



**Title:** Health-related quality of life following bariatric surgery: a systematic review of prospective long-term studies

John Roger Andersen <sup>1,2</sup>, Anny Aasprang <sup>2</sup>, Tor-Ivar Karlsen <sup>3,4</sup>, Gerd Karin Natvig <sup>2,7</sup>, Villy Våge <sup>1</sup>,  
Ronette L. Kolotkin <sup>1,2,3,5,6</sup>

<sup>1</sup> Department of Surgery, Førde Central Hospital, Norway

<sup>2</sup> Faculty of Health Studies, Sogn og Fjordane University College, Førde, Norway

<sup>3</sup> Morbid Obesity Centre, Vestfold Hospital Trust, Tønsberg, Norway

<sup>4</sup> Department of Health and Nursing Sciences, University of Agder, Grimstad, Norway

<sup>5</sup> Quality of Life Consulting, Durham, NC, US

<sup>6</sup> Duke University School of Medicine, Durham, NC, US

<sup>7</sup> Department of Public Health and Primary Health Care, University of Bergen, Norway

Corresponding author: John Roger Andersen. Department of surgery, Førde Central Hospital, Førde.

Vievegen 2, 6807 Førde. Norway. Telephone: 482 178 186. E-mail: johnra@hisf.no

**Short title:** Quality of life following bariatric surgery

**Abstract**

Impaired health-related quality of life (HRQoL) is common in bariatric surgery candidates and is often one of the motivating factors for seeking bariatric surgery. Although many studies have reported changes in HRQoL after bariatric surgery, few are long-term prospective studies and no systematic review has been conducted. A systematic database search identified studies reporting HRQoL preoperatively and  $\geq 5$  years after bariatric surgery. Change in HRQoL over time was the outcome variable, divided into primary and secondary outcomes. Seven prospective cohort studies met the inclusion criteria. Eight HRQoL measures and six surgical methods were identified. Long-term follow-up time ranged from 5-10 years, sample sizes from 44 to 655 patients, and follow-up rates from 61 to 92%. None of the seven studies were randomized controlled trials and only two studies used control groups. Six out of seven studies showed statistically significant improvements in all of the primary outcomes, while one study showed statistically significant improvements in one of two primary outcomes. Of the statistically significant HRQoL improvements, 92 percent were clinically meaningful. Peak improvements in primary HRQoL outcomes were typically observed during the first years of follow-up, followed by a gradual decline that seemed to stabilize five years postoperatively. Long-term HRQoL scores typically remained improved relative to preoperative scores but were somewhat below population norm scores. In conclusion, while bariatric surgery candidates reported impaired HRQoL pre-surgically, their HRQoL improved considerably after bariatric surgery and much of the initial HRQoL improvements were maintained over the long term.

**Keywords:** Health-related quality of life; quality of life; bariatric surgery; obesity; systematic review

## Background

Severe obesity, defined by a body mass index (BMI)  $\geq 40$  or  $\geq 35$  with obesity-related disease<sup>[1]</sup>, is associated with a multitude of comorbidities, shorter life span and impaired health-related quality of life (HRQoL)<sup>[1-3]</sup>. HRQoL is a multidimensional construct of the individual's perception of the negative impact of an illness, capturing the physical, psychological, and social dimensions of health<sup>[4, 5]</sup>. Among patients with severe obesity, bariatric surgery candidates are at a particularly increased risk of having reduced HRQoL<sup>[6]</sup>, and the desire for an improved HRQoL is often a major motivation for seeking bariatric surgery<sup>[7, 8]</sup>. Although many studies have reported significant improvements in HRQoL after bariatric surgery<sup>[9]</sup> there are few long-term prospective studies (e.g.,  $\geq 5$  years). Further, no systematic reviews of long-term HRQoL outcomes after bariatric surgery have been conducted. A recent systematic review of HRQoL outcomes in 53 randomized trials for weight loss interventions included only four bariatric surgery studies, with follow-ups ranging between 52 and 104 weeks<sup>[10]</sup>. Thus, we aimed to conduct a systematic review of prospective studies reporting HRQoL data at least five years after bariatric surgery.

## Methods

This review was conducted in accordance with the MOOSE Guidelines for Meta-Analyses and Systematic Reviews of Observational Studies<sup>[11]</sup>.

### *Search strategy*

We used OVID to search in Medline, Embase and PsycINFO for prospective studies published in the year 2000 or later. Searches were limited to full text articles written in English. The search terms were as follows: (bariatric surgery OR weight loss surgery OR obesity surgery OR weight reduction surgery OR biliopancreatic diversion OR duodenal switch OR laparoscopic band OR lap band OR gastric band OR gastric bypass OR gastroplasty OR gastric sleeve OR sleeve gastrectomy) AND

(obesity) AND (quality of life OR health-related quality of life OR health status). We also manually searched for additional articles. The last search was conducted on May 22, 2014.

#### *Selection criteria*

Studies were eligible and considered to be of acceptable quality if (1) study participants were 18 years or older, (2) study participants had severe obesity, defined by a body mass index (BMI)  $\geq 40$  or  $\geq 35$  with comorbidity, (3) HRQoL was measured with a validated generic and/or obesity-specific measure, or validated measures of physical or mental HRQoL (e.g. Hospital Anxiety and Depression Scale), (4) HRQoL was measured before and at least five years after the surgery (according to the American Society for Metabolic and Bariatric Surgery, a  $\geq$  five-year follow-up is defined as a long-term study<sup>[12]</sup>, (5) attrition rate was  $<50\%$ , and (6) the study had the power to detect  $\geq 0.5$  standard deviation change from baseline using a two-sided paired test providing 90% power,  $P < 0.05$ , indicating at least 44 paired observations [13].

#### *Data collection and analyses*

Two of the authors independently reviewed the literature to identify studies meeting our criteria. Disagreement was resolved by consensus. We successfully contacted the first authors of two included papers in order to obtain more information on HRQoL standard deviations<sup>[14]</sup>, and raw HRQoL scores<sup>[15]</sup>. Change in HRQoL  $\geq 5$  years postoperatively compared to baseline was the main outcome. We divided the HRQoL outcomes into primary and secondary outcomes to simplify reporting of the results and to minimize the problems associated with multiple comparisons<sup>[5]</sup>. Primary outcomes were defined as summary scores of generic or obesity-specific measures or a measure of overall health status or overall well-being. Secondary outcomes were defined as domain scores, such as social interaction, physical appearance, and self-regard. When available, comparisons with HRQoL scores from the general population (norm scores) were also studied for comparison. The

effect sizes of the HRQoL changes over time were obtained by dividing the change scores by the standard deviations (SD) of the baseline scores. Effect sizes are reported in SD units of change. Changes over time were considered to be trivial (<0.2 SD units), small (0.2 to <0.5 SD units), moderate (0.5 to <0.8 SD units) or large ( $\geq 0.8$  SD units) according to guidelines proposed by Cohen<sup>[16]</sup>. Changes over time were also considered to be clinically meaningful if the effect size was at least 0.5 SD units<sup>[13]</sup>. We also calculated the size of the difference between HRQoL scores of the patients compared to population norms by dividing the difference scores by the SDs of the patient group. Interpretation of these comparisons also followed Cohen's guidelines. A meta-analysis was considered inappropriate due to heterogeneity.

## Results

### *Overview of HRQoL Measures*

As can be seen in Table 1, eight HRQoL measures were used in the reviewed studies, resulting in nine primary HRQoL outcomes (SF-36 Physical Component Summary (PCS) and Mental Component Summary (MCS) are counted as separate primary outcomes). Of these eight HRQoL measures, five were generic (General Health Rating Index, General Well-being, 15 D, Nottingham Health Profile II, and SF-36) and three were obesity-specific (Obesity-Related Problems scale, Weight Distress, and Impact of Weight on Quality of Life-Lite). Despite heterogeneity of HRQoL instruments, three studies<sup>[15, 17, 18]</sup> reported SF-36 summary score changes, allowing us to illustrate the trajectory of change following bariatric surgery.

### *Overview of general findings*

The database research identified 1589 articles; an additional article was identified manually<sup>[18]</sup>. After some consideration, we also excluded two studies that used the gastrointestinal quality of life index<sup>[19, 20]</sup>, as this measure is neither an obesity-related nor a generic HRQoL instrument. The final sample

consisted of seven prospective cohort studies (N=1113). Details of these seven studies are presented in Table 2. All but one <sup>[18]</sup> of 1590 papers were identified through the database search. After removing ineligible papers, nine studies remained. None of the seven studies were randomized controlled trials (RCTs) and only two studies used control groups. Six surgical methods were identified. The total follow-up times ranged from 5-10 years, the sample sizes ranged from 44 to 655 patients, and the follow-up rates ranged from 60.8 to 92%.

In Table 3 changes in primary HRQoL measures are reported for each study, including effect sizes, statistical significance, and norm scores for interpretation. Six out of seven studies showed statistically significant improvements in all of the primary outcomes, while one study <sup>[18]</sup> showed statistically significant improvements in one of two primary outcomes (Table 3). Of the statistically significant HRQoL improvements, 92 percent were clinically meaningful (i.e. effect size  $\geq 0.5$  SD units). For the three studies that included population norms <sup>[15, 17, 21]</sup>, long term HRQoL scores were lower than population norms.

Although there was considerable heterogeneity of outcome measures, three of the studies reported changes in the PCS and MCS scores <sup>[15, 17, 18]</sup>, allowing us to examine their long-term HRQoL trajectories (noting that the studies did not report yearly results, so we are generalizing based on the available data). In Figure 1, SF-36 PCS and MCS scores are shown over a 6-year period. Results for the PCS are quite consistent across studies with the peak occurring in the first 1-2 years, followed by scores that decline somewhat at 5-6 years but are much higher than baseline scores. Results for SF-36 MCS appears less consistent across studies, with seemingly higher peak and nadir scores occurring in the Aasprang et al study <sup>[17]</sup>. Nevertheless, all three studies showed declines in SF-36 MCS at five to six years. Only one study <sup>[21]</sup> assessed HRQoL annually, describing peak improvements in HRQoL at six months and one year, with declines occurring until about five years and stabilization between six and ten years. However, this study did not administer the SF-36, so we cannot make a direct comparison with the three studies used to create the above trajectory.

*Overview of specific studies*

We will now describe the results of the included studies separately by study (Tables 2 and 3).

*Karlsson et al. 2007* <sup>[22]</sup>: The mean BMI in the surgical group changed from  $41.9 \pm 4.2$  at baseline to  $35.3 \pm 5.4$  at ten years, compared to no change in the conventionally treated control group. The primary outcomes of Current Health Perception and the OP scale improved significantly for the bariatric surgery patients, and these improvements were greater than those found in the control group ( $P < 0.001$ ). Only the change in the OP scale was clinically meaningful at ten years. Regarding secondary outcomes the surgically-treated patients showed significantly greater improvements compared to the control group for Social Interaction <sup>[23]</sup> and the HADS Depression score <sup>[24]</sup> ( $P < 0.05$ ). No significant changes between groups were observed for Overall Mood <sup>[25]</sup> and the HADS Anxiety score ( $P > 0.05$ ). The bariatric surgery patients' primary HRQoL scores were also compared to population norms from individuals with a BMI  $< 30$ , adjusted for age and gender, to reflect the same distribution as the study sample. The surgical patients' Current Health Perceptions score was 1.14 SD units lower than the population norm before surgery (indicating greater impairment) and 0.81 SD units lower after ten years. The surgical patients' OP scale score was 1.8 SD units higher than the population norm before surgery (indicating greater impairment), as well as 0.8 SD units higher after ten years.

*Mathus-Vliegen et al. 2007* <sup>[26]</sup>: The mean BMI changed from  $50.7 \pm 8.5$  at baseline to  $39.3 \pm 10.3$  five years after surgery. The primary outcomes of General Well-being and Weight Distress improved significantly, and the changes on both were clinically meaningful. Regarding secondary outcomes <sup>[27]</sup> statistically significant improvements were found for Depression (P value of overall change over time = 0.005), Physical Appearance ( $P < 0.001$  for overall change over time) and Self-regard ( $P = 0.001$  for overall change over time). No data were available on population norms for comparison.

*Helmiö et al. 2011*<sup>[28]</sup>: The mean excess weight loss was 57.9% ± 31.1 at five years' post-surgery. Mean BMI before surgery was 46.3 ± 6.3. The primary outcome of 15D total score showed significant improvement, and this change was clinically meaningful. The secondary outcomes (i.e. 15D domains) of Mobility, Breathing, Usual Activities, Depression, Vitality and Sexual Activity improved significantly from baseline to the five-year follow-up ( $P < 0.05$ ), while no significant changes were reported on the domains of Sleeping, Eating, Elimination, Mental functioning, Discomfort and distress ( $P > 0.05$ ). HRQoL was also assessed with an early version of the Moorehead-Ardelt Quality of Life Questionnaire that instructed patients to rate themselves retrospectively<sup>[29]</sup>. Since this HRQoL questionnaire was not administered prospectively, we did not include these results. No data were available on population norms for comparison.

*Schouten et al. 2011*<sup>[14]</sup>: The mean BMI changed from 47 at baseline to 35 seven years post-surgery. The primary outcome of NHPII improved significantly and the change was clinically meaningful. After seven years 19 patients in the LAGB group had undergone a conversion to Roux-en-Y Gastric Bypass. The NHPII scores of converted versus not converted patients was virtually identical after seven-years (10.0 versus 10.6). In terms of secondary outcomes the NHPII domains of Energy level and Physical Abilities improved ( $P < 0.05$ ), but no changes were observed in Pain, Emotional reaction, Sleep and Social Isolation from baseline to the seven-year follow-up ( $P > 0.05$ ). Significant improvements were also observed in the Sickness Impact Profile (SIP-68) [30] domains of Mobility Control, Social Behavior and Mobility Range ( $P < 0.05$ ), while no changes were observed in Somatic Autonomy, Psychic Autonomy and Communication and Emotional Stability ( $P > 0.05$ ).

*Kolotkin et al. 2012*<sup>[15]</sup>: The mean unadjusted BMI in the surgical group changed from 47.0 ± 7.6 at baseline to 33.6 ± 7.2 at six years, compared to 0.2% regain at six years in control group 1 (individuals evaluated for Roux-en-Y Gastric Bypass surgery but who did not have the surgery) and no change in control group 2 (obese individuals without a history of bariatric surgery randomly chosen from a population database). The primary outcomes of PCS, MCS, and IWQOL-Lite total

improved significantly, and all changes were clinically meaningful. Regarding secondary HRQoL outcomes, the surgically treated patients had significantly greater changes from baseline to six years compared to the control groups in all five IWQOL-Lite domains and in seven of the eight SF-36 subscales ( $P_s < 0.001$ ). No statistically significant differences in change between groups were observed for the SF-36 domain of Emotional Role Functioning ( $P = 0.320$ ). Changes in the surgical group were greater than in controls for the SF-36 PCS and IWQOL-Lite total ( $P_s < 0.001$ ), while no difference was found between groups for the SF-36 MCS ( $P = 0.880$ ). The patients' HRQoL scores were also compared to norm scores, representing the average US population. The PCS score was 2.0 SD units lower than the population norm before surgery (indicating greater impairment) and 0.4 SD units lower after six years. The MCS score was 0.7 SD units lower than the population norm before surgery (indicating greater impairment), and 0.2 SD units lower after six years. The IWQOL-Lite total score was 3.8 SD units lower (indicating greater impairment) than the population norm before surgery, and 0.9 SD units lower after six years.

*Aasprang et al. 2013*<sup>[17]</sup>: The mean BMI changed from  $51.7 \pm 7.5$  at baseline to  $32.9 \pm 6.6$  five years post-surgery. The primary HRQoL outcomes of PCS and MCS improved significantly, and these changes were clinically meaningful. For secondary HRQoL outcomes the eight SF-36 scores and the HADS-depression score also improved significantly ( $P_s < 0.05$ ), while the HADS-anxiety score did not ( $P = 0.126$ ). The patients' primary HRQoL scores were also compared to norm scores, adjusted for age and gender to reflect the same distribution as the study sample. The PCS score was 1.95 SD units lower than the population norm before surgery (indicating greater impairment) and 0.38 SD units lower after five years. The MCS score was 1.2 SD units lower than the population norm before surgery (indicating greater impairment), and 0.5 SD units lower after five years.

*Zijlstra et al. 2013*<sup>[18]</sup>: The mean BMI changed from  $46.5 \pm 4.9$  at baseline to  $36.8 \pm 6.2$  six years after surgery. The PCS score improved significantly after six years, but changes in the MCS were not statistically significant. In addition, only the change in PCS was clinically meaningful. No

secondary outcomes were reported in this study, nor were data reported on population norms for the PCS and MCS in the Netherlands.

## Discussion

The findings of this systematic review were quite consistent across studies despite heterogeneous HRQoL measures, surgery types, baseline BMIs, and countries. The principal finding was that six out of seven studies showed statistically significant improvements in all of the primary outcomes, while one study showed statistically significant improvements in one of two primary outcomes. Of the statistically significant HRQoL improvements, 92 percent were clinically meaningful. The improvements were especially large in the obesity-specific HRQoL measures. However, some declines in HRQoL were observed over time, despite initial improvements. Nevertheless, long-term HRQoL scores typically remained improved relative to preoperative scores, even though they tended to be somewhat below population norm scores. These findings are consistent with long-term cross sectional studies (13-25 years of follow up) showing that bariatric surgery is associated with good HRQoL provided there has been adequate and maintained weight loss without bothersome side-effects<sup>[31, 32]</sup>.

The first two years after surgery have often been named “the honey-moon period”. The patient at this stage usually, and often for the first time in life, experiences a meaningful amount of weight loss, as well as a feeling of being in control of the obesity. The highest HRQoL scores seem to coincide with the nadir weight. After this period there is a decline in scores, probably associated with weight regain<sup>[21]</sup>. Whether more support from health personnel at this stage will reduce weight gain or reduce the decline in HRQoL scores is unknown and could be explored in future studies. In treating patients over the long term, health providers need to be especially mindful to provide sincere and positive support during this period, as patients may experience weight regain, and perhaps declines in HRQoL, as a personal failure due to a history of obesity-related stigma<sup>[33]</sup>.

This review has several limitations that should be addressed. First, as no long-term RCTs were identified, we had to rely on information from prospective cohort studies<sup>[34]</sup>. Second, only seven published studies were included in this review, and the total sample size was relatively small (N=1113). Third, the attrition rates in the included studies ranged from 8 to 39.2%. Thus, the outcomes of a large number of patients are unknown. Some of the patients who dropped out of the follow-up programs might have had small HRQoL changes, while others may have felt that they were satisfied with the treatment and had no need to attend follow-ups<sup>[35]</sup>. Fourth, the included studies were heterogeneous with respect to baseline BMI, HRQoL instruments, and surgical methods. Furthermore, one of the studies does not fully represent modern bariatric surgery due to developments in surgery procedures over the last twenty years<sup>[21]</sup>. As a result of this heterogeneity, a meta-analysis was considered inappropriate. Our decision not to conduct a meta-analysis is in accordance with a recent systematic review that concluded it is extremely difficult to synthesize results of bariatric surgery studies due to poor study design and reporting of patient reported outcomes such as HRQoL<sup>[36]</sup>. Fifth, it should be noted that the standard deviations of the HRQoL scores may differ across studies, limiting exact direct comparisons. Thus, the SD units in this review should be considered as estimates of the clinical meaningfulness of the results<sup>[13]</sup>. Finally, testing for funnel plot asymmetry was not performed as only seven studies were included in this review<sup>[37]</sup>. We acknowledge that reporting bias may lead to over-representation of significant or positive studies<sup>[38]</sup>.

## Conclusion

This systematic review of seven studies of long-term HRQoL changes in bariatric surgery candidates indicates remarkably similar findings across studies, outcome measures, and countries. All studies showed impaired HRQoL prior to surgery, followed by initial post-surgery improvements (i.e. 1-2 years post-surgery), with much of the initial HRQoL improvements being maintained over the long

term, despite some decline. The consistency of the findings is encouraging. However, in order to strengthen the body of research in this field, further prospective long-term studies are warranted , especially RCT's or large, well designed, observational studies with high retention rates, comparator groups and carefully chosen generic and obesity-specific HRQoL measures <sup>[36, 39]</sup> .

**Disclosures**

XXXXXX has received royalties from Duke University for the IWQOL-Lite.

Accepted manuscript

## Legends

Figure 1. Changes in the physical component summary (PCS) of the SF-36 are presented in 1. A. Higher scores represent better HRQL. Changes in the mental component summary (MCS) are presented in 1. B. Kolotkin et al. [15] postoperative data were assessed after 2 and 6 years. Aasprang et al. [17] postoperative data were assessed after 1 and 5 years. Zijlstra et al. [18] postoperative data were assessed after 1 and 6 years. The PCS and MCS scores in the study by Aasprang et al. were converted to orthogonal scores [40] so that the scores were comparable to those in the other two studies. A score = 50 is the average score in the general population in the United States [40].

Table 1. Description of the primary health-related quality of life measures

Table 2. Long-term studies on health-related quality of life after bariatric surgery (N=1113)

Note: \*Control groups were not randomized.

Table 3. Long-term changes in health-related quality of life after bariatric surgery (N=1113)

Note: Abbreviations: NHP II: The Nottingham Health Profile II, PCS: physical component summary, MCS: mental component summary, IWQOL-Lite total: Impact of Weight on Quality of Life-Lite total, NA = not available.

\* These are unadjusted effect sizes obtained from Kolotkin. The Kolotkin et al [15] paper reports adjusted effect sizes.

**References**

1. WHO. Obesity: preventing and managing the global epidemic. Geneva, WHO, 2000.
2. Kolotkin RL, Meter K, Williams GR. Quality of life and obesity. *Obes Rev* 2001;2:219-29.
3. Whitlock G, Lewington S, Sherliker P, et al. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. *Lancet* 2009;373:1083-96.
4. Cramer JA, Spilker B. Quality of life and pharmacoeconomics an introduction. Philadelphia, Lippincott-Raven, 1998.
5. Fayers PM, Machin D. Quality of life: the assessment, analysis and interpretation of patient-reported outcomes. Chichester, John Wiley, 2007.
6. Kolotkin RL, Crosby RD, Williams GR. Health-related quality of life varies among obese subgroups. *Obes Res* 2002;10:748-56.
7. Meana M, Ricciardi L. Obesity surgery. Stories of altered lives. Reno, University of Nevada Press, 2008.
8. Munoz DJ, Lal M, Chen EY, et al. Why patients seek bariatric surgery: a qualitative and quantitative analysis of patient motivation. *Obes Surg* 2007;17:1487-91.
9. Sarwer DB, Wadden TA, Fabricatore AN. Psychosocial and behavioral aspects of bariatric surgery. *Obes Res* 2005;13:639-48.
10. Warkentin LM, Das D, Majumdar SR, Johnson JA, Padwal RS. The effect of weight loss on health-related quality of life: systematic review and meta-analysis of randomized trials. *Obes Rev* 2014;15:169-82.
11. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000;283:2008-12.
12. ASMBS. Updated position statement on sleeve gastrectomy as a bariatric procedure. *Surg Obes Relat Dis* 2012;8:e21-6.
13. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003;41:582-92.
14. Schouten R, Wiryasaputra DC, van Dielen FM, van Gemert WG, Greve JW. Influence of reoperations on long-term quality of life after restrictive procedures: a prospective study. *Obes Surg* 2011;21:871-9.
15. Kolotkin RL, Davidson LE, Crosby RD, Hunt SC, Adams TD. Six-year changes in health-related quality of life in gastric bypass patients versus obese comparison groups. *Surg Obes Relat Dis* 2012.
16. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 1 ed. New York, Academic Press, 1988.

17. Aasprang A, Andersen JR, Vage V, Kolotkin RL, Natvig GK. Five-year Changes in Health-Related Quality of Life after Biliopancreatic Diversion with Duodenal Switch. *Obes Surg* 2013.
18. Zijlstra H, Larsen J, Wouters ER, BV, Grenen R. The long-term course of quality of life and the prediction of weight outcome after laparoscopic adjustable gastric banding: a prospective study. *Bariatric surgical patient care* 2013;8:18-22.
19. Wang W, Yu PJ, Lee YC, Wei PL, Lee WJ. Laparoscopic vertical banded gastroplasty: 5-year results. *Obes Surg* 2005;15:1299-303.
20. Lee WJ, Ser KH, Lee YC, Tsou JJ, Chen SC, Chen JC. Laparoscopic Roux-en-Y vs. mini-gastric bypass for the treatment of morbid obesity: a 10-year experience. *Obes Surg* 2012;22:1827-34.
21. Karlsson J, Taft C, Ryden A, Sjöström L, Sullivan M. Ten-year trends in health-related quality of life after surgical and conventional treatment for severe obesity: the SOS intervention study. *Int J Obes* 2007;31:1248-61.
22. Sjöström L, Narbro K, Sjöström CD, et al. Effects of bariatric surgery on mortality in Swedish obese subjects. *N Engl J Med* 2007;357:741-52.
23. Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: development and final revision of a health status measure. *Med Care* 1981;19:787-805.
24. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res* 2002;52:69-77.
25. Sjöberg L, Svensson E, Persson LO. The measurement of mood. *Scand J Psychol* 1979;20:1-18.
26. Mathus-Vliegen EM, de Wit LT. Health-related quality of life after gastric banding. *Br J Surg* 2007;94:457-65.
27. Mathias SD, Williamson CL, Colwell HH, et al. Assessing health-related quality-of-life and health state preference in persons with obesity: a validation study. *Qual Life Res* 1997;6:311-22.
28. Helmiö M, Salminen P, Sintonen H, Ovaska J, Victorzon M. A 5-year prospective quality of life analysis following laparoscopic adjustable gastric banding for morbid obesity. *Obes Surg* 2011;21:1585-91.
29. Moorehead MK, Ardelt-Gattinger E, Lechner H, Oria HE. The validation of the Moorehead-Ardelt Quality of Life Questionnaire II. *Obes Surg* 2003;13:684-92.
30. de Bruin AF, Buys M, de Witte LP, Diederiks JP. The sickness impact profile: SIP68, a short generic version. First evaluation of the reliability and reproducibility. *J Clin Epidemiol* 1994;47:863-71.
31. Våge V, Solhaug JH, Viste A, Bergsholm P, Wahl AK. Anxiety, depression and health-related quality of life after jejunioileal bypass: a 25-year follow-up study of 20 female patients. *Obes Surg* 2003;13:706-13.
32. de Zwaan M, Lancaster KL, Mitchell JE, et al. Health-related quality of life in morbidly obese patients: effect of gastric bypass surgery. *Obes Surg* 2002;12:773-80.

33. Puhl RM, Heuer CA. The stigma of obesity: a review and update. *Obesity* (Silver Spring) 2009;17:941-64.
34. Sugerman HJ, Kral JG. Evidence-based medicine reports on obesity surgery: a critique. *Int J Obes* 2005;29:735-45.
35. Moroshko I, Brennan L, O'Brien P. Predictors of attrition in bariatric aftercare: a systematic review of the literature. *Obes Surg* 2012;22:1640-7.
36. Coulman KD, Abdelrahman T, Owen-Smith A, Andrews RC, Welbourn R, Blazeby JM. Patient-reported outcomes in bariatric surgery: a systematic review of standards of reporting. *Obes Rev* 2013.
37. Higgins J, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. 2011; [http://handbook.cochrane.org/chapter\\_10/10\\_4\\_3\\_1\\_recommendations\\_on\\_testing\\_for\\_funnel\\_plot\\_asymmetry.htm](http://handbook.cochrane.org/chapter_10/10_4_3_1_recommendations_on_testing_for_funnel_plot_asymmetry.htm). Accessed 7. May, 2014.
38. Song F, Parekh S, Hooper L, et al. Dissemination and publication of research findings: an updated review of related biases. *Health Technol Assess* 2010;14: 1-193.
39. Wolfe BM, Belle SH. Long-term Risks and Benefits of Bariatric Surgery: A Research Challenge. *JAMA* 2014.
40. Ware JE, Kosinski M. SF-36 physical and mental health summary scales: a manual for users of version 1. 1 ed. Lincoln, RI, QualityMetric Inc., 2001.
41. Davies A, Sherbourne C, Peterson J, Ware J. *Scoring Manual: Adult Health Status and Patient Satisfaction Measures Used in RAND's Health Insurance Experiment*. Santa Monica, The RAND Corporation, 1998.
42. 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 Relat Metab Disord* 2003;27:617-30.
43. Sintonen H. The 15D instrument of health-related quality of life: properties and applications. *Ann Med* 2001;33:328-36.
44. Schowalter M, Benecke A, Lager C, et al. Changes in depression following gastric banding: a 5- to 7-year prospective study. *Obes Surg* 2008;18:314-20.
45. Hunt SM, McEwen J, McKenna SP. Measuring health status: a new tool for clinicians and epidemiologists. *J R Coll Gen Pract* 1985;35:185-8.
46. Ware JE, Kosinski M, Gandek B. SF-36 health survey : manual & interpretation guide. 2 ed. Lincoln, RI, QualityMetric Inc, 2000.
47. Kolotkin RL, Crosby RD, Kosloski KD, Williams GR. Development of a brief measure to assess quality of life in obesity. *Obes Res* 2001;9:102-11.
48. Crosby RD, Kolotkin RL, Williams GR. An integrated method to determine meaningful changes in health-related quality of life. *J Clin Epidemiol* 2004;57:1153-60.

Table 1.

Studies	Health-related quality of life Measures
Karlsson et al. [21]	<b>Current health perception</b> was measured by the Current Health Scale selected from the General Health Rating Index [41]. This scale includes nine general statements on perceived health. Item responses are aggregated to a total score ranging between 0 and 100, where a higher score indicates more positive perceived health status.
Karlsson et al. [21]	<b>The obesity-related problems scale (OP)</b> is a condition specific instrument constructed to measure the impact of obesity on psychosocial functioning [42]. OP comprises eight items on a four-point scale. Subjects are asked to indicate how bothered they are by their obesity in a broad range of social activities. Item responses are aggregated to a total score ranging between 0 and 100, where a lower score indicates fewer obesity-related problems.
Mathus-Vliegen et al. [26]	<b>The General Well-being</b> score measures the general perception of health, current health status, and perception (harm, benefit and control) of the current weight condition [27]. Item responses are aggregated to a total score ranging between 2 and 42, where a lower score indicates a more positive general well-being.
Mathus-Vliegen et al. [26]	<b>Weight Distress</b> was measured by a modified version of the distress scale from the Medical Outcomes Study, with the wording modified to be weight specific [27]. Item responses were aggregated to a total score ranging between 3 and 36, where a lower score indicates less weight distress.
Hemiö et al. [28]	<b>The 15D</b> is a generic, 15-dimensional measure of HRQoL (physical, social and mental domains) that can be used as a single index score measure [43]. Item responses were aggregated to a total score ranging between 0-1, where a higher total score indicates a more positive 15D status.
Schouten et al. [44]	<b>The Nottingham Health Profile II (NHP II)</b> measures those areas of task performance most affected by health [45]. It contains seven statements that refer to the effects of health problems on occupation, ability to perform domestic tasks, hobbies, personal relationships, sex life, social life, and holidays. Item responses are aggregated to a total score ranging between 7 and 21, where a lower score indicates a more positive NHP II status.
Kolotkin et al.[15] , Aasprang et al. [17]and Zijstra et al. [18]	<b>The SF-36 physical component summary (PCS) and SF-36 mental component summary (MCS)</b> are derived from the Medical Outcomes Study Short-Form 36 (SF-36) that assesses eight dimensions of physical, social and mental health [46]. The PCS is a physical summary score and the MCS a mental summary score of the items in SF-36. PCS and MCS scores were standardized using a representative sample of the 1998 US general population, so that the mean score equals 50 and the standard deviation equals 10 [40]. Higher scores indicate more positive PCS and MCS status.

Table 1. Cont.

Studies	Health-related quality of life Measures
Kolotkin et al.[15]	<b>Impact of Weight on Quality of Life (IWQOL-Lite) total score</b> is derived from a 31-item measure of weight- related quality of life [47]. There are five domain scores that are aggregated to a total score ranging between 0 and 100, where a higher score indicates a more positive IWQOL-Lite status. Norms for the IWQOL-Lite total were based on a sample (n = 534) of normal weight and overweight individuals not currently enrolled in any weight loss treatment program [48].

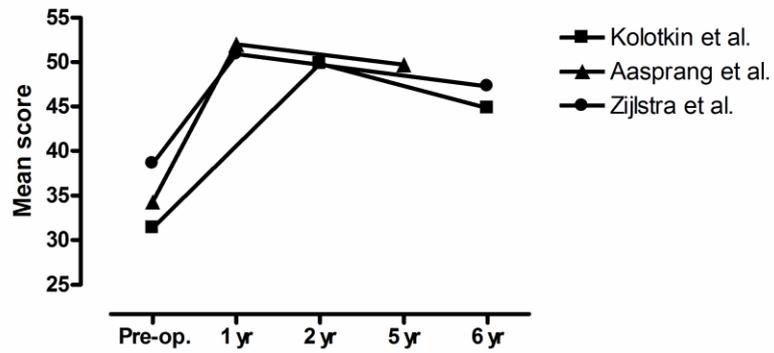
Table 2.

Study	Country	Follow-up time	Follow-up rate	Control group*	Patient characteristics.	Type of surgery
Karlsson et al. [21]	Sweden	10 years	N = 655 (77%)	Yes	Mean age at the inclusion was 47.0 ± 5.7 years in the surgical group and 48.4 ± 6.7 years in the conventional treatment group.	Fixed or variable banding, vertical banded gastroplasty or gastric bypass. Females (53.8%).
Mathus-Vliegen et al. [26]	Netherlands	5 years	N=44 (88%)	No	Females (68%). Mean age at the inclusion was 35.0 ± 7.4 years.	Laparoscopic adjustable gastric banding.
Helmiö et al. [28]	Finland	5 years	N = 49 (65.3%)	No	Females (68%). Mean age at the inclusion was 42.4 ± 10.7 years.	Laparoscopic adjustable gastric banding.
Schouten et al. [14]	Netherlands	7 years	N = 44 (84%)	No	Females (77.3%). Mean age at the inclusion was 39.0 ± 9.0 years.	Laparoscopic adjustable gastric banding and Vertical banded gastroplasty.
Kolotkin et al. [15]	USA	6 years	N = 230 (71.2%)	Yes	Females (83.9%). Mean age at the inclusion was 43.4 ± 10.7 years in the surgical group and 44.7 ± 10.9/49.7 ± 10.5 years in the control groups.	Roux-en-Y Gastric bypass.
Aasprang et al. [17]	Norway	5 years	N = 46 (92%)	No	Females (54%). Mean age at the inclusion was 37.8 ± 8.1 years.	Biliopancreatic Diversion with Duodenal switch.
Zijstra et al. [18]	Netherlands	6 years	N = 45 (60.8%)	No	Females (84.5%). Mean age at the inclusion was 48.0 ± 9.0 years.	Laparoscopic adjustable gastric banding.

Table 3.

Study and measures	Before	After 5+ years	Effect Sizes	P-value for change	Norm score
Karlsson et al. [21]					
Current health perception	51.8 ± 24.1	57.5 ± 26.8	0.2	<0.001	79.3 ± 21.1
Obesity related problem scale (OP)	58.0 ± 27.0	29.7 ± 27.3	1.1	<0.001	8.4 ± 14.5
Mathus-Vliegen et al. [26]					
General well-being	29.0 ± 3.9	25.9 ± 6.3	0.8	<0.001	NA
Weight distress	24.2 ± 6.3	12.6 ± 6.4	1.8	<0.001	NA
Helmiö et al. [28]					
15-D total score	0.859 ± 0.081	0.899 ± 0.087	0.5	<0.001	NA
Schouten et al.					
NHP II	13.0 ± 4.5	10.2 ± 4.5	0.6	<0.001	NA
Kolotkin et al. [15]					
SF-36 PCS	31.4 ± 9.3	44.8 ± 10.8	1.4*	<0.001	50.0 ± 10
SF-36 MCS	41.5 ± 11.7	47.4 ± 11.9	0.5*	<0.001	50.0 ± 10
IWQOL-Lite total	33.2 ± 16.4	79.4 ± 17.5	2.8*	<0.001	94.7 ± 7.6
Aasprang et al. [17]					
SF-36 PCS	32.6 ± 10.1	48.4 ± 10.4	1.6	<0.001	52.3 ± NA
SF-36 MCS	37.8 ± 12.8	45.5 ± 15.2	0.6	<0.001	52.5 ± NA
Zijstra et al. [18]					
SF-36 PCS	38.6 ± 12.0	47.3 ± 11.3	0.7	<0.001	NA
SF-36 MCS	46.0 ± 11.7	47.0 ± 14.3	0.1	0.090	NA

1. A: SF-36 PCS



1. B: SF-36 MCS

