland[1,2] | Andrew McAleavey[3,4] | Anny Aasprang[1] | Christian Moltu[5] | Ronette L. Kolotkin[1,6,7] | John Roger Andersen[1,4]

[1]Faculty of Health and Social Sciences, Western Norway University of Applied Sciences, Førde, Norway

[2]Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway

[3]Department of Psychiatry, Weill Cornell Medical College, New York, New York, USA

[4]Centre of Health Research, Førde Hospital Trust, Førde, Norway

[5]Department of Psychiatry, District General Hospital of Førde, Førde, Norway

[6]Department of Family Medicine and Community Health, Duke University School of Medicine, Durham, North Carolina, USA

[7]Quality of Life Consulting, PLLC, Durham, North Carolina, USA

Correspondence
Pål A. Hegland, Faculty of Health and Social Sciences, Western Norway University of Applied Sciences, Førde, Norway.
Email: pal.andre.hegland@hvl.no

Summary

The Norse Feedback (NF) is a questionnaire developed for patient-reported outcome monitoring with a clinical feedback system (PRO/CFS). As mental health is a concern after bariatric surgery, the use of the NF as part of PRO/CFS may be beneficial. The aim of this study is to test the reliability and validity of the NF in patients who have been accepted for or have undergone bariatric surgery. We performed separate robust confirmatory factor analyses (CFAs) to test the unidimensionality on 19 of the NF scales. We also performed correlation analyses on 19 of the NF scales with the Obesity-related Problems scale (OP). We included 213 patients. In the CFA analyses, three out of 12 scales with four or more items showed satisfactory psychometric properties in all goodness of fit indices (*Suicidality*, *Need for Control* and *Self-Criticism*). Four scales showed satisfactory psychometric properties in all indices but RMSEA (*Somatic Anxiety*, *Substance Use*, *Social Safety* and *Cognitive Problems*). Several of the scales demonstrated floor effects. In the correlation analyses, 18 of the 19 scales showed small-to-moderate correlation coefficients with the OP. Our demonstration of satisfactory psychometric properties on several important scales of the NF suggests that this tool may prove valuable in the routine follow-up of mental health in this population. However, further work is needed to innovate the NF for patients undergoing bariatric surgery.

KEYWORDS
bariatric surgery, clinical feedback system, mental health, patient-reported outcome measures

1 | BACKGROUND

Following bariatric surgery, the majority of patients experience sustainable weight loss and significant improvements to physical impairment, obesity-related comorbidities and health-related quality of life (HRQOL).[1–5] Furthermore, up to 80% of patients have reported

being satisfied with their overall treatment satisfaction 5 years after surgery.[6] Although bariatric surgery is followed by positive outcomes for most patients, a cohort study of almost 25 000 patients who had undergone bariatric surgery in Australia found an increase in psychiatric illness presentations after surgery, as well as increased occurrences of deliberate self-harm and suicidal deaths.[7] These findings are supported

in a recent systematic review exploring the risk of suicide and self-harm after bariatric surgery.[8]

The awareness of impaired mental health has led to an increased focus on evaluating such concerns after bariatric surgery. Mental illness preoperatively has been associated with lower total weight loss and the risk of increased weight regain in patients following Roux-en-Y gastric bypass.[9] Furthermore, mental illness can contribute to a higher risk of early readmission after bariatric surgery.[10,11] A review article found that psychological factors like depression and anxiety may increase the risk of drop-out from follow-up consultations,[12] which is disturbing, as lifelong follow-up after bariatric surgery is recommended.[13,14] This knowledge has led to changes to the guidelines for post-bariatric surgery follow-up, most recently to the British guidelines.[15] In the British guidelines, a preoperative and postoperative mental health evaluation to identify those patients in need of psychological follow-up is recommended. For patients in need of mental health treatment, the follow-up should be with trained personnel, preferably a psychologist.[15]

As a way to organize a person-centred follow-up after bariatric surgery, patient-reported outcome monitoring with a clinical feedback system (PRO/CFS) appears promising,[16–18] in particular regarding psychological health. A PRO/CFS is a system in which the patient completes various digital questionnaires prior to the consultation with a healthcare professional and their responses form the basis for the content of the communication between the patient and healthcare professional during the consultation.[16] This enables the healthcare professional to start the consultation by addressing what the patient has reported as most bothersome. The use of a digital PRO/CFS to provide an instant analysis of the patient's responses and present them in a visual report has been found to be advantageous for the clinical consultation[16] and the possibility to adapt the consultation to the patient's needs has been reported to be important to patients.[19]

The Norse Feedback (NF) is a multidimensional questionnaire developed for mental health treatment in Norway.[20] The system was developed for a PRO/CFS and is computer-adaptive in that it automatically selects the items relevant to the individual patient based on what he/she answered to a set of trigger questions following an initial administration of the full item set.[20] Traditional questionnaires developed with a nomothetic approach have been criticized for omitting important concerns experienced by many patients and for lacking relevance to particular patients because they are aimed at capturing problematic issues across the entire population.[21] An ideographic adaptation in which the questionnaire is adapted to the patient's situation has been recommended.[22] The NF contains a broad measure of symptoms and processes, many of which may be relevant for patients who have undergone bariatric surgery, such as suicidal ideation and substance use. However, which of these scales, if any, is most appropriate for use in bariatric surgery patients is not yet clear because the NF has thus far only been used and evaluated in mental and behavioural health settings.[20]

The aim of this study was to test the internal consistency and construct validity of the scales in the NF in a population of patients who had been accepted for or already undergone bariatric surgery. More specifically, we hypothesized that (a) the individual scales in the NF would demonstrate satisfactory internal consistency; (b) the items in each scale would demonstrate sufficient unidimensionality. To explore the relationship between the Obesity-related Problems scale (OP) and the NF we hypothesized that the pattern of correlations with the NF scales would be moderate and positive for scales that represent common mental health issues (such as Sad Affect and Somatic Anxiety), and food and socially relevant scales. This is consistent with previous research demonstrating that individuals with OP tend to experience a wide range of psychological issues. We anticipated lower correlations between the OP and less typically relevant scales of the NF.

2 | METHODS

2.1 | Ethical considerations

The study was approved by the Norwegian Centre for Research Data, Department of Data Protection Services (reference number 282738). All participants signed an informed consent form and were recruited to the study by an outpatient nurse. The PRO/CFS is integrated as a standard part of the consultation, so there was no additional burden on those patients who were willing to participate. The design of the implemented PRO/CFS is described elsewhere.[23]

2.2 | Population

Patients who were accepted for or had already undergone bariatric surgery were eligible for inclusion in the study. The criteria for undergoing bariatric surgery were a body mass index (BMI) of over 40 kg/m^2 or 35–39.9 kg/m^2 with obesity-related comorbidities. When accepted for surgery, the patient attends a preoperative consultation, and follow-up at 3, 12, 24, 60 and 120 months after surgery. Each time the patient completes the PRO/CFS prior to the consultation at a bariatric outpatient clinic.

The surgical population had undergone laparoscopic sleeve gastrectomy (LSG) or biliopancreatic diversion with duodenal switch (BPD/DS). The patients had their consultation in the bariatric outpatient clinic between February 2018 and January 2020. For patients with several consultations during the course of the study, we included only their first consultation in our analyses (the first consultation could be at any follow-up consultation). All patients met with an outpatient nurse.

2.3 | Outcome measures

The NF and the OP were a part of the PRO/CFS that the patients responded to prior to the consultations and were thereby administered at the same time.

2.3.1 | Mental health

The NF is a multidimensional questionnaire developed for use with a digital PRO/CFS. The NF version 2.0 consists of 102 potential items that load on 23 different scales in four domains, three treatment process scales and four single items that do not load on any scales (see Table 1).[20] Twelve of the NF scales consist of four or more items, whereas seven scales consist of three items. Higher scores indicate a higher symptom load. The scale composition of the NF is the result of an iterative process involving patient and provider feedback, as well as quantitative scale development methods. Since the NF scales are designed to be used somewhat independently and the full NF is rarely administered after the initial completion, we elected to maintain the scale structure of the existing NF, treating each scale of the NF as an independent measure. In other words, we do not conduct analyses of all items in the NF as a whole but treat each existing scale as an independent measurement. This is consistent with the development process of the NF in mental health settings.[20] After an initial complete assessment in mental-health settings, some NF scales are automatically skipped for certain patients when they are algorithmically determined to be likely not important for care. The specifics of that algorithm are not relevant in this case since patients only completed the measure one time and so could reasonably complete all scales. In this project, we applied 20 scales, four single items and one two-item scale. For the *Substance Recovery* scale, we did not perform a CFA analysis, as only 20 patients had responded to this scale. The *Substance Recovery* scale only opens when the *Substance Use* scale is elevated above an empirical threshold. After considerations in a user-panel, including two patients and a clinician, we excluded the scales *Alliance* and *Needs in treatment*, as the items in these scales are specific to mental health treatment and not directly relevant for patients undergoing obesity treatment.[23] Because the NF scales are designed to be used independently and in one score per scale, exclusion of these scales for research purposes should not affect the validity of the questionnaire.[20]

The items have seven-point response categories, ranging from *This is completely true for me* to *This is not at all true for me*, as well as the response choice *Not relevant/Do not know/Refuse to answer*.

When a patient completes the NF for the first time, the patient responds to all items. If the patient's scale score falls under the defined threshold for that scale, the patient will only answer one trigger item from that scale the next time he or she is administered the NF. This ensures an individualization of the questionnaire, as the patients are not asked items that are likely to be irrelevant to them. However, if the patient's response later switches to an empirically predefined threshold on the trigger item, the scale will re-open to its full length. This is considered a key feature in idiographic adaptation.[22] The assessment of the adaptive part of the NF does not fall within the scope of this study and will be performed at a later stage.

2.3.2 | Obesity-specific quality of life

The OP scale is a questionnaire developed to measure the psychosocial consequences of obesity.[24] The questionnaire has been validated

TABLE 1 Structure of domains, scales and number of items included in each scale in the Norse Feedback questionnaire

Domains and scales included in the Norse Feedback	Number of items
Symptom Expression	
Eating Problems	5
Sad Affect	3
Somatic Anxiety	5
Substance Use	4
Suicidality	4
Trauma Reactions	4
Resources	
Readiness for Recovery	3
Recovery Environment	5
Social Safety	6
Problem-maintaining Processes	
Need for Control	4
Hopelessness	5
Internal Avoidance	5
Irritability	3
Self-Criticism	8
Situational Avoidance	3
Social Avoidance	3
Worry	3
Personal Consequences	
Cognitive Problems	6
General Functioning	3
Substance Recovery	4
Treatment Process Scales	
Alliance[a]	4
Needs in treatment[a]	5
Medication	2
Individual Items	
I take care of my physical health	
I am sleeping very badly at the moment	
My sexuality and/or sex life is difficult for me	
I am impulsive in a way that troubles me	

[a]Scales not included in the clinical feedback system, as the items were not considered relevant for follow-up to bariatric surgery.

for use in Norway,[25] as well as validated for both a surgical and non-surgical population of people with obesity.[24] The questionnaire consists of eight items that measure the impact of obesity on social activities, public activities and intimate relations on a four-point scale (*definitely bothered*, *mostly bothered*, *not so bothered* and *definitely not bothered*). Lower scores indicate fewer psychosocial consequences of obesity. In research, the score is transformed to a 0–100 scale, where a lower score indicates better psychosocial functioning. In clinical consultations, the individual items provide the most useful

information. The OP has been shown to correlate positively with anxiety, depression and overall mood measured using the Hospital Anxiety and Depression Scale (HADS).[24]

2.4 | Analyses

Demographic data are presented in terms of frequency and percentages (%) for dichotomous data and mean with standardized deviation (SD) and median for continuous data.

The NF scales with four or more items were examined for structural validity (unidimensionality) using robust confirmatory factor analysis (CFA) for ordinal data, with weighted least squares and mean and variance adjustment (WLSMV) estimation, using the R package lavaan (version 0.6–7).[26] Patients who did not complete all items on the given scale were excluded from the analyses because the WLSMV estimator in lavaan does not handle missing values. Factor loadings ≥0.4 were considered satisfactory. We evaluated goodness of fit using the root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), the comparative fit index (CFI), and the Tucker-Lewis index (TLI). Model fit was indicated by RMSEA close to or below 0.06, SRMR close to or below 0.08, and CFI as well as TLI close to or above 0.95.[27,28] To evaluate the internal consistency of scale items, factor loadings from the CFA were used to calculate composite reliability (CR) values for the NF scales, with scores ≥0.7 being considered satisfactory.[29] CR is a CFA-based estimate of reliability and is not biassed as the more common Cronbach's alpha. Nevertheless, we also calculated Cronbach's alpha, as this is the most reported analysis for internal consistency with a threshold of 0.7 or higher considered as satisfactory,[30] and report mean inter-item correlation, where ideal range is considered between 0.2 and 0.4.[31] Thus, a total of at least 11 indices were compared for each scale with four or more items: three internal consistency metrics (CR, alpha, mean inter-item correlation), four CFA model fit indices, and factor loadings for each item. For scales with three items, we do not report model fit indices as they are uninformative, but the other seven indices were computed.

SPSS version 26 was used for other analyses. Hypothesis (c) was used to study the relationship between the OP and the NF, and we performed correlation analyses of the 19 different scales, as well as the four single-item indices and the one two-item index in the NF and the correlation to the total score of the OP, as the OP is designed to be used as a unidimensional measure. Given the large number of pairwise comparisons, we interpret effect sizes rather than statistical significance. Pearson's correlation coefficients < 0.1were considered trivial, 0.1–0.29 as small, 0.30–0.49 as moderate and ≥0.5 as strong.[32] SPSS version 26 was used for statistical analyses[33] and exact two-sided p-values reported.

3 | RESULTS

Out of 259 eligible patients, 213 (82%) are included in the analyses. The included patients had a mean age of 41.1 years, a mean BMI

TABLE 2 Demographic information on the 213 patients included in the analyses

Demographic	Included in the analyses (n = 213)
Age	
Mean (SD)	41.1 (10.7)
Median	41.0
Biological sex (female), n (%)	162 (76.1)
BMI before surgery	
Mean (SD)	43.5 (5.5)
Median	42.3
Type of surgery, n (%)	
LSG	198 (93.0)
BPD/DS	14 (6.6)
Depression[a], n (%)	27 (12.7)
Marital status, n (%)	
Married/cohabitant	134 (64.1)
Single	66 (31.6)
Divorced	8 (3.8)
Widow/widower	1 (0.5)
Education, n (%)	
University college or higher	67 (33.0)
Highschool or primary school	136 (67.0)
Time point for follow-up, n (%)	
Preoperative	44 (20.7)
3 months	27 (12.7)
12 months	22 (10.3)
24 months	23 (10.8)
60 months	53 (24.9)
120 months	44 (20.7)

Abbreviations: BDP/DS, biliopancreatic diversion with duodenal switch; LSG, laparoscopic sleeve gastrectomy; n, number. SD, standard deviation.
[a]The patients are asked whether they are undergoing treatment for depression.

before surgery of 43.5 kg/m^2, and 93% of the patients had undergone LSG. The majority of the patients (76.1%) were women and 27% of the patients reported receiving treatment for depression at the time of data collection. See Table 2 for demographic details on the patients included in the analyses. Two patients were excluded due to not giving consent to participate, whereas 44 patients were excluded because they had not completed the PRO/CFS when they met for their follow-up consultation. The patients included did not differ from those excluded from the study regarding age, biological sex, BMI before surgery or type of surgery. In the population not included in the analyses, more patients were divorced (11.4% vs. 3.8%), had undergone treatment for depression (20.5% vs. 12.7%), and fewer had education at university college level or higher (19.5% vs. 33.0%). In terms of the follow-up consultation, more patients in the population not included met for their 120-month follow-up (31.8% vs. 20.7%) and fewer for the preoperative consultation (15.9% vs. 20.7%).

For descriptive information on the scales and individual items of the NF in bariatric surgery follow-up, see Table 3. The analyses demonstrated that the floor effect ranged from 5.7% to 88.2% (with floor effects detected in the scales *Sad Affect*, *Somatic Anxiety*, *Substance Use*, *Suicidality*, *Trauma Reactions*, *Need for Control*, *Hopelessness*, *Irritability*, *Situational Avoidance*, *Social Avoidance*, *Worry*, *Cognitive Problems*, *General Functioning* and *Medication*, and the single items *Physical Health*, *Sleep*, *Sexuality*, and *Impulsivity*), whereas the ceiling effect ranged from 0% to 13.1%.

A summary of the psychometric properties of the NF is presented in Table 4.

The CFA analyses showed that three out of the 12 scales with four or more items had satisfactory psychometric properties in all indices (*Suicidality*, *Need for Control* and *Self-Criticism*). Six of the remaining scales showed satisfactory psychometric properties in all model fit indices but one (*Somatic Anxiety*, *Substance Use*, *Social Safety*, *Cognitive Problems*, *Trauma Reactions* and *Internal Avoidance*). The scale *Recovery Environment* showed CR 0.68 and Cronbach's alpha 0.54, while the scales *Eating Problems* and *Hopelessness* showed unsatisfactory Cronbach's alpha and factor loading below 0.4 in one item each.

The analyses of the scales with three items showed satisfactory CR, Cronbach's alpha and factor loadings above 0.4 in four out of seven scales (*Sad Affect*, *Situational Avoidance*, *Social Avoidance* and *Worry*). The scale *Readiness for Recovery* showed unsatisfactory CR, Cronbach's alpha and factor loading below 0.4 in one item, whereas the scales *Irritability* and *General Functioning* both showed unsatisfactory CR and Cronbach's alpha.

TABLE 3 Mean, standard deviations, median, minimum-maximum scores and floor and ceiling effects of the Norse Feedback

Scales	Mean (SD)	Median	Minimum-maximum score	Floor effect/ceiling effect %
Symptom Expression				
Eating Problems	2.38 (1.09)	2.2	1.00–6.00	10.5/0
Sad Affect	1.86 (1.29)	1.33	1.00–6.67	48.1/0
Somatic Anxiety	2.13 (1.27)	1.70	1.00–6.20	23.8/0
Substance Use	1.14 (0.60)	1.00	1.00–6.25	88.2/0
Suicidality	1.24 (0.68)	1.00	1.00–5.75	81.1/0
Trauma Reactions	1.81 (1.06)	1.25	1.00–6.00	41.0/0
Resources				
Readiness for Recovery	2.76(1.29)	2.67	1.00–7.00	13.0/0.5
Recovery Environment	2.31(1.09)	2.20	1.00–7.00	15.2/0.5
Social Safety	3.03 (1.29)	3.00	1.00–6.83	5.7/0
Problem-maintaining Processes				
Need for Control	2.42 (1.29)	2.25	1.00–6.50	17.2/0
Hopelessness	2.05(1.13)	1.75	1.00–7.00	25.4/0.5
Internal Avoidance	2.43(1.30)	2.20	1.00–6.40	13.8/0
Irritability	1.91 (0.92)	1.67	1.00–4.50	30.0/0
Self-Criticism	1.89 (1.03)	1.63	1.00–5.88	13.7/0
Situational Avoidance	1.67 (1.16)	1.00	1.00–7.00	51.7/0.5
Social Avoidance	2.22 (1.53)	1.67	1.00–7.00	35.5/1.4
Worry	2.54 (1.68)	2.00	1.00–7.00	28.4/1.4
Personal Consequences				
Cognitive Problems	2.16 (1.31)	1.67	1.00–6.50	22.7/0
General Functioning	2.50 (1.29)	2.33	1.00–6.33	20.0/0
Treatment Process Scale				
Medication	3.33(1.40)	4.00	1.00–7.00	18.7/1.6
Individual Items				
I take care of my physical health	2.98 (1.61)	3.00	1.00–7.00	25.5/0.5
I am sleeping very badly at the moment	3.15 (2.11)	3.00	1.00–7.00	36.0/10.3
My sexuality and/or sex life is difficult for me	2.89 (2.14)	2.00	1.00–7.00	40.7/13.1
I am impulsive in a way that troubles me	1.78 (1.45)	1.00	1.00–7.00	68.1/2.4

Abbreviation: SD, standard deviation.

TABLE 4 Psychometric properties of the 19 scales of the Norse Feedback

Norse Feedback scales, n = 213	Composite reliability	Cronbach's alpha	Mean inter-item correlation	Model fit indices	Factor loadings above 0.4
Eating Problems, n = 201	0.76	0.63	0.27	CFI = 0.99 TLI = 0.97 RMSEA = 0.07 (90% CI 0.00–0.13) SRMR = 0.05	Q46: 0.55 Q104: 0.76 Q57: 0.69 Q18: 0.69 Q63: 0.39
Sad Affect, n = 205	0.91	0.83	0.64	NA	In all items
Somatic Anxiety, n = 202	0.87	0.79	0.43	CFI = 0.99 TLI = 0.97 RMSEA = 0.11 (90% CI 0.05–0.17) SRMR = 0.04	In all items
Substance Use, n = 180	0.98	0.90	0.74	CFI = 1.00 TLI = 0.99 RMSEA = 0.13 (90% CI 0.04–0.23) SRMR = 0.02	In all items
Suicidality, n = 197	0.94	0.80	0.54	CFI = 1.0 TLI = 1.0 RMSEA = 0 (90% CI 0.00–0.13) SRMR = 0.02	In all items
Trauma Reactions, n = 186	0.83	0.67	0.35	CFI = 1.0 TLI = 1.0 RMSEA = 0 (90% CI 0.00–0.14) SRMR = 0.02	In all items
Readiness for Recovery, n = 181	0.64	0.40	0.20	NA	Q135: 0.19 Q134: 0.53 Q133: 1.00
Recovery Environment, n = 186	0.68	0.54	0.20	CFI = 0.98 TLI = 0.96 RMSEA = 0 (90% CI 0.00–0.12) SRMR = 0.04	In all items
Social Safety, n = 199	0.82	0.74	0.33	CFI = 0.95 TLI = 0.92 RMSEA = 0.13 (90% CI 0.09–0.18) SRMR = 0.06	In all items
Need for Control, n = 201	0.76	0.70	0.38	CFI = 0.99 TLI = 0.97 RMSEA = 0.08 (90% CI 0.00–0.18) SRMR = 0.03	In all items
Hopelessness, n = 182	0.83	0.66	0.32	CFI = 0.98 TLI = 0.96 RMSEA = 0.13 (90% CI 0.07–0.19) SRMR = 0.06	Q61: 0.82 Q15: 0.07 Q115: 0.79 Q24: 0.84 Q88 0.83
Internal Avoidance, n = 191	0.82	0.75	0.37	CFI = 0.99 TLI = 0.98 RMSEA = 0.07 (90% CI 0.00–0.14) SRMR = 0.03	Q34: 0.76 Q78: 0.67 Q122: 0.77 Q123: 0.82 Q10: 0.38
Irritability, n = 199	0.64	0.40	0.20	NA	In all items
Self-Criticism, n = 193	0.93	0.86	0.46	CFI = 0.99 TLI = 0.99 RMSEA = 0.06 (90% CI 0.02–0.09) SRMR = 0.05	In all items
Situational Avoidance, n = 200	0.81	0.74	0.49	NA	In all items

TABLE 4 (Continued)

Norse Feedback scales, n = 213	Composite reliability	Cronbach's alpha	Mean inter-item correlation	Model fit indices	Factor loadings above 0.4
Social Avoidance, n = 207	0.88	0.82	0.60	NA	In all items
Worry, n = 203	0.87	0.82	0.61	NA	In all items
Cognitive Problems, n = 197	0.93	0.88	0.56	CFI = 0.99 TLI = 0.98 RMSEA = 0.14 (90% CI 0.10–0.19) SRMR = 0.04	In all items
General Functioning, n = 204	0.63	0.51	0.26	NA	In all items

Note: Psychometric properties were considered satisfactory if the following criteria were fulfilled: Composite Reliability and Cronbach's alpha ≥0.7, CFI and TLI close to or above 0.95, RMSEA close to 0.08, SRMR < 0.06 and factor loadings above 0.4.
Abbreviations: CFI, comparative fit index; CI, confidence Interval; *N*, Number of patients included in the analyses; NA, not applicable; RMSEA, root mean squared error of approximation; SRMR, standardized root mean square residual; TLI, Tucker-Lewis Index.

TABLE 5 Pearson's correlation coefficient of the scales in the Norse Feedback and Obesity-related Problems scale

Norse Feedback scales	Obesity-related Problems scale Pearson's correlation coefficient (*p*-value)
Scales	
Eating Problems	0.33 (<0.001)
Sad Affect	0.33 (<0.001)
Somatic Anxiety	0.30 (<0.001)
Substance Use	0.03 (0.68)
Suicidality	0.26 (<0.001)
Trauma Reactions	0.21 (0.002)
Readiness for Recovery	0.21 (0.003)
Recovery Environment	0.29 (<0.001)
Social Safety	0.25 (<0.001)
Need for Control	0.35 (<0.001)
Hopelessness	0.30 (<0.001)
Internal Avoidance	0.38 (<0.001)
Irritability	0.16 (0.02)
Self-Criticism	0.39 (<0.001)
Situational Avoidance	0.45 (<0.001)
Social Avoidance	0.42 (<0.001)
Worry	0.24 (<0.001)
Cognitive Problems	0.14 (0.04)
General Functioning	0.29 (<0.001)
Two-item index	
Medication	−0.04 (0.70)
Single items	
I take care of my physical health	0.27 (<0.001)
I am sleeping very badly at the moment	0.24 (<0.001)
My sexuality and/or sex life is difficult for me	0.43 (<0.001)
I am impulsive in a way that troubles me	0.13 (0.06)

The analyses of the relationship between the NF and the OP showed moderate correlation for the scales *Somatic Anxiety*, *Sad Affect*, *Eating Problems*, *Social Avoidance*, *Situational Avoidance*, *Need for Control*, *Hopelessness*, *Internal Avoidance*, *Self-Criticism* and the single item *Sexuality*. Furthermore, small correlation was detected for the scales *Social Safety*, *Suicidality*, *Trauma Reactions*, *Readiness for Recovery*, *Recovery Environment*, *Irritability*, *Worry*, *General functioning* and *Cognitive Problems* and the single items *Physical Health*, *Sleep* and *Impulsivity*. Ultimately, the scales *Substance Use* and *Medication* showed trivial correlations. See Table 5 for further details on the correlation analyses.

4 | DISCUSSION

This is the first step in developing the NF for a population of patients undergoing bariatric surgery. In this study, the NF demonstrates satisfactory psychometric properties in several of the scales measuring domains that research have shown to be relevant for patients in follow-up after bariatric surgery (*Sad Affect*, *Suicidality*, *Situational Avoidance*, *Need for Control*, *Self-Criticism*, *Social Avoidance* and *Worry*).[7,9] These findings indicate that the NF may be a useful tool in the follow-up of patients after bariatric surgery.

Due to floor effects in 14 of the scales and all single items, this questionnaire should be adjusted to this population of patients. There were strong floor effects in several scales as, e.g., the scale *Sad Affect* and the single items of *Sleep* and *Sexuality*, suggesting that these items did not assess common concerns among bariatric patients. This deserves further scrutiny, as some issues related to these scales are clearly relevant to patients undergoing BS.[34,35] The NF was originally developed for patients receiving mental health treatment. Therefore, the high floor effects are likely a result of the population of patients undergoing BS having different characteristics and needs in treatment.[36] In a clinical setting with a population of patients undergoing bariatric surgery, several of these domains only provide important information for a small number of patients, but this is likely highly clinically relevant for those individuals. In contrast, the NF scales did not appear to have substantial ceiling effects, so there are few patients for whom the NF would fail to capture their clinical severity. This suggests that further adaptations of the NF for patients undergoing bariatric surgery should include fewer questions from scales with

large floor effects, while using the NF's adaptive capability to open additional relevant items for those patients who have related problems.

There are several mental health questionnaires that are often used in bariatric surgery care, such as the Beck Depression Inventory (BDI) and Hospital Anxiety and Depression Scale (HADS),[37] Binge Eating Scale (BES)[38] and Eating Disorder Examination Questionnaire (EDE-Q).[39] The NF was chosen in this project for several reasons, such as the fact that the NF is a multidimensional questionnaire that measures more aspects of mental health than the BDI or HADS. The various domains measured by the NF were considered relevant for this population of patients. Because the NF is designed for a digital PRO/CFS questionnaire, there is the possibility for idiographic adaptation and a digital summary report. The Minnesota Multiphasic Personality Inventory-3 (MMPI-3) is another multidimensional questionnaire assessing mental health and used for research in a bariatric surgery population[40]; however, the questionnaire does not yet have computer-adaptive capabilities.

Furthermore, the NF appears to be relevant for the group of patients undergoing bariatric surgery, as most scales of the NF demonstrate medium to small correlations with the well-used OP questionnaire. This indicates that patients experiencing greater problems with obesity also endorse higher distress across the broad range of scales on the NF. Importantly, these correlations are likely attenuated by the restricted range of the NF scale scores due to the high floor effect. The small–to-moderate correlations indicate that if a patient experiences impairment in mental health, the obesity-specific quality of life will also be affected. As examples, the *Situational Avoidance* and *Social Avoidance* scales in the NF demonstrate the highest correlation coefficients with the OP, which is natural, as the OP measures the negative psychosocial impacts of obesity, such as going to parties or public bathrooms. As the OP and NF measure different constructs, one would not expect perfect correlations between these two questionnaires, as evident with the scale *Substance Use*. The correlations of the *Sad Affect* and *Somatic Anxiety* scales in the NF and the total OP score should therefore increase the confidence in the reliability of the NF in this patient population.

Regarding unidimensionality and internal consistency, the scales of the NF showed overall positive results. The psychometric properties in some domains measuring constructs relevant for patients following bariatric surgery, such as *Eating Problems*,[41] still need to be refined to be maximally relevant for this population. This scale largely assesses restrictive eating and may be more appropriate for clinical settings where overly restrictive eating is more of a problem than in obesity surgery. Greater emphasis on symptoms of impulsive or binge eating may be preferable in future revisions. Furthermore, some scales (e.g., *Somatic Anxiety* and *Substance Use*) demonstrated satisfactory psychometric properties in all indices but the RMSEA. These scales require further investigation and development, since many of the NF scales measure constructs that are important to assess in this population.[42,43] We chose to apply the strict model fit cut-offs as proposed by Hu and Bentler, even though these have been critiqued for being too strict. Applying less strict cut-offs as CFI/TLI above 0.90 and

RMSEA below 0.8 would have resulted in the scales *Need for Control* and *Self-Criticism* showing satisfactory psychometric properties in all indices.

Patients highlight a wish for a structured follow-up of mental health concerns.[44] In our experience, the use of a digital PRO/CFS seems feasible in the follow-up process to bariatric surgery. As demonstrated, some patients may experience a variety of mental health challenges following surgery, whereas others experience only minor challenges post-surgery. The variety of challenges call for a broad assessment of the patients. This broad assessment often results in the administration of several questionnaires, which in turn may cause respondent burden. The advantage of a digital PRO/CFS is the possibility for an idiographic adaptation of the questionnaires.[16] After a broad assessment, the patient is asked questions that are likely to be relevant to his or her situation, as highlighted as important by both patients and healthcare professionals for the PRO/CFS to be experienced as useful in clinical consultations.[45] This reduces the risk of patients reading and responding to the items in an unfocused manner if they become tired of answering questions.

4.1 | Limitations of the study and future directions

A limitation to the study is that including more patients would have provided more robust estimates. In addition, the time since surgery varies among patients, from prior to surgery to 10 years after surgery. This resulted in a limited number of patients in each category, preventing us from performing analyses of reliability and validity at the different time points for the consultations. Only 20.7% of the patients in this sample were assessed during their pre-operative consultation, making the findings in this study less robust for patients who are awaiting bariatric surgery. Furthermore, more patients in the not-included sample had undergone treatment for depression. This might have affected the psychometric properties and especially the high floor effect on the scale *Sad Affect* as these patients' responses might had shown more diversity in symptoms for this scale.

Another limitation to the study is the lack of a gold standard questionnaire that measures mental health, thereby not enabling us to assess the criterion validity of the NF.[36]

In the current version of the NF, the Likert response *Not relevant/ Do not know/Refuse to Answer* is treated as one response and might have affected the psychometric performance. This Likert response is in the NF psychiatry version changed to *I prefer not to answer*, and *This is not relevant to me*, and will be changed accordingly in the new version of NF Obesity.

The NF needs further assessments of validity and reliability for use in bariatric surgery care. In a future study, we plan to use an iterative process to further develop the NF in bariatric surgery care.[23] For this iterative process, a panel of patients and healthcare professionals will be established to assess the face validity of the NF. To develop a new version of the NF for bariatric surgery care, larger data sets with greater diversity in patient presentations would be essential. This adaption of the questionnaire is important before the effectiveness of

a PRO/CFS can be explored. Future research on PRO/CFS methodology should focus on assessing whether a PRO/CFS is feasible in the follow-up after bariatric surgery, as this is still not established.

5 | CONCLUSION

The NF appears to be a feasible questionnaire in the follow-up after bariatric surgery. The NF shows satisfactory psychometric properties in scales measuring relevant constructs for patients after bariatric surgery. However, before the NF can be implemented more broadly in the follow-up after bariatric surgery, the questionnaire must be further refined, adapted and psychometrically tested for this population of patients.

CONFLICT OF INTEREST

C.M. is one of the developers of the NF and owns intellectual property to the system. R.K. is the owner of Quality of Life Consulting PLLC, and in that role receives royalties for IWQOL-Lite, IWQOL-Kids and IWQOL-Lite-CT. There is no commercial or financial interest related to the current project. The other authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

Pål A. Hegland, Anny Aasprang and John R. Andersen contributed to designing and implementing the intervention. Pål A. Hegland and John R. Andersen conducted the statistical analyses with the assistance of Andrew McAleavey. Pål A. Hegland wrote the first draft of the manuscript and all authors contributed to the manuscript revision and read and approved the submitted version.

ORCID

Pål A. Hegland https://orcid.org/0000-0001-6524-5744
John Roger Andersen https://orcid.org/0000-0001-6300-9086

REFERENCES

1. Courcoulas AP, King WC, Belle SH, et al. Seven-year weight trajectories and health outcomes in the longitudinal assessment of bariatric surgery (LABS) study. *JAMA Surg.* 2018;153(5):427-434.
2. Adams TD, Davidson LE, Litwin SE, et al. Weight and metabolic outcomes 12 years after gastric bypass. *N Engl J Med.* 2017;377(12): 1143-1155.
3. Jakobsen G, Småstuen M, Sandbu R, et al. Association of bariatric surgery vs medical obesity treatment with long-term medical complications and obesity-related comorbidities. *JAMA.* 2018;319(3):291-301.
4. Kolotkin RL, Kim J, Davidson LE, Crosby RD, Hunt SC, Adams TD. 12-year trajectory of health-related quality of life in gastric bypass patients vs. comparison groups. *Surg Obes Relat Dis.* 2018;14(9):1359-1365.
5. Panagiotou OA, Markozannes G, Adam GP, et al. Comparative effectiveness and safety of bariatric procedures in Medicare-eligible patients: a systematic review. *JAMA Surg.* 2018;153(11):e183326-e.
6. Hegland PA, Aasprang A, Kolotkin RL, Tell GS, Andersen JR. Overall treatment satisfaction 5 years after bariatric surgery. *Obes Surg.* 2020; 30(1):206-213.

7. Morgan DJR, Ho KM, Platell C. Incidence and determinants of mental health service use after bariatric surgery. *JAMA Psychiatry.* 2019;77:60.
8. Castaneda D, Popov VB, Wander P, Thompson CC. Risk of suicide and self-harm is increased after bariatric surgery—a systematic review and meta-analysis. *Obes Surg.* 2019;29:1-12.
9. Müller A, Hase C, Pommnitz M, de Zwaan M. Depression and suicide after bariatric surgery. *Curr Psychiatry Rep.* 2019;21(9):84.
10. Jalilvand A, Dewire J, Detty A, Needleman B, Noria S. Baseline psychiatric diagnoses are associated with early readmissions and long hospital length of stay after bariatric surgery. *Surg Endosc.* 2019; 33(5):1661-1666.
11. Litz M, Rigby A, Rogers AM, Leslie DL, Hollenbeak CS. The impact of mental health disorders on 30-day readmission after bariatric surgery. *Surg Obes Relat Dis.* 2018;14(3):325-331.
12. Hood MM, Corsica J, Bradley L, Wilson R, Chirinos DA, Vivo A. Managing severe obesity: understanding and improving treatment adherence in bariatric surgery. *J. Behav. Med.* 2016;39(6):1092-1103.
13. Busetto L, Dicker D, Azran C, et al. Practical recommendations of the obesity management task force of the European Association for the study of obesity for the post-bariatric surgery medical management. *Obes Facts.* 2017;10(6):597-632.
14. O'Kane M, Parretti HM, Hughes CA, et al. Guidelines for the follow-up of patients undergoing bariatric surgery. *Clin Obes.* 2016;6(3): 210-224.
15. Ogden J, Ratcliffe D, Snowdon-Carr V. British Obesity Metabolic Surgery Society endorsed guidelines for psychological support pre- and post-bariatric surgery. *Clin Obes.* 2019;9(6):e12339.
16. Boswell JF, Kraus DR, Miller SD, Lambert MJ. Implementing routine outcome monitoring in clinical practice: benefits, challenges, and solutions. *Psychother Res.* 2015;25(1):6-19.
17. Carlier IV, Meuldijk D, Van Vliet IM, Van Fenema E, Van der Wee NJ, Zitman FG. Routine outcome monitoring and feedback on physical or mental health status: evidence and theory. *J Eval Clin Pract.* 2012; 18(1):104-110.
18. Hegland PA, Aasprang A, Hjelle Øygard S, et al. A review of systematic reviews on the effects of patient-reported outcome monitoring with clinical feedback systems on health-related quality of life—implications for a novel technology in obesity treatment. *Clin Obes.* 2018;8(6):452-464.
19. Börjesson S, Boström PK. "I want to know what it is used for": Clients' perspectives on completing a routine outcome measure (ROM) while undergoing psychotherapy. *Psychother Res.* 2019;30:1-11.
20. McAleavey A, Nordberg S, Moltu C. Initial quantitative development of the Norse Feedback system: a novel clinical feedback system for routine mental healthcare. *Qual Life Res.* 2021;107:2411.
21. Sales CM, Neves IT, Alves PG, Ashworth M. Capturing and missing the patient's story through outcome measures: a thematic comparison of patient-generated items in PSYCHLOPS with CORE-OM and PHQ-9. *Health Expect.* 2018;21(3):615-619.
22. Sales CM, Alves PC. Patient-centered assessment in psychotherapy: a review of individualized tools. *Clin Psychol.* 2016;23(3):265-283.
23. Hegland PA, Aasprang A, Kolotkin RL, Moltu C, Tell GS, Andersen JR. A novel patient-reported outcome monitoring with clinical feedback system in bariatric surgery care: study protocol, design and plan for evaluation. *BMJ Open.* 2020;10(6):e037685.
24. Karlsson J, Taft C, Sjostrom L, Torgerson JS, Sullivan M. Psychosocial functioning in the obese before and after weight reduction: construct validity and responsiveness of the obesity-related problems scale. *Int J Obes.* 2003;27(5):617-630.
25. Aasprang A, Andersen JR, Våge V, Kolotkin RL, Natvig GK. Psychosocial functioning before and after surgical treatment for morbid obesity: reliability and validation of the Norwegian version of obesity-related problem scale. *PeerJ.* 2015;3:e1275.
26. Rosseel Y. Lavaan: an R package for structural equation modeling. *J. Stat. Softw.* 2012;48(2):1-36.

27. Brown TA. Confirmatory Factor Analysis for Applied Research: Guilford Publications; 2015.

28. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling*. 1999;6(1):1-55.

29. Raykov T. Estimation of composite reliability for congeneric measures. *Appl. Psychol. Meas.* 1997;21(2):173-184.

30. Bland JM, Altman DG. Statistics notes: Cronbach's alpha. *BMJ*. 1997;314(7080):572.

31. Piedmont RL, Hyland ME. Inter-item correlation frequency distribution analysis: a method for evaluating scale dimensionality. *Educ Psychol Meas*. 1993;53(2):369-378.

32. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates; 1988.

33. IBM Corp. *IBM SPSS Statistics for Windows, Version 24.0.* . IBM Corp; 2016.

34. Mechanick JI, Apovian C, Brethauer S, et al. Clinical practice guidelines for the perioperative nutrition, metabolic, and nonsurgical support of patients undergoing bariatric procedures – 2019 update: cosponsored by American Association of Clinical Endocrinologists/American College of Endocrinology, The Obesity Society, American Society for Metabolic & Bariatric Surgery, Obesity Medicine Association, and American Society of Anesthesiologists. *Surg Obes Relat Dis*. 2020;16(2):175-247.

35. Çaynak S, Boyacıoğlu NE, Temel M. Body perception and sexuality of bariatric surgery patients. *Perspect Psychiatr Care*. 2021;57(3):1266-1272.

36. HCW DV, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine*. Cambridge University Press; 2011.

37. Spirou D, Raman J, Smith E. Psychological outcomes following surgical and endoscopic bariatric procedures: a systematic review. *Obes Rev*. 2020;21(6):e12998.

38. Grupski AE, Hood MM, Hall BJ, Azarbad L, Fitzpatrick SL, Corsica JA. Examining the binge eating scale in screening for binge eating disorder in bariatric surgery candidates. *Obes Surg*. 2013;23(1):1-6.

39. Luce KH, Crowther JH. The reliability of the eating disorder examination—self-report questionnaire version (EDE-Q). *Int J Eat Disord*. 1999;25(3):349-351.

40. Marek RJ, Martin-Fernandez K, Heinberg LJ, Ben-Porath YS. An investigation of the eating concerns scale of the Minnesota multiphasic personality inventory–3 (MMPI-3) in a postoperative bariatric surgery sample. *Obes Surg*. 2021;31(5):2335-2338.

41. Dawes AJ, Maggard-Gibbons M, Maher AR, et al. Mental health conditions among patients seeking and undergoing bariatric surgery: a meta-analysis. *JAMA*. 2016;315(2):150-163. 14p.

42. Kalarchian MA, Marcus MD. Psychosocial concerns following bariatric surgery: current status. *Curr Obes Rep*. 2019;8(1):1-9.

43. King WC, Chen J-Y, Courcoulas AP, et al. Alcohol and other substance use after bariatric surgery: prospective evidence from a US multicenter cohort study. *Surg Obes Relat Dis*. 2017;13(8):1392-1402.

44. 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(1):88-107.

45. Philpot LM, Barnes SA, Brown RM, et al. Barriers and benefits to the use of patient-reported outcome measures in routine clinical care: a qualitative study. *Am J Med Qual*. 2018;33(4):359-364.