

## Supplementary file 1 Patient-reported outcome measures included in the PRO/CFS

### ***Obesity-specific HRQOL***

The Obesity-related Problems Scale (OP) is an 8-item HRQOL questionnaire measuring how obesity affects psychosocial functioning. It includes eight aspects of psychosocial functioning with a 4-point scale that has response categories ranging from *Definitely not bothered* to *Definitely bothered*. The scores range from 0 to 100, where lower scores indicate better psychosocial functioning. The eight individual items are useful in the clinical conversation with the patient. The questionnaire has been used in a Norwegian setting. Based on recommendations from the developer, cut-off levels for the summary report were set as follows: A score under 19 was designated green, 20-59.9 yellow, and 60-100 red. (Karlsson, J., Taft, C., Sjostrom, L., Torgerson, J. S., & Sullivan, M. (2003). Psychosocial functioning in the obese before and after weight reduction: construct validity and responsiveness of the Obesity-related Problems scale. *International Journal of Obesity*, 27(5), 617-630).

The Patient Reported Outcomes in Obesity (PROS) is a 10-item questionnaire measuring the impact obesity has on a person's daily life. It was developed and validated in a Norwegian setting. The questionnaire has eight questions on daily activities and two questions on the consequences of bariatric surgery, such as negative side effects and excess skin. The response categories are rated on a 4-point scale from *Considerably bothered* to *Not bothered*. The total score can be used to discuss overall obesity-specific HRQOL; however, the individual scores are preferred for use in clinical consultation as they give more specific information about the source of the negative impact of obesity. In addition to the 10 items measuring HRQOL, the questionnaire has one question on social support and one on overall treatment satisfaction, both measured on 4-point scales, with the response categories *Very satisfied*, *Satisfied*, *Unsure* and *Dissatisfied*. In the summary report each item is presented with set cut-off values as follows: *not bothered* green, *mildly bothered* yellow, and *moderately* and *considerably bothered* as red.

### ***Generic HRQOL***

The Short Form-36 (SF-36), Norwegian version 1.2, is a 36-item questionnaire measuring HRQOL, which is widely used for research on obesity. The questionnaire has eight dimensions of physical and mental functioning, with a total score ranging from 0 to 100,

where higher scores indicate better HRQOL. The dimensions are divided in two summary-scores — the Physical Component Score (PCS) and the Mental Component Score (MCS) — based on factor analysis with oblique rotation. For the analyses of both the sub-scores and summary scores, we used T-scores where a score of 50 approximates the average of the general population. Thresholds for the cut-off values in the summary report were set according to the norm-score (T-score 50), and previous research on clinical significance of impairment. This resulted in cut-off values as follows: Over 45 was designated as green, 42.1-45.0 yellow and under 42.0 red. (Loge et al., 1998).

### ***Mental health***

Norse Feedback (NF) is a multidimensional computer-adaptive questionnaire developed primarily to assess PROMS for use in mental health treatment (20). Structured interviews of the needs of patients and clinicians in treatment for mental health disorders were the basis for the development of Norse Feedback. Responding to clinicians' and patients' needs, Norse Feedback aims to measure discrete clinical phenomena, such as sad affect, rumination and interpersonal problems, for use in clinical conversations. In turn, multidimensional patterns of scores on discrete scales form hypotheses of higher order constructs, such as depression in a diagnostic context. Moreover, the measure adapts to the individual's presentation after an initial broad screening.

In addition to the factorial structure of the measure, information values within constructs for individual items in the questionnaire are evaluated through item-response theory (30, 31). The questionnaire consists of 93 items assessing symptoms related to mental health and addiction. Norse Feedback assesses 20 different dimensions in four domains (symptom expression, dysfunctional processes, functional consequences and resources). The questions use 7-point response categories, which range from *Is not correct for me at all* to *Is correct for me*; the patient can also choose the response *Not relevant/do not know/refuse to answer*.

The visual report of the NF is presented as raw scores in the obesity setting. Visually the categories green, yellow and red are carried over from a mental health treatment setting. In mental health the Y-axis presents standard deviations from a norm population, matching the colour categories. In this project, the presentation as raw scores in the summary report was chosen to represent face valid information for the healthcare professionals in the exploratory phase. Through the current project, data for establishing new standard deviations for a patient

population after bariatric surgery will be collected, and a presentation where Y-axis and colour categories match will result.

### ***Eating self-efficacy***

The Weight Efficacy Lifestyle Questionnaire - Short Form (WEL-SF) is a questionnaire that measures the confidence patients have in their ability to resist overeating in various situations. There are eight questions that are rated on a 10-point scale, from *Not confident* to *Very confident* about their eating self-efficacy, with a total score ranging from 0 to 80. Higher scores indicate better eating self-efficacy. The questionnaire is used for patients after bariatric surgery in Norway. Based on clinical judgement, we defined the following thresholds in the summary report: 70-80 green, 60-69.9 yellow, and under 60 red.

### ***Bowel symptoms***

The Gastrointestinal Symptoms Rating Scale (GSRS) measures gastrointestinal symptoms relevant to patients after bariatric surgery. The questionnaire measures 15 bowel symptoms on a 7-point scale (ranging from 1 to 7), with response categories ranging from *Not bothered at all* to *Severely bothered*. The 15 symptoms can be interpreted as 5 dimensions; *Abdominal pain*, *Reflux*, *Indigestion*, *Diarrhoea* and *Constipation*, or a total score that ranges from 15 to 105. The questionnaire has been used for research on patients after bariatric surgery in Norway. For the colour categories in the summary report we set the following thresholds for cut-off of the five dimensions, based on clinical judgement: 3-6 green, 9-12 yellow, and 15-21 red.