

Bariatric Surgery and its Impact on Long-term (≥ 5 years) Health Related Quality of Life

by
©Shannon Driscoll
200722718

B.Sc. (Honors) Memorial University of Newfoundland 2012

A thesis submitted to the
School of Graduate Studies
in partial fulfillment of the
requirements for the degree of
Master of Science in Medicine
Clinical Epidemiology
Faculty of Medicine

MEMORIAL UNIVERSITY OF NEWFOUNDLAND

May 2017

Abstract

Bariatric surgery results in significant weight loss in the majority of patients living with severe obesity ($\text{BMI} \geq 35 \text{ kg/m}^2$). Improvement in Health Related Quality of Life (HRQoL) is an equally important patient reported outcome; however, there are few studies reporting the impact of bariatric surgery on long-term (≥ 5 years) HRQoL outcomes. The main aim of this thesis was two-fold: first to conduct a SR in order to assess the quality of evidence and effectiveness of bariatric surgery on $\text{HRQoL} \geq 5$ years in patients ≥ 18 years compared to non-surgical control groups and second to conduct a meta-analysis (MA) of studies that have been deemed appropriate.

PubMed, Cochrane Review, EmBase, CINANL, PsycInfo, obesity conference abstracts, and reference lists of published papers were searched. Keywords were bariatric surgery, obesity, and quality of life. Studies were included if (1) there was ≥ 5 years follow-up, (2) patients had class II or III obesity, (3) individuals completed a validated HRQoL survey, and (4) there was a non-surgical comparison group with obesity. Two reviewers independently assessed each study.

From the initial 1376 articles, 9 studies were included in the SR and 6 in the MA.

Inconsistent results for long-term improvements in physical and mental health emerged from the SR. However, in contrast, the MA found significant improvements in these domains ≥ 5 years after bariatric surgery. These study findings provide evidence for a substantial and significant improvement in physical and mental health favoring the surgical group compared with controls spanning 5 to 25 years after surgery, an important finding for patients, clinicians and decision-makers.

Acknowledgements

I would like to acknowledge and thank my supervisor Laurie Twells for her guidance and support throughout my Masters program; you have helped me along greatly both academically and professionally. I can't thank you enough! I would also like to express my gratitude to Kendra Lester, Data Analyst at the Translational Personalized Medicine Initiative of Memorial University for her assistance with data input and putting up with my endless questions. I would also like to thank my committee members Drs' Deborah Gregory and John Fardy for their input and feedback on my Masters work. Finally, I would like to thank my family and friends for their support during my graduate studies.

Table of Contents

Abstract	ii
Acknowledgements	iii
List of Tables	vi
List of Figures	vii
List of Abbreviations	viii
List of Publications	ix
List of Appendices	x
Co-authorship Statement	xi
Chapter 1 Introduction	1
1.1 Background and Rationale	1
1.1.1 <i>Epidemiology of Obesity</i>	1
1.1.2 <i>Health Related Quality of Life (HRQoL)</i>	1
1.1.3 <i>Severe Obesity and HRQoL</i>	4
1.1.4 <i>Bariatric Surgery as a Treatment for Severe Obesity</i> ...	6
1.1.5 <i>Bariatric Surgery and its Impact on HRQoL</i>	8
1.1.6 <i>Gaps in the Literature</i>	9
1.2 Purpose of the Study	10
1.3 Significance of the Study	10
1.4 Specific Research Objectives	10
1.5 References	12
Chapter 2 Protocol	17
2.1 Unpublished <i>A Priori</i> Protocol for Bariatric Surgery and its Impact on Quality of life in Adults: A Systematic Review	17
2.1.1 <i>Introduction</i>	17
2.1.2 <i>Study Measurement Tools</i>	19
2.1.3 <i>Reason for a Systematic Review</i>	21
2.1.4 <i>Study Selection and Data Extraction</i>	23
2.1.5 <i>Risk of Bias Assessment</i>	24
2.1.6 <i>Measurement of Treatment Effect</i>	24
2.1.7 <i>Strategy for Data Synthesis</i>	24
2.2 Complete Search Strategy	25
2.3 References	28
Chapter 3 Long-Term Health-Related Quality of Life in Bariatric Surgery Patients: A Systematic Review and Meta-Analysis	30
3.1 Introduction	31
3.2 Methods	32
3.2.1 <i>Search Strategy</i>	32
3.2.2 <i>Risk of Bias (Quality) Assessment</i>	34

3.2.3	<i>The US Preventative Services Task Force (USPSTF)</i>	
	<i>Quality Criteria Rating</i>	34
3.2.4	<i>Meta-Analysis</i>	36
3.2.5	<i>Data Analysis for Meta-Analysis</i>	37
3.3	Results.....	39
3.3.1	<i>Systematic Review</i>	39
3.3.2	<i>The Short-Form 36</i>	39
3.3.3	<i>Meta-Analysis</i>	43
3.3.4	<i>Synthesis of Results</i>	50
3.4	Discussion	54
3.5	References	56
Chapter 4	Summary	59
4.1	Overview.....	59
4.2	Strengths and Limitations	59
4.3	Implications of Current Study Findings.....	62
4.4	Future Research	62
4.5	Conclusions.....	63
4.6	References.....	64
Appendix A	Short-Form (SF) 36.....	65
Appendix B	USPSTF Quality Criteria Rating For RTC and Cohort Studies.....	69

List of Tables

Table 1: Included studies for review and meta-analysis.....	44
Table 2: Risk of bias assessment	49

List of Figures

Figure 1. PRISMA flow diagram of selection process and included studies.....	41
Figure 2. Forest plot for SF-36 physical domains.....	52
Figure 2.a. Forest plot for SF-36 physical function domain.....	52
Figure 2.b. Forest plot for SF-36 role physical domain.....	52
Figure 2.c. Forest plot for SF-36 bodily pain domain.....	52
Figure 2.d. Forest plot for SF-36 general health domain	52
Figure 3. Forest plots for SF-36 mental domains.....	53
Figure 3.a. Forest plots for SF-36 vitality domain.....	53
Figure 3.b. Forest plots for SF-36 social functioning domain.....	53
Figure 3.c. Forest plots for SF-36 role emotional domain.....	53
Figure 3.d. Forest plots for SF-36 mental health domain.....	53

List of Abbreviations

- BAROS-** Bariatric Analysis of Reporting Outcome System
- BaSCo-** Bariatric Surgery Cohort
- BMI** - Body mass index
- CI-** Confidence intervals
- CINAHL** - Cumulative index to nursing and allied health literature
- EQ5D-** EuroQoL 5 Dimension
- EQ-5D-3L-** EuroQoL 5 Dimension 3 Level
- HRQoL-** Health related quality of life
- IQR-** Inter Quartile Range
- IWQoL-Lite-** Impact of Weight on Quality of Life- Lite
- MA-** Meta-analysis
- MeSH-** Medical search headings
- MCS-** Mental Component Score
- MOS-** Medical Outcomes Study
- OPS-** Obesity Problems Scale
- OWLQOL-** Obesity and Weight Loss Quality of Life
- PCS-** Physical Component Score
- PRISMA-** Preferred Reporting Items for Systematic Reviews and Meta-Analyses
- SF-36-** Short Form 36
- SF-12** - Short Form 12
- SOS-** Swedish Obesity Study
- SR-** Systematic review
- TOS-** The Obesity Society

List of Publications

Published Manuscript

Driscoll S, Gregory DM, Fardy JM, Twells LK. Long-term Health Related Quality of life in bariatric surgery patients: A systematic review and meta-analysis. *Obesity* January 2016, doi:10.1002/oby.21322.

Published Abstract

Driscoll S, Gregory DM, Fardy JM, Twells LK. Long-term Health Related Quality of life in bariatric surgery patients: A systematic review and meta-analysis. *CJD* 2015; 39:S30. doi: 10.1016/j.jcjd.2015.01.123.

Conference Presentation

Driscoll S. A systematic review and meta-analysis of long-term health related quality of life in bariatric surgery patients. Canadian Obesity Network Summit, Toronto Ontario, April 2015. [Oral Presentation]

List of Appendices

Appendix A	The Short Form (SF) -36 questionnaire.....	65
Appendix B	USPSTF Quality Criteria Rating For RTC and Cohort Studies.....	69

Co-authorship Statement

Chapter 2

Dr. Laurie Twells, Dr. Deborah Gregory and Dr. John Fardy conceived the study. Shannon Driscoll wrote the systematic review protocol with input from other authors, designed the search strategy with the help of a librarian and conducted the initial search.

Chapter 3

Dr. Laurie Twells, Dr. Deborah Gregory and Dr. John Fardy conceived the study. Shannon Driscoll completed the systematic review. Both Shannon Driscoll and Dr. Twells assessed the quality of the articles for inclusion in the systematic review and meta-analysis. Shannon Driscoll completed the systematic review and the meta-analysis and drafted the manuscript. Drs' Twells and Gregory critically revised the manuscript. All authors read and approved the final manuscript.

Chapter 1 Introduction

1.1 Background and Rationale

1.1.1 Epidemiology of Obesity

The increasing prevalence of obesity is a global health problem. Obesity, most often defined as a body mass index or BMI $\geq 30 \text{ kg/m}^2$ ⁽¹⁾, is associated with several comorbid conditions. Obesity can be divided into classes that indicate the associated health risks that are linked to excess body weight. Class I is defined as BMI=30-34.9 kg/m², Class II is BMI=35-39.9 kg/m² and Class III is BMI $\geq 40 \text{ kg/m}^2$ ^(1,2). Class II and III are associated with increased morbidity, higher risk of death and a greater burden on the healthcare system⁽³⁻⁵⁾. Class II and III obesity, commonly referred to as severe obesity, are associated with an even greater risk of developing comorbid conditions that include but are not limited to: insulin resistance and type 2 diabetes mellitus, cardiovascular disease, stroke, hypertension, dyslipidemia, sleep apnea, gallbladder disease, gout, osteoarthritis and a variety of cancers⁽¹⁾. In addition, severe obesity impairs health related quality of life (HRQoL) and shortens life expectancy^(1,5,6). The combination of comorbid conditions and their associated burdens have the greatest impact on a person living with severe obesity. In Canada, as in many other countries, severe obesity (BMI $\geq 35 \text{ kg/m}^2$) has increased. It now affects 5% of the population or an estimated 1.2 million adults and it is projected to increase to 6.4% by 2019⁽⁷⁾.

1.1.2 Health Related Quality of Life

HRQoL encompasses measures of well-being and physical and psychosocial functioning. It measures a patient's self reported state of physical, social and psychosocial well-being⁽⁸⁾. As with many chronic conditions or illnesses, the change in a patient's condition can be measured by various evaluations or end points such as changes in disease severity and improvement or

resolution of associated comorbidities or mortality, but these outcome measures are not reported by patients and do not always reflect treatment success that is meaningful to patients. This dissonance between patients and practitioners has led to the development of a quantifiable method for measuring patient HRQoL. Often administered in the form of a questionnaire or survey, HRQoL can be measured and quantified, which is useful for researchers and clinicians to better understand patients' perceptions of what matters most to them in terms of their quality of life. Being able to quantify HRQoL has helped researchers to evaluate changes in HRQoL before and after an intervention.

A number of tools have been developed and validated to measure HRQoL domains (e.g., mobility, pain, anxiety, usual activities). These tools can be specific to a condition such as measuring HRQoL associated with obesity, CVD, COPD and cancer or more generic by measuring HRQoL in general either in a population with a diseases or a broader general population. It is also important to measure the HRQoL of the general population in order to obtain population normative data; so that comparisons can be made with diseased-populations in order to better understand the impact of disease on HRQoL. The HRQoL of those living with obesity is often measured by both generic and specific (to obesity) validated tools. The most common generic tools include the EQ-5D and the SF-36 or one of its shortened versions such as the SF-12. Specific tools include the Impact of Weight on Quality of Life; BAROS (bariatric surgery specific) and Obesity Problems Scale (OPS).

The EQ-5D-3L⁽⁹⁾ is an indirect preference-based health survey that consists of 5 dimensions assessing mobility, self care, pain, usual activity and anxiety, each of which is rated on one of three levels; no problems, some problems or extreme problems. This combines to create 243 possible health states. The descriptive system is then scored using a set of weights that describe

the general population's preferences and from this a single summary preference based utility index or score is calculated. The utility scores range between 1 and -0.59, with 1 representing full health where the respondent has no problems for any dimension and -0.59 representing that the respondent has reported the lowest level for each dimension⁽⁹⁾. A visual analogue scale (VAS) is also calculated which shows an overall health score for the EQ-5L respondents. The VAS ranges from 0 to 100 representing worst imaginable health and perfect health respectively. The results are presented as a mean and standard deviation⁽⁹⁾.

The SF-12v2 (version 2) is a shortened version of the Short-Form 36 (SF-36) which is a common generic health status tool. The SF-12 has been validated against the SF-36 for patients with and without obesity⁽¹⁰⁾. The survey has 12 questions that examine 8 different quality of life domains: physical function, role physical, bodily pain, general health, vitality, social function, role emotional and mental health. Overall physical and mental component scores (MCS) are then calculated from these domains⁽¹¹⁾.

The IWQoL-Lite is the shortened form of the Impact of Weight on Quality of Life (IWQoL). This validated survey specifically assesses the impact of obesity on the quality of life in individuals seeking weight loss treatment and is the first survey of its kind to do so⁽¹²⁾. The IWQoL-Lite incorporates 31 statements that begin with the phrase "Because of my weight..." and each statement provides 5 response options from (1) "Never true" to (5) "Always true". These measure the impact of weight on 5 different domains: physical function, self-esteem, sexual life, public distress and work life. Each domain has an associated score calculated for each patient as long as they have answered at least 50% of the questions in that domain. An overall score is calculated if patients have responded to at least 26 of 31 questions. The raw scores are converted into a T-score (0-100) with 0 representing worst possible health and 100

representing best possible health. Results are reported as mean and standard deviation for each domain and total score⁽¹²⁾.

1.1.3 Severe Obesity and Health Related Quality of Life

Severe obesity has been linked to a significantly lowered HRQoL⁽¹³⁻¹⁷⁾. Many studies demonstrate a significant link between increasing BMI and a deterioration in HRQoL^(13, 15-17). In the systematic review (SR) and meta-analysis (MA) conducted by Van Nunen et al⁽¹³⁾, patients with obesity that were actively seeking surgical treatment showed the most reduced HRQoL when compared to both a general populations and other obese non-surgical treatment seeking populations using measures from the IWQoL (measured in mean deviation in standard deviation units from the norm, $-5.5 < d < -2.8$)⁽¹³⁾. The same SR and MA looked at studies using the SF-36 survey and found that while there were differences in results when compared to the IWQoL results, the group seeking surgical treatment still had the worst reported HRQoL (measured in mean deviation and in standard deviation units from the norm, $-1.6 < d < -0.5$)⁽¹³⁾. The SF-36 results showed a significant decrease in HRQoL in 5 of 8 physical domains and 2 mental domains. The authors suggest that these results indicate that impaired HRQoL and limitations in daily life in severely obese patients are not solely determined by excess weight⁽¹³⁾. This helps to highlight the validity of the SF-36 as a partly weight independent outcome measure for HRQoL. The results of these studies also suggest that a decrease in physical function may be a reason why patients seek out surgical treatment more than other obese patients that have similar BMIs.⁽¹³⁾

Herpertz et al.⁽¹⁴⁾ conducted a SR that focuses on whether bariatric surgery increases psychosocial functioning. In this review, 40 studies were reviewed with follow up times

ranging from 12-185 months. Various types of bariatric surgery and a variety of tools were used to measure psychosocial functioning.

When examining populations at baseline, studies have demonstrated that treatment seeking populations have higher levels of anxiety and depression as well as decreased self esteem before surgery versus after ^(14,18).

One study of patients examined a list of 187 HRQoL items that were then ranked by 100 patients who were classified as severely obese ($BMI \geq 40 \text{kg/m}^2$) ⁽¹⁵⁾. These items were categorized and given a clinical impact score which was determined by the important score reported by patients and the proportion of patients who deemed it important. The most important domains were identified as: activity and mobility, symptoms, personal hygiene/clothing, emotions, social interactions, sexual life, and eating behaviour. These represented the most severely impacted HRQoL items for the patients ⁽¹⁵⁾. This study showed that the impact of severe obesity is not limited to activity/mobility. Severe obesity contributes to the impairment of many different areas in HRQoL ⁽¹⁵⁾.

Two other studies ^(16,17) focus on the impact of severe obesity on HRQoL. Jia and Lubetkin ⁽¹⁶⁾ studied a range of different US populations and administered the SF-12 and EQ-5D surveys. Fontaine et al. ⁽¹⁷⁾ compared different classes of obese patients using the SF-36 survey. Using a multivariate linear regression, Jia and Lubetkin ⁽¹⁶⁾ found that when compared to a normal population, the severely obese population had the largest decrements in the SF-12 physical component scores (PCS) (4.00), EQ-5D index (0.073) and EQ VAS (4.68) scores. They also found a smaller but still significant decrement in SF-12 MCS (1.07) ⁽¹⁶⁾. The study also showed an inverse relationship with BMI and HRQoL scores. This effect is seen both with and without obesity-related comorbid conditions ⁽¹⁶⁾. Both studies found that physical domain scores were

more greatly impacted in the severely obese populations compared to normal and other overweight and obese populations^(16,17). Results for Fontaine et al.⁽¹⁷⁾ also demonstrate that the bodily pain score of the SF-36 is significantly more impaired ($P < 0.005$) and to a greater degree in severely obese patients (mean bodily pain score 52.8 ± 26.5) than those with the chronic conditions such as depression (58.8 ± 26.7), HIV (59.1 ± 23.2) or congestive heart failure (62.2 ± 30.9)⁽¹⁷⁾.

In summary, HRQoL is significantly impaired in severely obese populations. Studies demonstrate an inverse relationship between HRQoL and BMI and it appears that severe obesity impacts more than just the physical domains or weight-related domains of HRQoL such as mobility. A reduction in HRQoL in the severely obesity cannot be explained by excess weight or comorbid conditions alone.

1.1.4 Bariatric Surgery as a Treatment for Severe Obesity

Severe obesity is challenging to treat. “Standard” weight loss methods including lifestyle changes that focus on exercise and changes in diet or counseling or behaviour modifications, medical management and pharmaceutical solutions have limited effectiveness either on significant weight loss or sustained weight loss in the long term. These approaches to weight loss result in an average weight loss of 5-10% of initial body weight. This is modest weight loss and although it can often improve a patient’s comorbid profile, it is often not enough to result in significantly improved quality of life⁽¹⁹⁻²²⁾. The most effective treatment for severe obesity is bariatric surgery which results in significant and sustained weight loss, improvement in comorbid conditions, improved quality of life in the short term and a reduction in the risk of death⁽²³⁻³¹⁾. In Canada, the clinical practice guidelines recommend surgical treatment for adults

living with severe obesity. Patients must have a BMI \geq 35kg/m² including an obesity-related comorbidity (e.g., diabetes, hypertension, sleep apnea) or a BMI \geq 40kg/m²⁽³⁾.

There are different types of bariatric surgery, but all types are categorized as malabsorptive, restrictive or a combination of both^(3,28-31). Malabsorptive procedures limit caloric and nutrient absorption in the digestive tract. An example is the biliopancreatic diversion/duodenal switch (BPD/OS) in which part of the small intestine is removed or rerouted to avoid nutrient and calorie uptake. Restrictive procedures physically limit the amount of food that can be consumed. The gastric volume is reduced either through surgically removing a portion of the stomach (laparoscopic sleeve gastrectomy or LSG) or by banding with a device that can be inflated to restrict the volume of the stomach (adjustable gastric banding or AGB). The current gold standard for bariatric surgery is the roux-en-y gastric bypass (RYGB) which is a combination of both a restrictive and malabsorptive procedure as it reduces the size of the gastric pouch as well as re-routing a portion of the intestinal tract⁽²⁹⁾.

In the past 10-15 years there has been a shift in the type and volume of bariatric surgery performed^(32,33). In 2003, 146,301 bariatric surgeries were performed globally and in 2013 that number grew to 468,609, with the majority of surgeries performed in the United States and Canada⁽³³⁾. The type of surgery performed has also changed over time. In 2003, 85% were RYGB, 9% AGB and 4.5% BPD/DS. In 2013 the percentages were 45.0% RYGB, 37.0% LSG, 10% AGB and 1.5% BPD/DS⁽³³⁾. This trend shows exponential growth in LSG in the United States and Canada as well as globally⁽³³⁾. The increase in laparoscopic surgeries makes clinical sense as these procedures are minimally invasive and reduce surgical risk, hospital stay and recovery time compared to open techniques. These reasons have all contributed to the increasing popularity of this procedure⁽³⁴⁾.

Bariatric surgery is associated with a decreased rate of mortality when compared to non-surgical treatments for obesity^(24-26, 35-37). A MA comparing non-surgical interventions and bariatric surgery reported that bariatric surgery when compared to controls reduced the risk of global mortality (OR = 0.55, CI, 0.49–0.63), cardiovascular mortality (OR = 0.58, CI, 0.46–0.73), and all cause mortality (OR 0.70, CI, 0.59–0.84) through a reduction of myocardial infarction, diabetes, and cancer-related deaths⁽³⁵⁾. A reduction in mortality post-surgery has been reported in several studies in different countries^(24, 36, 37). The results from the Swedish Obesity Study (SOS)⁽²⁴⁾ show that after 15 years follow-up, the surgical group had a hazard ratio of 0.76 (95%CI, 0.59 to 0.99; P=0.04), compared with the control group that used conventional weight loss methods. Over the follow-up period, 129 subjects (6.3%) in the control group died, while 101 (5.0%) of the surgical group died. Adams et al.⁽³⁶⁾ found a reduction in all-cause mortality after 7.1 years of follow-up in a surgical group as compared to an obese population group (37.6 vs. 57.1 deaths per 10,000 person-years, P<0.001)⁽³⁶⁾. Peeters et al.⁽³⁷⁾ report that their surgical group had a 72% lower hazard of death compared to the control population group (HR 0.28; 95% CI 0.1-0.85) after a median follow up time of 4 years.

1.1.5 Bariatric Surgery as a Treatment for Severe Obesity and its Impact on HRQoL

There is substantial evidence to suggest that bariatric surgery for the treatment of severe obesity improves HRQoL in patients' post-surgery compared to severely obese groups either waiting for surgery or not seeking treatment in the short term^(27, 38). The SOS⁽²⁷⁾ reports the largest improvement in HRQoL at 1 year post surgery. In the short term, significant improvements have been reported post-surgery for general health, physical HRQoL (e.g.,

mobility, pain, usual activities) and weight-specific HRQoL with significant but less impact on mental or emotional health^(27, 38).

The short term results are mixed on the type and magnitude of the association between the degree of weight loss and degree of HRQoL improvement. The first 6-18 months is the peak time for weight loss post bariatric surgery⁽²⁵⁾. A inverse association between the amount of excess weight lost and an improvement in HRQoL has been reported showing the peak weight loss and the largest improvement on several HRQoL surveys at 2 years post-surgery⁽²⁷⁾. In contrast, another study reports that the degree of body mass reduction does not influence the degree of HRQoL improvement⁽³⁹⁾. However this study also found that the surgical population showed positive HRQoL results that exceeded those of the general population at one year post surgery. A recent SR also reported that out of 24 studies with varying follow-up time intervals most treatment seeking populations demonstrated a significant improvement in HRQoL at 1 year post surgery⁽⁴⁰⁾.

There are very few studies that report on HRQoL 5 years or greater post-surgery, but in general in those that do, the authors suggest that bariatric surgery is associated with greater improvements in both general and obesity specific measures of HRQoL between six to ten years after surgery when compared to non-surgical care^(27, 41, 42). However, the differences in association with weight loss and HRQoL also necessitate a need for long term results as weight loss is known to slow down or reverse in the long-term post bariatric surgery⁽²⁷⁾.

1.1.6 Gaps in the Literature

Little is known about the impact of bariatric surgery on HRQoL in patients in the long term (≥ 5 years) post bariatric surgery⁽⁴³⁾. Study results are often limited to assessing physical quality of life with less focus and reporting of mental and emotional quality of life⁽⁴¹⁾. A recent

study has shown that bariatric surgery results in a greater improvement in physical HRQoL than in mental health scores⁽⁴⁴⁾. In the long term, at least one study demonstrates that post- surgery HRQoL varies and is dependent on the maintenance of weight loss⁽²⁷⁾. Given the increasing number of bariatric surgeries being performed, it is important to know if improved HRQoL scores observed in the early stages post bariatric surgery are maintained over time.

1.2 Purpose of Study

The overall purpose of the current study was to examine the impact of bariatric surgery as a treatment for severe obesity on long term (≥ 5 years) HRQoL in patients who have undergone bariatric surgery.

1.3 Significance of the Study

Severe obesity significantly impairs HRQoL, and although bariatric surgery results in significant weight loss in the majority of patients, there is far less known about its impact over the long term (≥ 5 years), an equally important outcome for patients. Although a short term study has assessed the impact of bariatric surgery on HRQoL and have reported physical and mental health improvements⁽⁴⁵⁾, very few studies report longer term outcomes, and of those published there are inconsistent results in physical and mental health^(27, 41, 46-53).

Severe obesity is increasing disproportionately in Canada and many other countries^(7, 54, 55-57), and in response so is the number of bariatric surgeries performed⁽³³⁾. It is increasingly important to determine the effectiveness of surgery not only on weight loss and clinical outcomes but on patient reported outcomes such as HRQoL

1.4 Specific Research Objectives

The specific research objectives of the current study are:

- 1.** to develop a protocol and search strategy focused on conducting a SR
- 2.** to assess the quality of evidence and effectiveness of bariatric surgery on the long term (\geq 5 years) HRQoL of patients (\geq 18 years) compared to non-surgical controls that include patients living with severe obesity not seeking treatment or those patients waiting for surgery, and
- 3.** to conduct a MA of appropriate studies to examine the impact of bariatric surgery on mental and physical health domains, if feasible.

1.5 Chapter 1 References

1. WHO Consultation on Obesity: Obesity: preventing and managing the global epidemic. Report of a WHO consultation. *World Health Organ Tech Rep Ser.* 2000; No. 894:i-xii, 1-253.
2. Health Canada: Canadian *guidelines for body weight classifications in adults.* 2003; Available from: http://preventdisease.com/pdf/weight_book-livres_des_poids_e.pdf [Accessed January 20th, 2015]
3. Lau DC, Douketis JD, Morrison KM, et al.: 2006 Canadian clinical practice guidelines on the management and prevention of obesity in adults and children [summary]. *CMAJ.* 2007; 176(Suppl 8):S1-13.
4. Twells LK, Bridger T, Knight JC, et al.: Obesity predicts primary health care visits: a cohort study. *Popul Health Manag.* 2012; 15(1):29-36.
5. Flegal KM, Kit BK, Orpana H, Graubard BI: Association of all-cause mortality with overweight and obesity using standardized body mass index categories. A systematic review and meta-analysis. *JAMA.* 2013; 309(1):71-82.
6. Sarwer DB, Lavery M, Spitzer JC: A review of the relationships between extreme obesity, quality of life and sexual function. *Obes Surg.* 2012; 22(4):668-676.
7. Twells L, Gregory DM, Reddigan J, Midodzi W: Current prevalence and future predictions of obesity in Canada: A trend analysis. *CMAJ Open.* 2014; 2(1): E18-26.
8. Guyatt GH, Feeny DH, Patrick, DL: Measuring Health Related Quality of Life. *Ann Intern Med.* 1993; 118(8): 622-629
9. EuroQol Research Foundation: EQ-5D is a standardized instrument for use as a measure of health outcomes. Available from: <http://www.euroqol.org/about-eq-5d.html> [Accessed October 15th 2015]
10. Wee CC, Davis RB, Hamel MB: Comparing the SF-12 and the SF-36 health status questionnaires in patients with and without obesity. *Health Qual Life Outcomes.* 2008; 6:11
11. Ware J, Snow K, Kosinski M, Gandek B. SF-36 health survey: Manual and interpretation Guide. Boston: Health Institute, New Eng-land Medical Center: 1993
12. Kolotkin RL: Quality of Life Consulting: Impact of Weight on Quality of Life-Lite (IWQOL-Lite) Available from: <http://www.qualityoflifeconsulting.com/iwqol-lite.html> [Accessed October 15th 2015]

13. van Nunen AM, Wouters EJ, Vingerhoets AJ, Hox JJ, Geenen R: The Health-related quality of life of obese persons seeking or not seeking surgical or non-surgical treatment: a meta-analysis. *Obes Surg.* 2007; 17(10):1357-1366
14. Herpertz S, Kielmann R, Wolf AM, Langkafel M, et al.: Does obesity surgery improve psychosocial functioning? A systematic review. *Int J Obes Relat Metab Disord* 2003; 27(11):1300-1314
15. Duval K, Marceau P, Lescelleur O, Hould FS, Marceau S, Biron S, Lebel S, Pérusse L, Lacasse Y: Health-related quality of life in morbid obesity. *Obes Surg* 2006; 16(5): 574-579
16. Jia H & Lubetkin EI: The impact of obesity on health-related quality-of-life in the general adult US population. *J Public Health. (Oxf)* 2005; 27(2): 156-164
17. Fontaine KR, Cheskin LJ & Barofsky I: Health-related quality of life in obese persons seeking treatment. *J Fam Pract.* 1996; 43(3): 265-270
18. Sullivan M, Karlsson J, Sjöström L, Backman L, Bengtsson C, Bouchard C, Dahlgren S, Jonsson E, Larsson B & Lindsedt S et al: Swedish obese subjects (SOS) - an intervention study of obesity. Baseline evaluation of health and psychosocial functioning in the first 1743 subjects examined. *Int J Obes Relat Metab Disord.* 1993; 17(9): 503-512
19. Moyer VA, US Preventive Services Task Force: Screening for and management of obesity in adults: US Preventive Services Task Force recommendation statement. *Ann Intern Med.* 2012; 157(5):376-378.
20. Avenell A, Broom J, Brown TJ, et al.: Systematic review of the long term effects and economic consequences of treatments for obesity and implications for health improvement. *Health Technol Assess.* 2004; 8(21):iii-iv, 1-182.
21. LeBlanc E, O'Connor E, Whitlock EP, Patnode C, Capka T: Screening for and management of obesity and overweight in adults. *Rockville (MD): Agency for Healthcare Research and Quality (US).* 2011; Report No. 11-05159-EF-1
22. 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(3):169-182
23. Padwal RS, Rueda-Clausen CF, Sharma AM, Aqhorsangaya CB et al.: Weight loss and outcomes in wait-listed, medically managed and surgically treated patients enrolled in a population-based bariatric program: prospective cohort study. *Med Care.* 2014; 52(3):208-215

24. Sjöström L, Nabro K, Sjöström CD et al.: Effects of bariatric surgery on mortality in Swedish obese subjects. *N Engl J Med.* 2007; 357(8):741-752
25. Chang SH, Stoll CR, Colditz GA: Cost-effectiveness of bariatric surgery: should it be universally available? *Maturitas.* 2011; 69(3):230-238
26. Christou NV, Sampalis JS, Liberman M et al.: Surgery decreases long-term mortality, morbidity, and health care use in morbidly obese patients. *Ann Surg.* 2004; 240(3):416-423; discussion 423-424
27. Karlsson J, Taft C, Rydén A et al.: 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(8):1248-1261
28. Belle SH, Berk PD, Courcoulas AP, Flum DR, Miles CW, Mitchell JE, et al.: Safety and efficacy of bariatric surgery: Longitudinal assessment of bariatric surgery. *Surg Obes Relat Dis.* 2007; 3(2):116-126
29. Padwal R, Klarenbach S, Wiebe N, Hazel M, Birch D, Karmali S, et al.: Bariatric surgery: a systematic review of the clinical and economic evidence. *J Gen Intern Med.* 2011; 26(10):1183-1194
30. Picot J, Jones J, Colquitt JL, Gospodarevskaya E, Loveman E, Baxter L, et al.: The clinical effectiveness and cost-effectiveness of bariatric (weight loss) surgery for obesity: a systematic review and economic evaluation. *Health Technol Assess.* 2009; 13(41):1-190, 215-357, iii-iv
31. Stefater MA, Wilson-Pérez HE, Chambers AP, Sandoval DA, & Seeley RJ: All bariatric surgeries are not created equal: insights from mechanistic comparisons. *Endocr Rev.* 2012; 33(4):595-622
32. Buchwald H, Oien DM: Metabolic/bariatric surgery worldwide 2008. *Obes Surg.* 2009; 19(12):1605-1611
33. Angrisani L, Santonicola A, Iovino P, Formisano G, Buchwald H, Scopinaro N. Bariatric surgery worldwide 2013. *Obes Surg.* 2015; 25(10):1822-1832
34. Mechanick JI, Youdiu A, Jones DB et al.: Clinical practice guidelines for the perioperative nutritional, metabolic, and nonsurgical support of the bariatric surgery patient-2013 update: cosponsored by American Association of Clinical Endocrinologists, the Obesity Society, and American Society for Metabolic & Bariatric Surgery. *Endocr Pract.* 2013; 19(2):337-372
35. Pontiroli AE, Morabito A: Long-term prevention of mortality in morbid obesity through bariatric surgery. A systematic review and meta-analysis of trials performed with gastric

banding and gastric bypass. *Ann Surg.* 2011; 253(3):484-487

36. Adams TD, Gress RE, Smith SC et al.: Long-term mortality after gastric bypass surgery. *N Engl J Med.* 2007; 357(8):753-761
37. Peeters A, O'Brien PE, Laurie C et al.: Substantial intentional weight loss and mortality in the severely obese. *Ann Surg.* 2007; 246(6):1028-1033.
38. Kolotkin RL, Crosby RD, Gress RE, Hunt SC, Adams TD: Two-year changes in health-related quality of life in gastric bypass patients compared with severely obese controls. *Surg Obes Relat Disord.* 2009; 5:250–256.
39. Major P, Matłok M, Pędziwiatr M, Migaczewski M, Budzyński P, Stanek M, Kisielewski M, Natkaniec M, and Budzyński A.: Quality of life after bariatric surgery. *Obes Surg.* 2015; 25(9): 1703-1710 doi: [10.1007/s11695-015-1601-2](https://doi.org/10.1007/s11695-015-1601-2) [Accessed Novemembr 12th 2015]
40. Raaijmakers LCH, Pouwels S, Thomassen SEM, Nienhuijs SW: Quality of life an bariatric surgery: a systematic review of short- and long-term results and comparison with community norms. *European Journal of Clinical Nutrition.* 2016; 1-9
41. 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; 8(5):625–633
42. Nickel MK, Loew TH, Bachler E: Change in mental symptoms in extreme obesity patients after gastric banding part II: six year follow-up. *International J of Psychiatry Med.* 2007; 37(1): 69-79
43. Arterburn DE, Fisher DP: The current state of the evidence for bariatric surgery. *JAMA.* 2014; 312(9):898-899
44. Aasprang A, Andersen, JR, Våge V, Kolotkin RL, Natvig GK. Five-year changes in health-related quality of life after biliopancreatic diversion with duodenal switch. *Obes. Surg.* 2013;23(10):1662-1668
45. Magallares A, Schomerus G. Mental and physical health-related quality of life in obese patients before and after surgery: A meta-analysis. *Psychol Health Med.* 2015;20(2):165-176
46. deZwaan M, Lancaster KL, Mitchell JE, et al. Health-related quality of life in morbidly obese patients: effect of gastric bypass surgery. *Obes Surg* 2002;12:773-780.
47. Kiewiet RM, Durian MF, Cuijpers LPLH, Hesp FLEM, van Vliet ACM. Quality of life after gastric banding in morbidly obese Dutch patients: long-term follow-up. *Obes Res Clin Pract* 2008;2:151-158.

48. Laurino Neto RM, Herbella FAM. Changes in quality of life after short-term and long-term follow-up of roux-en-Y gastric bypass for morbid obesity. *Arg Gastroenterol* 2012;**50**(3):186-190.
49. Sanchez-Santos R, Delarrio MJ, Gonzalez C, et al. Long-term health-related quality of life following gastric bypass: influence of depression. *Obes Surg* 2006;**16**:580-585.
50. Våge V, Solhaug JH, Viste A, et al. Anxiety, depression and health-related quality of life after jejunioileal bypass: a 25-year follow-up study of 20 female patients. *Obes Surg* 2003;**3**:706-713.
51. Aftab H, Risstad H, Sovik TT, et al. Five-year outcome after gastric bypass for morbid obesity in a Norwegian cohort. *Surg Obes Relat Dis*. 2014;**10**:71-78.
52. Raouf M, Näslund I, Rask E, et al. Health-related quality-of-life (HRQoL) on average of 12 years after gastric bypass surgery. *Obes Surg* 2015;**25**:1119-1127.
53. Andersen JR, Aasprang A, Karlsen TI, Vage V, Kolotkin RI. Health related quality of life after bariatric surgery: a systematic review of prospective long-term studies. *Surg Obes Relat Dis* 2014;**2**:466-473.
54. Sassi F, Devaux M, Cecchini M, Rusticelli E :The obesity epidemic: analysis of past and projected future trends in selected OECD countries. *Organisation for Economic Co-operation and Development (OECD), Directorate for Employment, Labour And Social Affairs, Health Committee*, 2009
55. WHO. *World Health Report: Reducing risks, promoting healthy life*. World Health Organization. 2002
56. WHO. *Global strategy on diet, physical activity and health: a framework to monitor and evaluate implementation*. World Health Organization. 2008
57. Finucane M.M, Stevens M.A, Cowan M.J, et al. National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9·1 million participants *Lancet*. 2011; 377: 557–567

Chapter 2 Protocol

2.1 Unpublished *A Priori* Protocol for Bariatric Surgery and its Impact on Quality of life in Adults: A Systematic Review

Anticipated or Actual Start Date: May 2014

Anticipated Completion Date: August 2014

Review Team Members:

Shannon Driscoll

Dr. Laurie Twells

Dr. John Fardy

Dr. Deborah Gregory

Funding Sources/Sponsors: Healthcare Foundation, Eastern Health

2.1.1 Introduction

The World Health Organization has acknowledged the prevalence of obesity as a worldwide epidemic^(1,2). The WHO defines overweight and obesity as “abnormal or excessive fat accumulation that presents a risk to health”⁽²⁾. Overweight and obesity are commonly measured using the BMI calculated from weight in kilograms divided by height in metres squared (kg/m^2). Overweight is considered to be a $\text{BMI} \geq 25 \text{kg}/\text{m}^2$ and obesity is considered to be a $\text{BMI} \geq 30 \text{kg}/\text{m}^2$ ^(2,3). Obesity is something that affects both men and women and is prevalent in adult and child/adolescent populations. It has been suggested that obesity is caused by the imbalance between high energy intake generally associated with fat and carbohydrate rich diets, a sedentary lifestyle that translates to low energy expenditure and in some cases a level of genetic susceptibility that leads to obesity^(1,4). Excessive fat is associated with a range of serious non-communicable disease including diabetes mellitus, various cardiovascular diseases, musculoskeletal disorders, some cancers and others that lead to premature mortality⁽²⁾. Along with these comorbidities, there is a reduction in HRQoL^(1,5).

Bariatric surgery has proved to be the only most effective treatment for obesity and is associated with significant and sustained weight loss⁽⁶⁾. Most bariatric surgeries are performed on patients who are considered severely obese and have been assessed by a team of medical professionals. Severe obesity is defined as having a BMI \geq 40kg/m² or a BMI \geq 35kg/m² with at least one comorbid condition^(1,7). The ideal bariatric surgery patients also have the following characteristics⁽⁷⁾: acceptable operative risk, documentation of failure with non-surgical weight-loss programs, thorough understanding of procedures, realistic expectations and well-informed and motivated patients.

Bariatric surgery can be restrictive, mal-absorptive or a combination of the two. In restrictive surgeries, the size of the stomach is reduced to limit food intake and slow down the digestive process. Mal-absorptive procedures normally reduce the exposure of food to the intestine which limits the amount of calories absorbed by the body. This is achieved through bypassing or removal of parts of the intestine. Procedures that use a combination of the two methods both restrict the stomach size as well as remove or by-pass portions of the intestine⁽⁸⁾.

The gold standard of bariatric surgeries has been the Roux-en-y gastric bypass (RYGB) surgery^(7,9). In a RYGB, the stomach is reduced to a small pouch and attached directly to the jejunum, bypassing the duodenum completely. It produces weight loss by both restricting the ability to consume large amounts of food as well as by being a mal-absorptive procedure through a reduced small intestine⁽⁷⁾.

Another common procedure is the laparoscopic adjustable gastric band. In this procedure an inflatable device is placed around the top portion of the stomach to reduce the size and slow the intake of food. It is a restrictive, minimally invasive procedure⁽⁸⁻¹⁰⁾.

The sleeve gastrectomy has recently increased in popularity as a bariatric procedure. In this procedure, the stomach is surgically reduced to a sleeve-like tube by removing around 80% of the stomach. In this procedure, the intestinal pathway stays intact^(8,9,11).

Regardless of what type of bariatric surgery is used, there are various benefits and complications associated with each. Surgeries can be either “open” or “laparoscopic”. Open surgeries are associated with more complications post-surgery as they are more invasive procedures. While laparoscopic surgeries are less invasive, it is not always a suitable procedure for patients that are too obese or have had previous abdominal surgeries^(12,13).

The primary outcome of bariatric surgeries for obesity is weight loss. The largest weight loss period post-surgery is the first 6-12 months. After this period, there is a tapering off of weight loss and in some cases weight regain occurs^(7,8,10). The resolution of co-morbidities is also an important outcome of bariatric surgery. There is also general improvement of mobility, which in turn can help with increased overall fitness levels and in maintaining weight loss⁽⁶⁻⁹⁾.

HRQoL measurements encompass measurements of well-being, functioning and health under physical, social and psycho-social domains or categories⁽¹⁴⁾. The HRQoL scores are most often recorded through a variety of questionnaires, each having their own strengths and areas of focus. There are general and specific quality of life measurements. General surveys are designed to assess to general well-being and can be administered to a variety of populations. Disease specific surveys relate directly to a certain disease or condition, obesity or excess weight in this case, that may impact quality of life.

2.1.2 Study Measurement Tools

The Newfoundland and Labrador Bariatric Surgery Cohort Study (NL BaSCo Study) uses a combination of 3 validated HRQoL measurements to assess enrolled patients. They are the

SF-12v2 (adapted from the Short Form 36 tool), the Impact of Weight on Quality of Life-Lite (IQWOL-Lite) and the European Quality of Life-5 dimensions-3Level (EQ-5D-3L). The instruments are described below.

The SF-12 (Adapted from the SF-36 tool)⁽¹⁵⁾ is a validated questionnaire that has been well established to reflect patient's overall health. The survey is short to complete and gives an overall view of the health status of the patient. The survey has 12 questions that examine 8 different quality of life domains: physical function, role physical, bodily pain, general health, vitality, social function, role emotional and mental health.

The EQ-5D-3L⁽¹⁶⁾ is a generic measure of health and is useful for population health surveys as well as clinical and economic healthcare evaluations. The survey is scored on 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. These are then ranked at 3 levels: no, some or extreme problems. The responses are pooled and represented on a visual analog scale ranging from 'Worst imaginable health state' to "best imaginable health state". This is a single-index value that can be used as a quantitative measure of health judged by the individual patient.

The IWQOL-Lite assesses a patient's perception on how their weight affects their daily life. It is comprised of 31 questions/items and is divided into 5 scales: physical function, self-esteem, sexual life, public distress and work. Each item begins with the phrase: "Because of my weight..." and patients have 5 response options from 1 = never true to 5 = always true.

It has been used in bariatric surgery patient populations^(17, 18).

There is variation in the maintenance of weight loss in long term post- operative results. Quality of life outcomes are less studied and it is important to know if the improved scores seen in early stages post-bariatric surgery are maintained over time. Improvements in HRQoL have

become an important outcome of these surgeries. Since HRQoL is a secondary outcome for bariatric surgery, the reporting is often sporadic, and a wide variety of measurements that are not always validated are used. A recent study has shown that weight loss shows a greater improvement in physical HRQoL than in mental scores⁽¹⁷⁾. These findings may be attributed to the use of specific versus generic tools that may result in significant differences in the HRQoL scores recorded. Second, the type of study may also affect the amount of HRQoL improvement seen. For example, a randomized control trial is much stricter in inclusion and exclusion criteria than an observational study. For these reasons the HRQoL changes seen in an RCT may not accurately reflect real world results.

This SR will pool the results of both randomized control trials and observational studies to compare generic and specific tools, mental and physical scores and summary component scores.

2.1.3 Reason for a Systematic Review

An SR of the literature will help assess the overall impact of bariatric surgery on quality of life in obese patients. If possible, a MA of validated questionnaires will be performed to analyze the pooled results. While there have been some general reviews of the literature, and at least one SR looking at the effects of a variety of weight loss methods on HRQoL⁽¹⁹⁾, to our knowledge there are no previously published SRs addressing this question. This review will compare specific and generic tools, mental and physical components in randomized control trials and observational studies.

Review Objective:

To assess the impact of bariatric surgery for obesity on quality of life in adults 18 and older.

Types of study to be included:

- (1) Randomised controlled trials (RCTs) (published or unpublished).
- (2) Non-randomised studies (NRS; concurrent and non-concurrent cohort studies)

Participants/Population: adults (age 18+) who have undergone bariatric surgery for the treatment of either class II (with at least one comorbid condition) or class III obesity.

Intervention/exposure: Bariatric surgery AND participate in a general or weight specific measurement of QOL.

Comparator/Control: obese population/wait listed patients or pre-surgery results.

Primary Outcomes (At the time of protocol submission –May 2014):

- Measures of quality of life

Secondary Outcomes:

- Weight loss
- Resolution of comorbid conditions
- Surgical complications

Electronic Searches

PubMed, Embase, CINAHL, PsycINFO

MeSH : bariatric surgery, quality of life

The search strategy will be designed with the assistance of a medical librarian and approved by the members of the research team before starting. Articles will be filtered by age for 18+. There will be no language or publication date restrictions. Reference lists of included studies and any relevant SRs identified will also be searched.

Searching Other Resources

Authors of major papers/studies will be contacted to see if there is any unpublished material or additional conference or abstract reports. Grey material will be searched for bariatric surgery abstracts and surgery protocols. Major conference meeting notes will also be searched. These include the Canadian Obesity Conference, European Congress on Obesity and The Obesity Society (TOS) Annual Scientific Meeting. Reference lists of included papers will be thoroughly searched for any additional studies missed in the preliminary searches.

2.1.4 Study Selection and Data Extraction

One reviewer (SD) will perform a preliminary search using the search strategy approved by other members of the research team. After the preliminary search, two reviewers (SD,LT) will independently scan study titles, abstracts, and keywords of every non-duplicate record retrieved from the literature search. Irrelevant titles will be excluded and full-text papers will be obtained where titles are deemed to be relevant or where eligibility is unclear. A record of why each study is rejected will be kept to include in the final SR for the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart and ‘characteristics of excluded studies’ chart. Two independent reviewers will identify studies using a study eligibility/relevance form. Discrepancies will be resolved by consensus. Where pertinent information is missing, authors of the papers will be contacted for the additional information. Following identification of the included studies, each reviewer will independently extract relevant study data. Relevant missing data will be obtained from authors, if feasible. Where discrepancies arise, consensus will be reached, and a third reviewer will adjudicate if necessary. References found from online databases will be managed and stored using RefWorks.

Data extraction will follow the altered Cochrane Public Health Group data extraction forms. From each paper outcomes involving weight/BMI, comorbid conditions and HRQoL will be recorded as well as population demographics for both intervention and control/comparison groups.

2.1.5 Risk of Bias Assessment

Two reviewers will assess the risk of bias or ‘quality’ of each study. Randomized control trials will be assessed using the ‘The Cochrane Collaboration’s tool for assessing risk of bias’ Observational studies will be assessed using the Newcastle-Ottawa Scale for assessing the quality of nonrandomized studies. Each reviewer will make notes on hard copies of the papers with the reasoning for their quality assessment scores to be included in the SR.

2.1.6 Measure of Treatment Effect

Standardized mean differences will be used where possible as studies may use different HRQoL tools. Weight loss and resolution of co-morbidities will be reported as changes from mean scores.

2.1.7 Strategy for Data Synthesis

Results between studies will be compared by study design (RCT versus observational), physical versus mental components and specific versus generic tools. The results will include a summary of findings table including the outcomes for each study on HRQoL, weight loss, and reduction in co-morbidities. The table will also include the type of surgery, number of participants in each study and the quality of the study.

Assessment of Heterogeneity

The I^2 statistic will be used to quantify the level of heterogeneity.

2.2 Complete Search Strategy

PubMed/MEDLINE (834)

- 1) "Bariatric Surgery"[Mesh] OR "bariatric surgery" [Title/Abstract]
- 2) "gastrectomy"[mesh] OR gastrectomy [Title/Abstract]
- 3) #1 OR #2

- 4) "Obesity"[Mesh] OR obesity [Title/Abstract]
- 5) "Morbid Obesity" [Mesh] OR "Morbid Obesity" [Title/Abstract]
- 6) #4 or #5

- 7) "Quality of Life"[Mesh] OR "Quality of life" [Title/Abstract]
- 8) "psychometrics"[mesh] OR psychometrics [Title/Abstract]
- 9) "health-related quality of life" [Title/Abstract]
- 10) "Short-Form" OR "Short Form" SF36 OR "SF-36" OR SF12 OR "SF-12" OR "SF-6D" OR SF6D
- 11) EQ-5D OR "Euroqol*" OR "visual analogue scale"
- 12) "health state preference"
- 13) "BAROS"
- 14) "obesity and weight loss quality of life" OR OWLQOL
- 15) #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14

- 16) #3 AND #5 AND #16

Embase (691)

- 1) 'bariatric surgery'/exp OR "Bariatric Surgery":ti,ab
- 2) 'gastrectomy'/exp OR "gastrectomy":ti,ab
- 3) 'stomach bypass'/exp OR 'gastric bypass':ab,ti
- 4) 'roux y anastomosis'/exp OR 'roux en y':ab,ti
- 5) #1 OR #2 OR #3 OR #4

- 6) 'obesity'/exp OR 'obesity':ab,ti
- 7) 'morbid obesity'/exp OR 'morbid obesity':ab,ti
- 8) #6 OR #7

- 9) 'quality of life'/exp OR 'quality of life':ab,ti
- 10) 'psychometry'/exp OR 'psychometry':ab,ti
- 11) 'eq 5d' OR 'euroqol' OR 'visual analogue scale'
- 12) OR 'health state preference'
- 13) 'baros'
- 14) 'obesity and weight loss quality of life' OR owlqol
- 15) #9 OR #10 OR #11 OR #12 OR #13 OR #14

- 16) #5 AND #8 AND #16

- 17) #17 AND ([adolescent]/lim OR [adult]/lim OR [aged]/lim OR [middle aged]/lim OR [young adult]/lim)

CINAHL (148)

- 1) (MH "Bariatric Surgery+") OR (MH "Bariatric Patients")OR TI "bariatric surgery" AND AB "bariatric surgery"
- 2) (MH "Gastrectomy") OR TI "gastrectomy" OR AB gastrectomy
- 3) (MM "Gastric Bypass") OR TI gastric bypass OR AB gastric bypass
- 4) (MM "Anastomosis, Roux-en-Y") OR TI roux en y OR AB roux en y
- 5) #1 OR #2 OR #3 OR #4

- 6) (MM "Obesity+") OR (MM "Obesity, Morbid")
- 7) TI (obesity OR "morbid obesity") OR AB (obesity OR "morbid obesity")
- 8) #6 OR #7

- 9) (MM "Quality of Life+") OR TI "Quality of Life" OR AB "Quality of Life"
- 10) (MM "Psychometrics") OR TI "Psychometrics" OR AB "Psychometrics"
- 11) (MM "Short Form-36 Health Survey (SF-36)") OR TI Short Form-36 (SF-36) OR AB Short Form-36 (SF-36) OR TI "short-form 12" OR SF-12 OR AB "short-form 12" OR SF-12
- 12) TX "eurqol" OR TX eq5d OR TX "visual analogue scale"
- 13) TX ("health state preference")
- 14) TX BAROS
- 15) TX 'obesity and weight loss quality of life' OR owlqol
- 16) #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15
- 17) #5 AND #8 AND #17

PsycInfo (95)

- 1) Bariatric surgery AND Obesity AND Quality of life

Cochrane Review (132)

- 1) "Bariatric Surgery"[Mesh] OR "bariatric surgery" [Title/Abstract]
- 2) "gastrectomy"[mesh] OR gastrectomy [Title/Abstract]
- 3) #1 OR #2
- 4) "Obesity"[Mesh] OR obesity [Title/Abstract]
- 5) "Morbid Obesity" [Mesh] OR "Morbid Obesity" [Title/Abstract]
- 6) "Quality of Life"[Mesh] OR "Quality of life" [Title/Abstract]
- 7) "psychometrics"[mesh] OR psychometrics [Title/Abstract]
- 8) "health-related quality of life" [Title/Abstract]

- 9) "Short-Form" OR "Short Form" SF36 OR "SF-36" OR SF12 OR "SF-12" OR "SF-6D"
OR SF6D
- 10) EQ-5D OR "Euroqol*" OR "visual analogue scale"
- 11) "health state preference"
- 12) "BAROS"
- 13) "obesity and weight loss quality of life" OR OWLQOL
- 14) # 6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14

2.3 Chapter 2 References

1. WHO Consultation on Obesity: Obesity: preventing and managing the global epidemic. Report of a WHO consultation. *World Health Organ Tech Rep Ser.* 2000; No. 894:i-xii, 1-253.
2. World Health Organization. Obesity and Overweight Fact Sheet. 2013. Available from: <http://www.who.int/mediacentre/factsheets/fs311/en/index.html> [Accessed May 23rd 2014]
3. Health Canada: Canadian guidelines for body weight classifications in adults.2003; Available from: http://preventdisease.com/pdf/weight_book-livres_des_poids_e.pdf [Accessed May 23rd 2014]
4. National Institutes of Health fact sheet: What causes overweight and obesity? 2012. Available from: <http://www.nhlbi.nih.gov/health/health-topics/topics/obe/causes.html> [Accessed May 23rd 2014]
5. Kopleman, P. Health risks associated with overweight and obesity. *Obesity Reviews.* 2007; 8(s1) 13-17
6. Colquitt JL, Picot J, Loveman E, Clegg AJ. Surgery for obesity. *Cochrane Database Syst Rev*2009;**15**(2):CD003641. doi:[10.1002/14651858.CD003641.pub3](https://doi.org/10.1002/14651858.CD003641.pub3). [Accessed September 2014]
7. Karmali, S., Johnson Stoklossa, C., Sharma, A., Christiansen, S., Cottreau, D. & Birch, D.W. Bariatric Surgery: A primer. *Can Fam Physician.* 2010; 56(9), 873-879
8. Canadian Agency for Drugs and Technologies in Health (CADTH): Bariatric surgery for severe obesity: systematic review and economic evaluation. 2010. Available from: <https://www.cadth.ca/bariatric-surgery-severe-obesity-systematic-review-and-economic-evaluation> [Accessed September 2014]
9. Padwal, R., Klarenbach, S., Wiebe, N., Hazel, M., Birch, D., Karmali, S. & Tonelli, M. Bariatric surgery: a systematic review of the clinical and economic evidence. *Journal of General Internal Medicine.* 2011; 26(10), 1183-1194.
10. Dixon, J.B., Straznicky, N.E., Lambert, E.A., Slaich, M.P. & Lambert, G.W. Laparoscopic adjustable banding and other devices for the management of obesity. *Circulation.* 2012; 126, 774-785
11. Himpens, J., Dobbeleir, J. & Peeters, G. Long-term results of laparoscopic sleeve gastrectomy for obesity. *Surgery* 2010. 252 (2)

12. Velanovich, V. Laparoscopic vs. open surgery: a preliminary comparison of quality-of-life outcomes. *Surg. Endosc.* 2000; *14*(1) 16-21
13. Reoch, J., Mottillo, S., Shimony, A., Filion, K.B., Christou, N.V., Josheph, L., Poirier, P. & Eisenberg, M.J. Safety of laparoscopic versus open bariatric surgery: a systematic review and meta-analysis. *Archives of Surgery.* 2011; *146*(11) 1314-1322
14. Guyatt G, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993;**118**:622-629
15. Ware J, Snow K, Kosinski M, Gandek B. *SF-36 Health Survey: Manual and Interpretation Guide.* Boston: Health Institute, New Eng-land Medical Center; 1993.
16. EuroQol Research Foundation: EQ-5D is a standardized instrument for use as a measure of health outcomes. Available from: <http://www.euroqol.org/about-eq-5d.html> [Accessed October 15th 2015]
17. Kolotkin, R.L., Crosby, R.D., Gresse, R.E., Hunt, S.C. & Adams, T.D. Two year changes in health-related quality of life in gastric bypass patients compared with severely obese controls. *Surg. Obes. Rel. Dis* 2009; *5*(2) 250-256
18. Aasprang, A., Andersen, J.R., Vage, V., Kolotkin, R.L. & Natvig, G.K. Five-year changes in health-related quality of life after biliopancreatic diversion with duodenal switch. *Obes. Surg.* 2013; *23*(10) 1662-1668
19. 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**:169-182

Chapter 3 Long-Term Health-Related Quality of Life in Bariatric Surgery Patients: A Systematic Review and Meta-Analysis

A version of this chapter has been published in *Obesity* (2016) 24, 60–70.
doi:10.1002/oby.21322

IF (impact Factor) 3.614

Shannon Driscoll¹, Deborah M. Gregory^{1,2}, John M. Fardy², and Laurie K. Twells^{1,3}

¹Department of Clinical Epidemiology, Faculty of Medicine, Memorial University of Newfoundland

²Department of Medicine, Faculty of Medicine, Memorial University of Newfoundland

³School of Pharmacy, Memorial University of Newfoundland

3.1 Introduction

Bariatric surgery is considered to be the only treatment for severe obesity that results in significant weight loss⁽¹⁾. Individuals are eligible for bariatric surgery if they have a BMI \geq 40kg/m² or a BMI \geq 35kg/m² with at least one obesity-related comorbidity^(2,3). Impaired quality of life is common in bariatric surgery candidates and is often one of the motivating factors for seeking surgery. Along with weight loss and resolving co-morbidities, improving HRQoL is an important outcome. HRQoL encompasses measures of well-being, functioning and health under physical, social and psycho-social domains⁽⁴⁾.

Short to mid-term (<5 years) HRQoL results post-surgery are well documented and show a significant improvement in physical health scores often reaching population normative values⁽⁵⁻⁹⁾. However, after this “honeymoon” period, there can be a plateau or weight recidivism⁽¹⁰⁾. Longer-term HRQoL data are scarce and in the published studies available limitations noted include: selective reporting, non- standardized presentation of results, use of non-validated measurement tools, and reporting of HRQoL as a secondary outcome with incomplete and inconsistent data. The long-term impact of bariatric surgery on HRQoL is unknown^(11,-14). Studies have shown a greater improvement in physical HRQoL than in mental health scores as these domains more directly benefit by a reduction in body weight⁽¹⁵⁾. The mental health outcomes are inconsistent^(11, 16-24). There have been general reviews published and at least one SR examined the effects of a variety of weight loss methods on HRQoL⁽²⁵⁾, but to our knowledge there is only one previously published SR addressing long term (i.e., \geq 5years) HRQoL after bariatric surgery⁽²⁴⁾. In that SR, the authors found that HRQoL after bariatric surgery improved significantly in the short term, declined slightly after 2 years and appeared to stabilize 5 years post-operatively. The authors included long-term, prospective studies, however not all studies

used a control or comparator group and some used a normal weight comparator while others used baseline results. There was no MA performed because the authors considered it inappropriate due to heterogeneity. The aim of our study was to conduct a SR of studies reporting health-related quality life data at least five years after bariatric surgery in patients 18 years of age and older and in non-surgical obese comparison control groups. An MA of the HRQoL results was planned if deemed appropriate.

3.2 Methods

3.2.1 Search Strategy

Eligibility criteria for studies included patients 18 or older who received bariatric surgery (type of surgery was not limited) for severe obesity defined as a BMI \geq 35kg/m² with a comorbidity or a BMI \geq 40kg/m²(²). For a study to be included in the review, a non-surgical obese comparison group with similar demographics to the surgical group at baseline had to be available. Follow-up time post-surgery was at least 5 years. The completion of a validated HRQoL survey at follow up was required. There were no publication type or publication date restrictions. In the protocol the plan was to include all languages. Once the initial searches were conducted, the articles that came back in other languages were available in English or did not meet the other study criteria based on the results. Therefore the language of study was limited to English. The study type was not limited and reviews including SRs were included in the search.

The initial search strategy was designed with the assistance of a medical librarian and approved by members of the research team. Keywords were formatted for each search engine. PubMed, Cochrane Review, EmBase, CINAHL and PsycInfo were searched. The Medical Search headings (MeSH) used were: “bariatric surgery”, “gastrectomy”, “obesity” “morbid obesity”, “quality of life” and “psychometrics”. Each of these was also searched as a keyword in

[ti] and [ab] so as not to miss un-indexed articles. In addition to the MeSH terms, the following quality of life questionnaires were also included: health-related quality of life, "Short-Form" or "Short Form", "SF36", "SF-36", SF12, "SF-12", "SF-6D", "SF6D", "EQ-5D", "Euroqol", "visual analogue scale", "health state preference", "BAROS", "obesity and weight loss quality of life", "OWLQOL". It was decided after the initial search was completed to exclude studies that used surveys measuring only one aspect of quality of life such as depression or anxiety, and studies that used BAROS as it was applicable to the surgical treatment group only following surgery. For the current review, HRQoL surveys were included that assessed both physical and mental domains. In the protocol, the intent was to combine the results from a variety of HRQoL tools, but the differences in scales and presentation of results made the standardization of the means not possible. This is why the SF-36 survey was chosen after completing the search as there were the largest number of papers using this tool that fit all the other selection criteria. For this reason standardized mean difference was not used for the results as it was only the one survey type used and the results were all measured on a uniform scale. The full search strategy for each electronic database has been described in Chapter 2. Major conference meeting notes were searched. These included the Canadian Obesity Summit, European Congress on Obesity and The Obesity Society (TOS) Annual Scientific Meeting. Reference lists of included papers were examined for additional studies missed in preliminary searches. Authors of papers were contacted for missing data, where possible. All references were managed and stored using RefWorks⁽²⁶⁾.

The final search was conducted on June 15th, 2015. The search yielded the following papers from online databases: PubMed (892), Cochrane Review (147), EmBase (762), CINAHL (168), PsycInfo (115). After duplicates were removed, 1376 references were identified from searching electronic and alternative sources. Upon reviewing titles and abstracts, 1355 were

removed for not having adequate follow-up, inadequate comparison groups, if they were commentaries/editorials, comparing different bariatric procedures, using an invalidated HRQoL tool, being published in a language other than English and for not meeting patient inclusion criteria. The remaining 21 articles were read in full text. Of these, 12 were excluded: due to the HRQoL tool used (i.e., not the SF-36) (n=3), non-surgical weight loss procedure (n=1), inadequate follow-up time (n=3), studies not published in English (n=3), no comparison group (n=2).

3.2.2 Risk of Bias (Quality) Assessment

One reviewer (SD) performed a preliminary search using the approved search strategy. After the preliminary search, one reviewer (SD) scanned study titles, abstracts, and keywords of every non-duplicate record retrieved from the literature search. Irrelevant titles were excluded and full-text papers were obtained where titles were deemed to be relevant or where eligibility was unclear. Two reviewers (SD and LT) independently assessed the remaining full text articles for eligibility criteria as well as risk of bias using the US Preventative Task Force quality rating criteria for assessing the quality of nonrandomized studies⁽²⁷⁾.

3.2.3 The US Preventative Task Force (USPSTF) Quality Criteria Rating

When assessing any type of scientific study, there must be both external and internal validity. External validity is the extent to which the evidence is relevant and generalizable to the population under study. Individual studies can increase their external validity by using randomized sampling methods and ensuring that the study population accurately reflects the larger population it is trying to reflect. Internal validity is the degree to which the study provides valid or accurate evidence for the population and setting in which it was conducted. The US Preventative Services Task Force (UPSTF) represents the efforts of the government and other

organizations committed to the development of clinical practice guidelines through a more evidence based approach. The USPSTF has developed an objective system for assessing the internal validity of various types of scientific studies⁽²⁷⁾. This system uses quality criteria checklists used to assess the quality of SRs, case-control studies, randomized control trials (RCT), cohort studies, and diagnostic accuracy studies. The current SR included cross-sectional studies, cohort studies and one prospective controlled trial. For this reason, the criteria checklists used for RCT's and cohort studies was the most appropriate to assess the internal validity of the studies. The checklist contains 8 categories each of which is given a poor, fair or good rating as well as an overall rating. The categories are described below:

- Initial assembly of comparable groups: For cohort studies: consideration of potential confounders with either restriction or measurement for adjustment in the analysis; consideration of inception cohorts
- Maintenance of comparable groups (includes attrition, crossovers, adherence, contamination)
- Important differential loss to follow-up or overall high loss to follow-up
- Measurements: equal, reliable, and valid (includes masking of outcome assessment)
- Clear definition of interventions
- All important outcomes considered
- Analysis: adjustment for potential confounders for cohort studies

Each reviewer reads the study and independently awards a ranking for each paper in each category. Studies are given a rating of good, fair or poor. Studies given a “*good*” rating meet the following criteria: comparable groups are assembled initially and maintained throughout the study (including follow-up at least 80 percent); reliable and valid measurement instruments are used and applied equally to the groups; interventions are spelled out clearly; important outcomes

are considered; and appropriate attention to confounders in analysis. Studies are graded a “*fair*” rating if any or all of the following problems occur, without the important limitations noted in the “*poor*” category below: comparable groups are assembled initially but some question remains whether some (although not major) differences occurred in follow-up; measurement instruments are acceptable (although not the best) and generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are accounted for. Studies will be graded as “*poor*” if any of the following major limitations exists: groups assembled initially are not close to being comparable or maintained throughout the study; unreliable or invalid measurement instruments are used or not applied at all equally among groups (including not masking outcome assessment); and key confounders are given little or no attention⁽²⁷⁾.

Only papers that receive an overall ranking of “*fair*” or “*good*” are included. Papers that receive an overall “*poor*” rating are excluded for further review. Using the quality criteria ratings helps to increase the likelihood that only papers with low systematic error are included which make the results of a SR and MA more accurate.

In the current study, the two reviewers (SD, LT) discussed any discrepancies and a consensus on quality was reached for each paper. The level of inter-rater agreement was 100%. Following identification of the included studies, each reviewer independently extracted relevant study data. Study location, surgery type, study population age, number of participants in both groups, BMI at baseline (where provided) and follow up, time to follow up, and quality of life survey used were extracted from each paper. Where feasible, missing data was obtained from authors through email correspondence.

3.2.4 Meta-analysis

All MAs were performed using Review Manager v. 5.3⁽²⁸⁾. Results were entered as mean domain score with standard deviation for the surgical and comparison group at time of long-term follow up. In the current study, HRQoL scores in patients ≥ 5 years post surgery were compared with baseline scores of a non-surgical control group (i.e., eligible patients waiting for bariatric surgery or those seeking treatment). This approach was based on our hypothesis that the intervention would result in significant improvements in HRQoL in the long-term compared with those who were not exposed to the intervention.

Three studies used multiple comparison groups. For the purpose of the MA, one group was selected for comparison. For Kolotkin et al.⁽¹⁸⁾ and Raoof et al.⁽²³⁾, the wait-listed populations were chosen as they were similar to the other comparison groups in the selected studies. In the case of Våge et al.⁽²¹⁾, the group with the smaller population was chosen as it had a small surgical group and the smaller comparison was less likely to skew the results than the second comparison group which had over a thousand individuals. To be analyzed in the MA the SF-36⁽²⁹⁾ results had to be entered in the same format of mean domain score and standard deviation. Authors were contacted by email if the data were published in a different format. Kolotkin et al.⁽¹⁸⁾, and Aftab et al.⁽²²⁾, responded with the data from their respective studies in the desired format. In other cases^(19,21), the results were converted from the format given into the desired format of mean and standard deviation. Våge et al.2003 provided the mean and 95% CIs converted to a standard deviation. Laurino Neto & Herbella⁽¹⁹⁾ provided the median score an IQR which were changed to mean and standard deviation using: $(\text{lower limit of IQR} + 2(\text{median}) + \text{upper limit of IQR})/4$ with a sample size ≥ 25 . When results were provided in a different format in a study they were either (a) converted using a similar approach as that used by Laurino Neto &

Herbella⁽¹⁹⁾ i.e., the conversion of IQR/median to mean with standard deviation, or (b) entered as the mean at follow up if available from the authors upon request.

3.2.5 Data Analysis for Meta-analysis

In the current MA, the random effects model was chosen to analyze the results. Under the random effects model, the assumption is that differences between studies are not all the same but that they do follow some sort of similar distribution. This model incorporates the lack of knowing why there are real or potential differences in the treatment effect by treating them all as random⁽³⁰⁾. The statistics used to measure heterogeneity were X^2 , I^2 and Tau^2 . The X^2 examines if observed result differences are congruent with chance alone. A large X^2 statistic relative to its degrees of freedom gives evidence of heterogeneity of intervention effects⁽³¹⁾. The I^2 describes the percentage of variability in effect estimates that is due to heterogeneity rather than chance⁽³¹⁾. The I^2 results can sometimes be misleading as it is a calculated value that depends on several factors. In general: 0% to 40%: might not be important; 30% to 60%: may represent moderate heterogeneity; 50% to 90%: may represent substantial heterogeneity; 75% to 100%: may represent considerable heterogeneity⁽³¹⁾. The Tau^2 is an estimate of between study variance for a random-effects model⁽³¹⁾. If $\text{Tau}^2 > 1$ it is likely there is substantial heterogeneity. This MA includes studies with small populations as well as non-randomized groups which can influence the heterogeneity. Since statistical heterogeneity may differ, for these reasons, it is also important to present results of each study within a forest plot in order to visually assess heterogeneity. This is so that both qualitative and quantitative heterogeneity can be assessed.

In the current study, results were analyzed as mean differences using a random effects model. Analysis was presented in forest plots. Heterogeneity was measured using Tau^2 , Chi^2 and the I^2 statistics as well as visual inspection of the forest plots. A p-value of <0.05 was considered

significant for the overall mean difference. All possible efforts were made to include all available published and unpublished studies identified in the search strategy. Authors used all resources at their disposal to search both published and unpublished results in an attempt to limit the effects of possible publication bias. Funnel plots were created but not included as there were too few studies and the plots could not be interpreted for bias. When funnel plots contain only a few data points, they may appear asymmetrical due to chance alone. With so few included studies, there may be publication bias but the funnel plots would likely show this regardless whether it was truly present.

3.3 Results

3.3.1 Systematic Review

The nine papers^(11, 16-24) included in the SR consisted of 7 cross-sectional studies, 1 prospective cohort study and 1 non-randomized clinical trial. The full study selection process is presented in the PRISMA flowchart (Figure 1) and the data extracted from the included papers are presented in Table 1. All papers included in the SR used a validated HRQoL assessment tool (e.g., the SF 36 or equivalent), but not all tools could be meaningfully combined for the MA. The standard mean difference could not be used for analysis as each survey used different questions and the scales were not comparable, even though they all measured quality of life. Only studies applying the generic SF-36 questionnaire were selected in order to give comprehensive information and allow for comparisons of study results for both physical and mental health domains. Generic HRQoL was assessed by the SF-36 health survey, which comprises 8 health

domains; physical function, role physical, bodily pain, general health, vitality, social function, role-emotion and mental health. All scale scores range from 0-100 with higher scores indicating higher health status.

3.3.2 The Short-Form 36

The SF-36 (The Medical Outcome Study 36-Item Short Form Health Survey) survey was developed as a result of the Medical Outcomes Study (MOS) as there was a need in the medical field for a tool that was easily administered, well documented and had good psychometric properties from a patient's point of view⁽²⁹⁾. The MOS was designed for two purposes. The first was to determine whether a patient's outcomes could be explained by differences in a clinician's specialty, technical and interpersonal styles and system of care and secondly to develop more practical tools for the more regimented monitoring of patient outcomes in medical practice⁽³²⁾. The MOS assesses 40 physical and mental health components. From this, 8 domains deemed the most important were chosen to be included in the SF-36. The SF-36 is considered a generic versus a specific tool as it evaluates health concepts that apply to basic human ideals that apply to everyone's well-being and functional status⁽³²⁾. These domains (e.g., physical and mental health domains) are both universally valued and are not age, disease or treatment specific. All of the domains for the SF-36 fit within one or more of the following 4 operational definitions: a) behavioural functioning, b) perceived well-being c) social and role disability d) personal evaluations (perceptions) of health in general⁽²⁹⁾. Specifically, the 8 included domains of the SF-36 are: 1) physical functioning, 2) role limitations due to physical health problems 3) bodily pain 4) general health 5) vitality (energy/fatigue) 6) social functioning 7) role limitations due to emotional problems 8) mental health (psychological distress and psychological well-being). The survey is scored such that a high score indicates optimal health for that domain (i.e., a high score

on a pain scale indicates freedom from pain). Once patients have completed the survey there is a 3-part process to determine the scores. First, 10 items need to be recoded to reflect the low-high scoring system. Secondly, a raw scale score is calculated for each domain. Finally, the scores are converted to a 0-100 scale⁽²⁹⁾. Since its creation, the SF-36 has been validated for a wide variety of chronic conditions, including obesity^(34,35). Both studies found that the SF-36 is suitable for use with an obese population, however there were some limitations noted. The physical and mental health summary scores had satisfactory construct validity; however the subscales (i.e., the individual domains) may require different aggregate groupings for proper interpretation for an obese population. The studies recommended the use of an obesity or weight specific HRQoL tool in conjunction with the SF-36 in order to assess HRQoL in a population living with obesity.

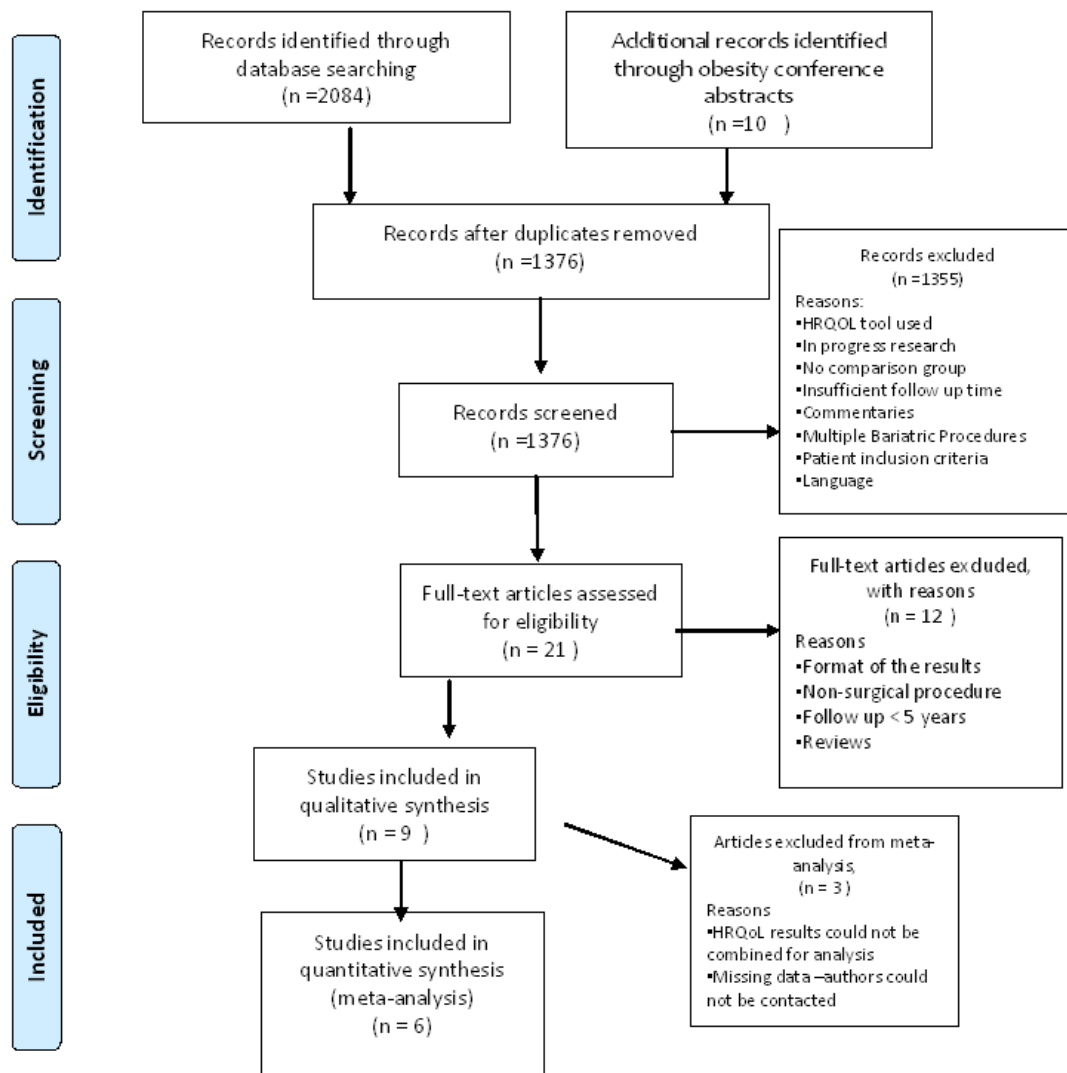


Figure 1 PRISMA flow diagram of selection process and included studies

There were three articles not included in the MA. An important article that analysed trends and effects of weight loss treatment on HRQoL in severely obese patients over a ten-year period was excluded due to the use of different tools (e.g., Current Health Scale, Mood Adjective Check list, the Hospital Anxiety and Depression Scale, Social Interaction category from the Sickness Impact Profile and the Obesity Problems Scale) that could not be combined with the SF-36⁽²⁹⁾. The Swedish Obesity Study (SOS) is a controlled, longitudinal trial of the health effects of weight loss in the severely obese. In a SOS study by Karlsson et al.⁽¹¹⁾ the authors presented data on HRQoL at 10 years follow-up post-surgery on 655 of 851 surgically treated patients. This study found that HRQoL was associated with the amount of weight lost. Peak improvements in the surgical group were observed in the first year of weight loss, whereas the weight regain phase (mainly between 1-6 year follow up) was accompanied by a gradual decline in HRQoL. The period for 6-10 follow-up was characterized by relatively stable observations in both weight and HRQoL. At ten years, net gains were noted in all HRQoL domains compared to baseline. Comparisons of treatment effects on HRQoL of the surgical versus conventional group after ten years showed significantly better outcomes in the surgery group on current health perception, social interaction, psychosocial functioning and depression, whereas no significant differences were found for overall mood and anxiety.

The study by Kiewiet et al.⁽¹⁷⁾ measured HRQoL using the Rand-36, a version of the SF-36⁽²⁹⁾ six years post-surgery. This paper could not be included as the results were presented as graphical data. The authors could not provide the numerical data. In this study, the physical functioning domain was the only domain that showed an improvement at follow-up for the surgical group versus the comparison group. All other domains (both physical and mental) showed no difference between groups. This study also included the Dutch population norms for

the Rand-36, and it is interesting to note that there were no significant differences between the surgical group and the population norms for 4 out of the 8 domains.

The study by Sanchez-Santos et al.⁽²⁰⁾ was also excluded. Although the paper focused on long term (>5 years) HRQoL post bariatric surgery and compared a surgical group with a control group, the study examined factors such as depression or stressful life events and how they influenced HRQoL. In addition, the tools used to measure HRQoL were not comparable. HRQoL in this study was assessed at five years post-surgery using the Bariatric Analysis of Reporting Outcome System (BAROS) which specifically examines outcomes in the bariatric surgery patient (i.e., not administered to the control group), the EQ-5D and a generic tool used to measure HRQoL.

3.3.3 Meta-analysis

The MA comprised 6^(16, 18, 19, 21-23) of these 9 studies. The risk of bias or “quality” of each study was assessed and is presented in Table 2. Of the 9 studies included in the SR, 5 were given a fair rating and 4 received a good rating. No papers were deemed poor by the reviewers and there were no discrepancies between the two reviewers.

Table 1 Included studies for review and MA

Study	Assessed Quality	Design	Baseline characteristics of treatment group N=population at time of follow up	Surgery Type	Control Group	Mean follow up time post-surgery	HRQOL Tool
Raof et al., 2015⁽²³⁾ Örboro, Uppsala, Sweden	Good	Cross-sectional	N=485 Baseline: BMI: 43.5 ± 6.7 Age: not given % Female: 84% Controls were age and sex matched. Average weight loss at follow-up	Roux-en-Y Gastric Bypass	N= 1590 (Control Group 1) General Swedish population N=972(Control Group 2) Scandinavian Surgery Registry, morbidly obese patients accepted to bariatric surgery	11.5 years	SF-36 Obesity Related Problems Scale
Aftab et al., 2014^{(22)*} Oslo, Norway	Good	Cross-sectional	N=177 Baseline: BMI: 46±5 kg/m ² Age: 38±9 years %Female:75% Average %weight loss at follow-up 27.0%± 11.0%	Laparoscopic Gastric Bypass	N=288 Patients on wait list at same clinic at time of follow up	63±5 months	SF-36 Obesity Related Problems Scale

DeZwaan et al. 2002^{(16)*} North Dakota, USA	Fair	Cross-sectional	N=78 Baseline BMI: 43.8±5.9 kg/m ² Age: 40.3 (17-61 years) %Female:83% Average BMI at follow-up 32.8 kg/m ² (22.7-49.5)	Roux-en-Y Gastric Bypass	N= 110 Pre-operative obese patients Baseline: BMI: 48.4±8.3 kg/m ² Age: 39.6(19-62 years) %Female: 87%	13.8 years	SF-36
Karlsson et al. 2007⁽¹¹⁾ Sweden	Good	Prospective controlled clinical trial	N= 655 Baseline: BMI: 41.9±4.2 kg/m ² Age: 47.0± 5.7 Average BMI at follow-up 35.3 kg/m ² ± 5.3	Gastric banding, gastric bypass, vertical gastrectomy	N= 621 conventional treatment Baseline: BMI:39.9 (4.6) kg/m ² Age 48.4±6.7 years	10 years	Current Health scale, Mood Adjective Check List, Hospital Anxiety and Depression scale, Sickness Impact Profile, and Obesity Related Problems Scale

<p>Kiewiet et al. 2008⁽¹⁷⁾</p> <p>The Netherlands</p>	<p>Fair</p>	<p>Cross-sectional</p>	<p>N=59</p> <p>Baseline BMI: 44.9±5.9 kg/m² Age: 42.4±9.7 %Female: 93.0%</p> <p>Average BMI at follow-up 33.3 kg/m²± 6.0</p>	<p>Adjustable Gastric Banding</p>	<p>N=28 Pre-operative patients</p> <p>Baseline BMI:41.8 ± 3.4 kg/m² Age: 39.8±8.5 years %Female 86.0%</p> <p>Average BMI at follow-up 41.8 ± 3.4 kg/m²</p>	<p>74.7 months</p>	<p>RAND 36-Item Health Survey</p>
<p>Kolotkin et al. 2012^{(18)*}</p> <p>Utah, United States</p>	<p>Good</p>	<p>Prospective cohort</p>	<p>N= 323</p> <p>Baseline: BMI: 47.4±7.7 kg/m² Age:43.4 10.7 %Female: 84%</p> <p>Average %EWL at follow-up 56.4% ±21.4%</p>	<p>Gastric Bypass</p>	<p>N1=257(Control Group 1) Obese patients on wait list **</p> <p>BMI: 45.9±(7.9kg/m² Age: 44.7 10.9 %Female 84%</p> <p>Average %EWL at follow-up 0.3% ±22.2%</p> <p>N2=272(Control Group 2) Obese population</p> <p>BMI: 43.6±(6.4</p>	<p>6 years</p>	<p>SF-36 IWQOL-Lite</p>

					kg/m ² Age 49.7± 10.5 years % Female 77% Average %EWL at follow-up 0.2% ±23.3%		
Laurino Neto & Herbella. 2013^{(19)*} Brazil	Fair	Cross sectional	N= 50 Baseline BMI:51.± 6.3 kg/m ² Age: 42.0 ±10.8 %Female: 88% Average BMI at follow-up 36.0 ±6.3 kg/m ²	Roux-en y Gastric Bypass	N=50 Obese patients on waitlist Baseline: BMI: 52.0±8.0 kg/m ² Age 42.0±11.4 years %Female 80%	112 months	SF-36
Sanchez- Santos et al. 2006⁽²⁰⁾ Spain	Fair	Cross- sectional	N=50 Baseline BMI: 50.5±9.0kg/ m ² Age: 40.5±9.0 %Female: 90.2%	Roux-en- Y Gastric Bypass	N= 78 Non-operated morbidly obese patients Baseline: BMI: 47.8±7.0 kg/m ² Age: 46 0±.8.8 years %Female: 9%	>5 years	EuroQOL-5D BAROS

<p>Våge et al. 2003^{(21)*}</p> <p>Norway</p>	<p>Fair</p>	<p>Cross-sectional</p>	<p>N=20</p> <p>Baseline: BMI: 41.0 (33-64 kg/m²) Age:31 (20-51 years) %Female:100%</p> <p>Average BMI at follow-up BMI: 33 kg/m² (23-75 kg/m²)</p>	<p>Jejunioileal Bypass</p>	<p>N1= 1 118 (Control Group 1) Obese patients waitlisted for surgery</p> <p>BMI: 46.6 ± 8.3 kg/m² Age : 38.3± 9.8 years</p> <p>N2= 80 (Control Group 2) Obese patients seeking treatment</p> <p>BMI: 48.7 ± 8.1kg/m² Age 38.0 (16-70 years)</p>	<p>25 years</p>	<p>SF-36</p>
--	-------------	------------------------	---	----------------------------	---	-----------------	--------------

Table 2 Risk of bias assessment*

Study	Assembly of comparable groups	Maintenance of comparable groups	No important differential loss to follow-up or overall high loss to follow up	Measurements: equal, reliable, valid (includes masking of outcome assessment)	Clear definition of interventions	All important outcomes considered	Analysis: adjustment for potential confounders	Overall assessed quality
Raof et al. 2015⁽²³⁾	Good	Good	Fair	Good	Good	Good	Good	Good
Aftab et al. 2014⁽²²⁾	Fair	Fair	Good	Good	Good	Good	Fair	Good
DeZwaan et al. 2002⁽¹⁶⁾	Good	Poor	Fair	Good	Good	Good	Fair	Fair
Karlsson et al. 2007⁽¹¹⁾	Good	Good	Good	Fair	Good	Good	Good	Good
Kiewiet et al. 2008⁽¹⁷⁾	Poor	Good	Fair	Good	Good	Good	Fair	Fair
Kolotkin et al. 2012⁽¹⁸⁾	Good	Good	Good	Good	Good	Good	Good	Good
Laurino Neto & Herbella 2013⁽¹⁹⁾	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair
Sanchez-Santos et al. 2006⁽²⁰⁾	Fair	Fair	Fair	Good	Good	Poor	Fair	Fair
Våge et al. 2003⁽²¹⁾	Poor	Fair	Fair	Good	Good	Fair	Fair	Fair

*Risk of bias or ‘quality’ of each study was assessed using the US Preventative Task Force quality rating criteria for assessing the quality of nonrandomized studies⁽²⁷⁾.

All six studies report some improvements in HRQoL for the surgical patients in the long-term compared to the obese comparison groups. The physical health scores and total scores showed improvements in all studies. Mental health scores had more varied results. Aftab et al.⁽²²⁾ study participants showed a significant improvement in all mental health domains of the SF-36 after five years of follow up compared to a population of patients wait-listed for bariatric surgery. DeZwaan et al.⁽¹⁶⁾ found no difference in the mental health domain but there was a significant difference favouring the surgical group for all other domains. Kolotkin et al.⁽¹⁸⁾ found significant improvements for the surgical group in all domains compared to the control groups with the exception of role emotional. Laurino Neto & Herbella⁽¹⁹⁾ found a significant improvement in only the vitality score for the surgical group in the long term. Raoof et al.⁽²³⁾ showed improvements in the vitality domain for surgical patients over the waitlisted comparison group. Våge et al. 2003 found lower social functioning and role emotional scores for their surgical group.

3.3.4 Synthesis of Results

The results of the MA can be seen in Figures 2a-d (four domains of physical health) and Figures 3a-d (four domains of mental health) as forest plots of the domain scores from the SF-36 survey⁽²⁹⁾. The z -score generated for each forest plot shows the overall effect of surgery on the domain of HRQoL. As the meta analysis looks at the mean difference, the positive value of all the z-scores demonstrates that the surgical group showed an improvement in HRQoL when compared to the non-surgical control groups, where the p value for the associated z score is ≤ 0.05 are considered statistically significant. The z-scores and therefore the overall effect of surgery on the HRQoL domains show significant improvement for all groups with the exception of role-physical which has a p value = 0.1. The forest plots favour ($p \leq 0.05$) the surgical group

for 3 out of 4 physical domain scores and 4 out of 4 mental domain scores. Mean difference and 95% CI for the physical function, bodily pain and general health were 31.29 (21.37, 41.21), 12.56 (6.51, 18.61), and 17.54 (6.60, 28.48), respectively. For the mental health scores of vitality, social functioning, role emotional and mental health, total scores (mean difference (95%CI)) were 18.18(6.50, 23.85), 14.66 (3.34, 25.97), 12.58 (4.06, 21.10) and 6.13 (0.17, 12.09). Role physical 12.15 (-1.73, 26.02) was not significantly different between the two groups. These measures exhibited a fairly high level of heterogeneity ($I^2 \geq 79\%$) for all domains. The statistical tests of heterogeneity show a significant degree of heterogeneity of the results from the various studies. Examination of the forest plots show that this is quantitative as opposed to qualitative heterogeneity. Despite the statistical heterogeneity, the results are similar from study to study. The similarity of the results from study to study somewhat mitigates the lack of randomization and proper control groups in these studies.

Figure 2a:

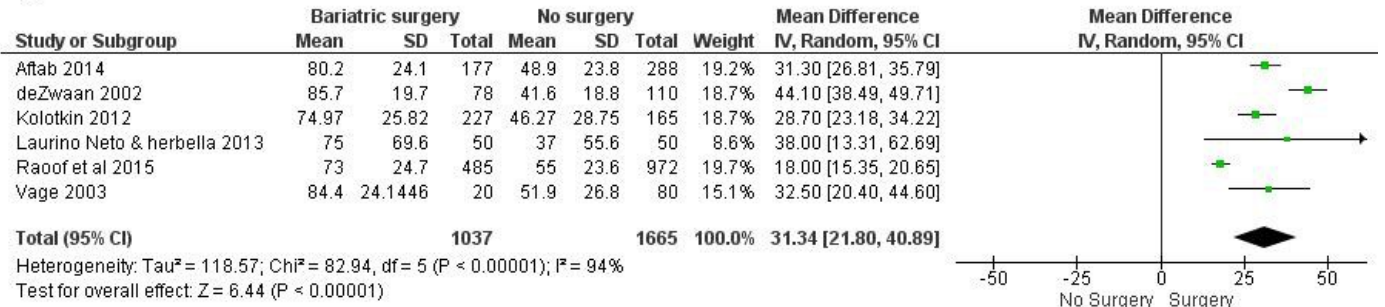


Figure 2b:

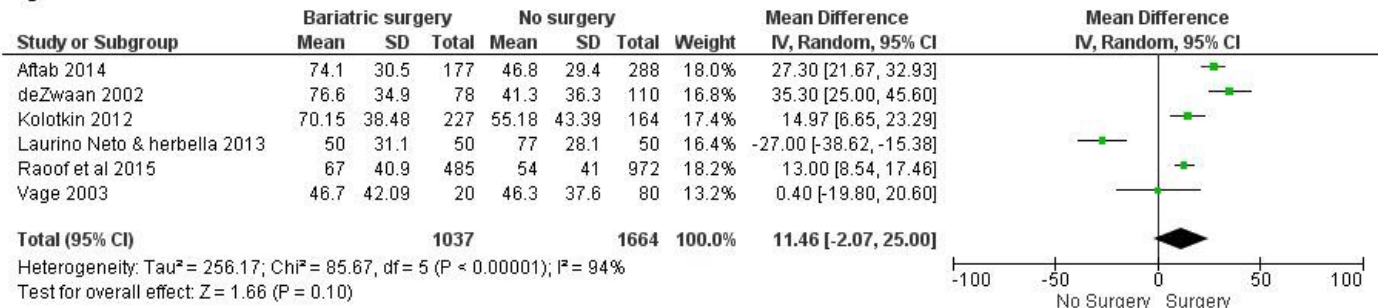


Figure 2c:

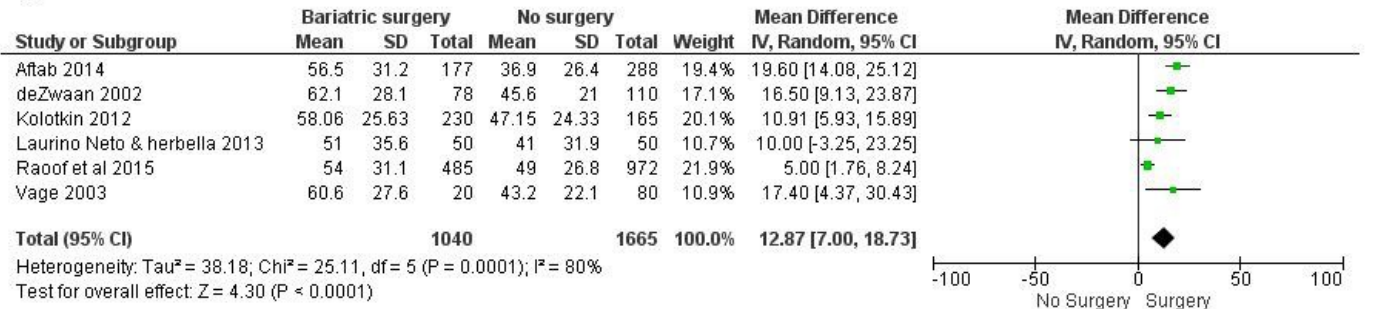


Figure 2d:

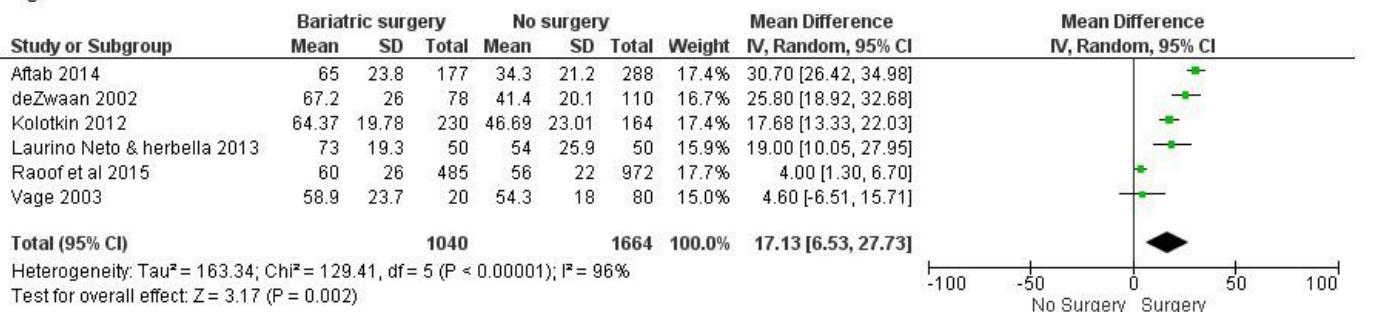


Figure 2 Forest plots for SF-36 physical domains. Each forest plot represents the non-surgical population groups with obesity on the left side and the surgical groups at time of long-term follow-up on the right side. The values illustrated in each forest plot are mean differences in domain scores between surgery and no surgery. (a) SF-36 physical function domain, (b) SF-36 role physical domain, (c) SF-36 bodily pain domain, (d) SF-36 general health domain.

Figure 3a:

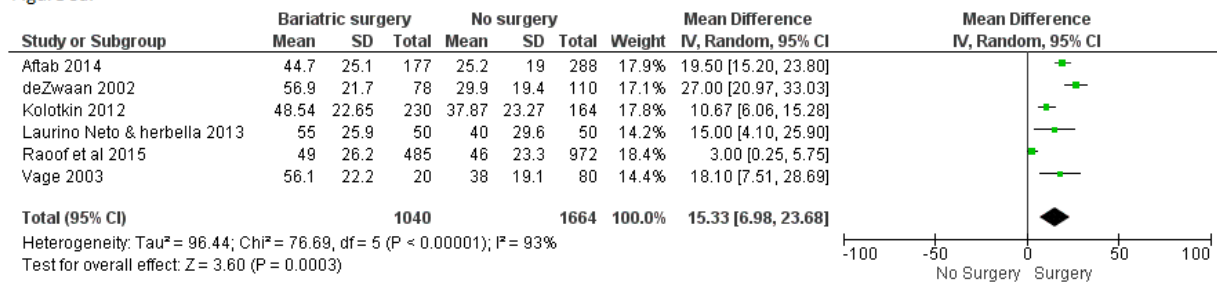


Figure 3b:

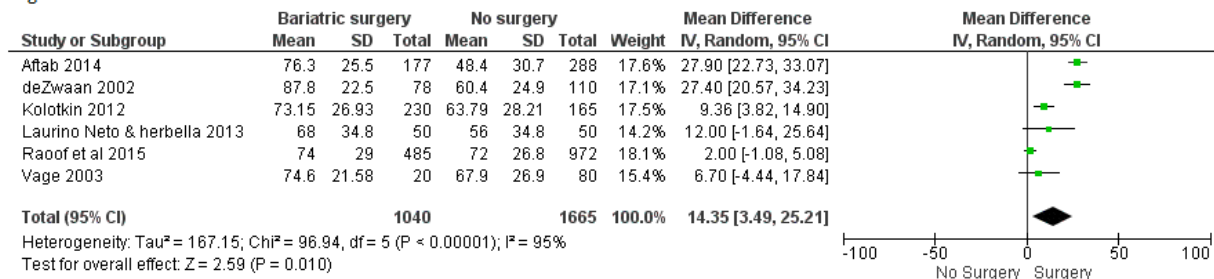


Figure 3c:

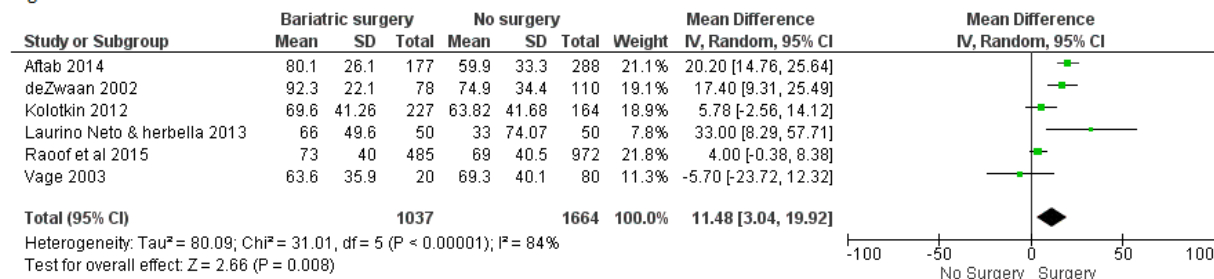


Figure 3d:

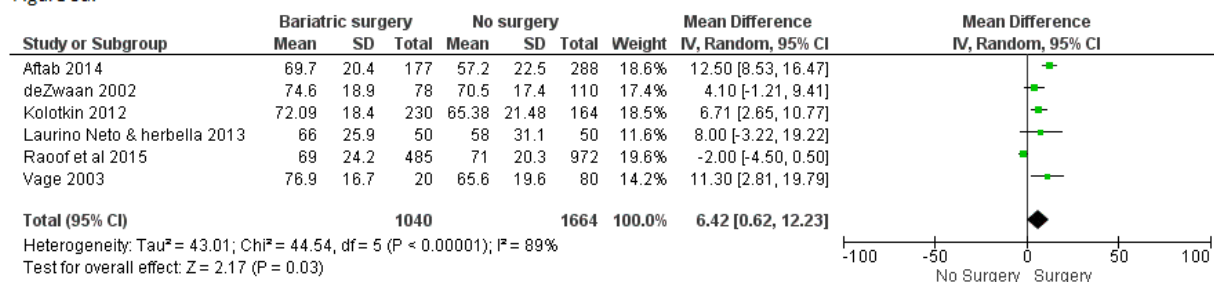


Figure 3 Forest plots for SF-36 mental domains. Each forest plot represents the non-surgical population groups with obesity on the left side and the surgical groups at time of long-term follow-up on the right side. . The values illustrated in each forest plot are mean differences in domain scores between surgery and no surgery. (a) SF-36 vitality domain, (b) SF-36 social functioning domain, (c) SF-36 role emotional domain, (d) SF-36 mental health domain.

3.4 Discussion

Overall greater improvements in HRQoL (e.g., physical and mental health domains) were reported in the SR after five years for the surgical group compared to the control group (i.e., waitlist bariatric surgery patients, population-based individuals not seeking bariatric surgery). However, there were inconsistencies in results (i.e., favouring surgical group, favouring control group, and no difference) in the individual physical and mental health domains.

The MA provides evidence that bariatric surgery significantly improves physical HRQoL scores, with the exception of only one domain, role physical. The studies by Laurino Neto & Herbella⁽¹⁹⁾ and Våge et al.⁽²¹⁾ favour no surgery and no effect, respectively, as shown in the MA. Laurino Neto & Herbella⁽¹⁹⁾ suggest that a decrease in role physical score is related to a lack of motivation as age increases in the surgical group. Våge et al.⁽²¹⁾ do not address the role physical scores between control and surgical groups, but have a 25 year follow up. It is possible that the findings are also related to the age of the surgical group as suggested by Laurino Neto & Herbella⁽¹⁹⁾.

Contrary to what is reported in the SR, the current MA provides evidence for a significant improvement in *all* mental health domains after five years favouring the surgical group compared with the controls. This MA documents improvement in mental health at least five years after bariatric surgery; a finding that has been inconsistently reported in the general literature of HRQoL in individuals undergoing bariatric surgery for the treatment of severe obesity^(17, 19, 20,22).

The magnitude of improvement in the surgical groups compared to the control groups was smaller for the mental health domains compared to the physical health domains, when examining the mean differences, but they are all statistically significant with the exception of

role physical. This MA demonstrates that long-term HRQoL significantly improves among bariatric surgery patients compared to an obese general population or those waiting for bariatric surgery. The similarity of the magnitude and direction of the results in each of the domains from study to study speaks very strongly to the validity of these results.

3.5 Chapter 3 References

1. Colquitt JL, Picot J, Loveman E, Clegg AJ. Surgery for obesity. *Cochrane Database Syst Rev* 2009;**15**(2):CD003641. doi:[10.1002/14651858.CD003641.pub3](https://doi.org/10.1002/14651858.CD003641.pub3).
2. World Health Organization. *Obesity: Preventing and Managing the Global Epidemic*. World Health Organization. Report number series:894, 2000.
3. Karmali S, Johnson Stoklossa C, Sharma A, Christiansen S, Cottreau D, Birch DW. Bariatric surgery: a primer. *Can Fam Phys* 2010;**56**:873-879.
4. Guyatt G, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993;**118**:622-629.
5. Dymek MP, LeGrange D, Neven K, Alverdy J. Quality of life and psychosocial adjustment in patients after Roux-en-Y gastric bypass: a brief report. *Obes Surg* 2001;**11**:32-39.
6. van Gemert WG, Adang EMM, Greve JWM, Soeters PB. Quality of life assessment of morbidly obese patients: effect of weight-reducing surgery. *Am J Clin Nutr* 1998;**67**:197-201.
7. Karlsson J, Sjöström L, Sullivan M. Swedish obese subjects (SOS): an intervention study of obesity: two-year follow-up of health-related quality of life (HRQL) and eating behavior after gastric surgery for severe obesity. *Int J Obes Relat Metab Disord* 1998;**22**:113-126.
8. Nguyen NT, Goldman C, Rosenquist CJ, et al. Laparoscopic versus open gastric bypass: a randomized study of outcomes, quality of life, and costs. *Ann Surg* 2001;**234**:279-289.
9. O'Brien PE, Dixon JB, Brown W, et al. The laparoscopic adjustable gastric band (Lap-Band): a prospective study of medium-term effects on weight, health and quality of life. *Obes Surg* 2002;**12**:652-660.
10. Magro DO, Geloneze B, Delfini R, Pareja BC, Callejas F, Pareja JC. Long-term weight regain after gastric bypass: a 5-year prospective study. *Obes Surg* 2008;**18**:648-651.
11. Karlsson J, Taft C, Ryden A, Sjoström 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**:1248-1261.
12. Nickel MK, Loew TH, Bachler E. Change in mental symptoms in extreme obesity patients after gastric banding, part II: six-year follow-up. *Int J Psychiatry Med* 2007;**37**:69-79.

13. 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**:1585-1591.
14. Suter M, Donadini A, Romy S, Demartines N, Giusti V. Laparoscopic Roux-en-Y gastric bypass: significant long-term weight loss, improvement of obesity-related comorbidities and quality of life. *Ann Surg* 2011;**254**:267-273.
15. Aasprang A, Andersen JR, Våge V, Kolotkin RL, Natvig GK. Five-year changes in health-related quality of life after biliopancreatic diversion with duodenal switch. *Obes Surg* 2013;**23**:1662-1668.
16. deZwaan 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**:773-780.
17. Kiewiet RM, Durian MF, Cuijpers LPLH, Hesp FLEM, van Vliet ACM. Quality of life after gastric banding in morbidly obese Dutch patients: long-term follow-up. *Obes Res Clin Pract* 2008;**2**:151-158.
18. 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;**8**:625-633.
19. Laurino Neto RM, Herbella FAM. Changes in quality of life after short-term and long-term follow-up of roux-en-Y gastric bypass for morbid obesity. *Arg Gastroenterol* 2013;**50**(3):186-190.
20. Sanchez-Santos R, Delarrio MJ, Gonzalez C, et al. Long-term health-related quality of life following gastric bypass: influence of depression. *Obes Surg* 2006;**16**:580-585.
21. Våge V, Solhaug JH, Viste A, et al. Anxiety, depression and health-related quality of life after jejunioileal bypass: a 25-year follow-up study of 20 female patients. *Obes Surg* 2003;**3**:706-713.
22. Aftab H, Risstad H, Sovik TT, et al. Five-year outcome after gastric bypass for morbid obesity in a Norwegian cohort. *Surg Obes Relat Dis*. 2014;**10**:71-78.
23. Raoof M, Näslund I, Rask E, et al. Health-related quality-of-life (HRQoL) on average of 12 years after gastric bypass surgery. *Obes Surg* 2015;**25**:1119-1127.
24. Andersen JR, Aasprang A, Karlsen TI, Vage V, Kolotkin RL. Health related quality of life after bariatric surgery: a systematic review of prospective long-term studies. *Surg Obes Relat Dis* 2014;**2**:466-473.

25. 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**:169-182.
26. *COS Online Research Management Tool*. Available from: <https://www.refworks.com/>. RefWorks: Bethesda, MD, 2009. [Accessed October 10th 2014]
27. Russell P. Harris, Mark Helfand, Steven H. Woolf, Kathleen N. Lohr, Cynthia D. Mulrow, Steven M. Teutsch, David Atkins. for the Methods Work Group, Third U.S. Preventive Services Task Force. Current Methods of the U.S. Preventive Services Task Force A Review of the Process . *Am J Prev Med*. 2001; 20(3S)
28. RevMan. *The Cochrane Collaboration's Review Manager (Vers.5.1)*. Oxford: The Cochrane Collaboration;2011.
29. Ware J, Snow K, Kosinski M, Gandek B. *SF-36 Health Survey: Manual and Interpretation Guide*. Boston: Health Institute, New England Medical Center; 1993
30. Kontopantelis, E.; Reeves, D. "Performance of statistical methods for meta-analysis when true study effects are non-normally distributed: A simulation study.". *Statistical Methods in Medical Research*. 2012; 21(4): 409–26.[doi:10.1177/0962280210392008](https://doi.org/10.1177/0962280210392008). [PMID 21148194](https://pubmed.ncbi.nlm.nih.gov/21148194/). [Accessed November 12th 2014]
31. Higgins, J., Green, S. (ed.) *Cochrane Handbook for Systematic review of interventions*. Version 5.1.0 Updated March 2011. (chapter 9 &10)
32. Tarlov, A.R, Ware, J.E., Greenfield, S. Nelson E.C, Perrin, E. & Zubkoff M. The Medical Outcomes Study. An application of methods for monitoring the results of medical care. *Jama*. 1989; 262 (7) 925-30
33. Ware, J.E. Standards for validating health measures. Definition and Content. *Journal of chronic diseases*. 1987; 40 473-480
34. Karlsen T.I., Tveita, E.K, Natvig, G.K, Tonstad, S. & Hjelmessaeth, J. Validity of the SF-36 in patients with morbid obesity. *Obes Fact*. 2011
35. Corica, F., Corsonello, A., Apolone, . Lucchetti, M. Melchionda, N., Marchesini, G. Construct validity of the short form-36 health survey and its relationship with BMI in obese outpatients. *Obesity*. 2006; 14(8): 1429-37

Chapter 4 Summary

4.1 Overview

The results of this SR and MA show an improvement in both mental and physical scores for bariatric patients in the long term. This chapter will discuss the strengths and limitations of the SR/MA, the implications of the study findings, future research and overall study conclusions.

4.2 Strengths and Limitations

There are both strengths and limitations of the current study. The limitations include only using observational studies as there were no available RCT's, the limited number of available studies, publication bias, how the results were reported, the study populations used, and the wide range of HRQoL tools.

A limitation of this review was that all included papers were observational studies. It is difficult to determine whether improvements or declines in quality of life were due to the intervention (bariatric surgery) or other confounders. The surgery seeking obese population tends to be worse off in terms of health compared with other population obese groups^(1,2). This would decrease the bias seen between the surgical and comparison groups in the studies for the MA, but may show differences between studies such as Karlsson et al.⁽³⁾ who use obese groups seeking conventional treatment. This limits the ability to directly compare the MA studies with the Karlsson et al.⁽³⁾, study in the SR.

There were a small number of comparable studies to use in the MA. This limits the usefulness of publication bias tests. The literature search produced a large number of studies overall, but many did not use a HRQoL tool that could be used in the MA. The various HRQoL surveys used across studies measured different aspects of quality of life, making it difficult to use

the standard mean difference to compare different surveys. The SF-36 was the most frequently used measure in our review. However, as a generic measure of HRQoL, it may not be the most sensitive questionnaire to examine HRQoL in bariatric surgery patients⁽⁴⁾.

Publication bias exists when the findings of a study influence whether it is published or not. In research, negative or “non” results often go unpublished but may include important findings. Publication bias can occur in cases like this when only the positive outcomes are reported which can lead to an information bias. It may also occur in reverse when positive results are not published for a certain reason⁽⁵⁾. In a MA, funnel plots are a graphical way of assessing publication bias. A funnel plot is a scatter plot showing the intervention effect estimates from each study against a measure of each study’s size or precision. The horizontal axis commonly represents the effect estimates and the vertical axis represents the study size on the vertical axis. Without bias present, the plot should resemble an inverted “funnel” with smaller studies evenly distributed across the horizontal (effect) axis but near the bottom of the vertical (study size) axis. Larger studies will be clustered closer to the center of the horizontal axis but higher up on the vertical axis. Visual assessment of funnel plots can be subjective and some argue minimally useful⁽⁵⁾. It is also hard to assess when there are a limited number of studies as the funnel shape will not be visually apparent with only a few data points on the plot. In this MA, a funnel plot was generated but not included as there were only 9 studies and the plot was not useful in showing any kind of trend in publication distribution. The data points were too sparse to show either a definite symmetrical or asymmetrical shape limiting interpretation of results in terms of publication bias. It is possible that this study suffers from publication bias which is a limitation of the study and the results.

The study results were presented in a variety of formats. In most cases authors were able to be reached and appropriate data obtained directly from the original research team. Laurino Neto & Herbella⁽⁶⁾ reported their results using median and IQR and the results which seemed a questionable choice on the part of the authors based on the recommended form of reporting suggested by the SF-36 instructions⁽⁷⁾. The results were converted to mean and standard deviation using an excel conversion file, but their original data was presented in a format different from the usual format for SF-36 results, which could impact the results of the MA. The results shifted to favouring surgery when the Laurino Neto & Herbella⁽⁶⁾ study was removed for the role physical domain in a sensitivity analysis.

The use of historical subjects versus contemporaneous comparison groups may also cause a bias in the results. The wait listed comparison groups were not followed prospectively as were the surgical groups. This may cause differences in baseline results between the groups as well as impact group scores on the HRQoL surveys. The comparison groups for the included studies did not come from randomized control trials therefore there is the possibility that the control groups were inherently different at baseline than the groups seeking surgery. Should this be the case, there may be a bias in the comparison of the 2 groups.

The limited number of studies using the same standardized and validated HRQoL tools hindered the degree to which studies could be meaningfully compared. There is no official “gold standard” for HRQoL measurements^(8,9). Of the 9 studies that met inclusion criteria for the SR, 9 different HRQoL tools were used. The majority of these tools were used in a single study, the SOS, which is the largest and most comprehensive of the review studies, using multiple generic and obesity specific tools⁽³⁾. The studies that helped validate the SF-36 for use with obese populations^(10,11) make the observation that a weight specific and generic tool combined would

give the most accurate view of HRQoL. On closer examination of the PRISMA flowchart, there are a number of studies eliminated due to the type of HRQoL tool. This generally meant that a formal tool was not used, or that the tool was not validated. There may be studies that have comparable study populations and study design but because of the tool selected the HRQoL results cannot be used in comparisons. The SF-36 has been validated for use in a population with obesity, but there are some limitations that have been raised in other studies^(10,11) mainly surrounding the discrepancies amongst the subscales versus the component scores. The majority of papers did not calculate the physical or MCSs and therefore they were not compared in the MA. These scores may provide additional insights in to the HRQoL of patients.

The strengths include the thorough nature of the literature review as well as the homogeneity in the direction of the results for the forest plots in the meta-analyses. The strength of the MA comes from the large qualitative homogeneity of the results across the forest plots. The statistical heterogeneity was fairly high in all studies, but through visual assessment of the forest plots, the results were incredibly similar and had the same directionality, despite having a large range in population size and time to follow up.

4.3 Implications of the Current Study Findings

These results are important for patients, health professionals and decision makers. For patients, they provide comfort that in the long term, bariatric surgery has a positive impact on general HRQoL as it relates to physical and mental health domains. For health professionals the study provides evidence that an intervention to treat severe obesity that results in short term improvements in HRQoL is sustained in the long term. For decision makers, given the increases

in the prevalence of severe obesity and the provision of bariatric surgery as a treatment, these findings help to inform evidence-based care in the area of bariatric care.

4.4 Future Research

The importance and relevance of future research on the long-term HRQoL of bariatric surgery patients will increase as the number of bariatric surgeries continues to increase.

Standardization of a preferred reporting method for HRQoL for both obese and bariatric patients is essential for future research. Without a standard method of HRQoL reporting, the presentation of results can be varied and make it difficult to compare across studies.

It is also important to study the relation between the primary and other secondary outcomes of bariatric surgery with HRQoL in the long term. There is a trend toward at least a partial regain of weight in the long term for patients' post-bariatric surgery. It is important to understand the correlation of comorbid conditions and weight gain as it pertains to a maintenance in improvement in quality of life.

4.5 Conclusions

The main aim of this thesis was two-fold: first to conduct a SR in order to assess the quality of evidence and effectiveness of bariatric surgery on HRQoL ≥ 5 years in patients ≥ 18 years compared to non-surgical control groups and second to conduct a MA of appropriate studies. The SR found inconsistent results for long-term improvements in physical and mental health. In contrast, the MA found significant improvements in these domains ≥ 5 years after bariatric surgery. These study findings provide evidence for a substantial and significant improvement in physical and mental health favoring the surgical group compared with controls

spanning 5 to 25 years after surgery, an important finding for patients, clinicians and decision-makers.

4.6 Chapter 4 References

1. Fitzgibbon, M.L, Stolley, M.R., Kirschenbaum, D.S. Obese people who seek treatment have different characteristics than those who do not seek treatment. *Health Psychology*, 1993; 12(5) 342-345. <http://dx.doi.org/10.1037/0278-6133.12.5.342> [Accessed November 20th 2016]
2. Kral JG, Sjostrom LV, Sullivan MBE. Assessment of quality of life before and after surgery for severe obesity. *Am J Clin Nutr* 1992; **55**: 611S–614S.
3. Karlsson, J., Taft, C., Rydén, 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 (Lond)* 2007; 31(8):1248-1261
4. Coulman KD, Abdelraham 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;**14**:707-720.
5. Higgins, J., Green, S. (editors) Cochrane Handbook for Systematic review of interventions. Version 5.1.0 Updated March 2011; (chapter 10)
6. Laurino Neto RM, Herbella FAM. Changes in quality of life after short-term and long-term follow-up of roux-en-Y gastric bypass for morbid obesity. *Arg Gastroenterol* 2012; **50**(3):186-190.
7. Ware J, Snow K, Kosinski M, Gandek B. *SF-36 Health Survey: Manual and Interpretation Guide*. Boston: Health Institute, New Eng-land Medical Center; 1993
8. Guyatt, G. Insights and limitations from health related quality of life research. *J Gen Intern Med*. 1997; 12(11): 720-721
9. Pickard, A.S, Yang, Y., Lee, T.A. Comparison of health related quality of life measures in chronic obstructive pulmonary disease. *Health Qual Life Outcomes*.2011; 9:26.

published online 2011 April 18 doi: 10.1186/1477-7525-9-26 [Accessed November 20th 2016]

10. Karlsen T.I., Tveita, E.K, Natvig, G.K, Tonstad, S., Hjelmesaeth, J. Validity of the SF-36 in patients with morbid obesity. *Obes Fac.* 2011; 4(5):346-351
11. Corica, F., Corsonello, A., Apolone, . Lucchetti, M. Melchionda, N., Marchesini, G. (2006) Construct validity of the short form-36 health survey and its relationship with BMI in obese outpatients. *Obesity (silver spring)* 14(8): 1429-37

Appendix A
The Short Form (SF)-36

SF-36 QUESTIONNAIRE

Name: _____ **Ref. Dr:** _____ **Date:** _____

ID#: _____ **Age:** _____ **Gender: M / F**

Please answer the 36 questions of the **Health Survey** completely, honestly, and without interruptions.

GENERAL HEALTH:

In general, would you say your health is:

Excellent Very Good Good Fair Poor

Compared to one year ago, how would you rate your health in general now?

Much better now than one year ago

Somewhat better now than one year ago

About the same

Somewhat worse now than one year ago

Much worse than one year ago

LIMITATIONS OF ACTIVITIES:

The following items are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much?

Vigorous activities, such as running, lifting heavy objects, participating in strenuous sports.

Yes, Limited a lot Yes, Limited a Little No, Not Limited at all

Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf

Yes, Limited a Lot Yes, Limited a Little No, Not Limited at all

Lifting or carrying groceries

Yes, Limited a Lot Yes, Limited a Little No, Not Limited at all

Climbing several flights of stairs

Yes, Limited a Lot Yes, Limited a Little No, Not Limited at all

Climbing one flight of stairs

Yes, Limited a Lot Yes, Limited a Little No, Not Limited at all

Bending, kneeling, or stooping

Yes, Limited a Lot Yes, Limited a Little No, Not Limited at all

Walking more than a mile

Yes, Limited a Lot Yes, Limited a Little No, Not Limited at all

Walking several blocks

Yes, Limited a Lot Yes, Limited a Little No, Not Limited at all

Walking one block

Yes, Limited a Lot Yes, Limited a Little No, Not Limited at all

Bathing or dressing yourself

Yes, Limited a Lot Yes, Limited a Little No, Not Limited at all

PHYSICAL HEALTH PROBLEMS:

During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of your physical health?

Cut down the amount of time you spent on work or other activities

Yes No

Accomplished less than you would like

Yes No

Were limited in the kind of work or other activities

Yes No

Had difficulty performing the work or other activities (for example, it took extra effort)

Yes No

EMOTIONAL HEALTH PROBLEMS:

During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?

Cut down the amount of time you spent on work or other activities

Yes No

Accomplished less than you would like

Yes No

Didn't do work or other activities as carefully as usual

Yes No

SOCIAL ACTIVITIES:

Emotional problems interfered with your normal social activities with family, friends, neighbors, or groups?

Not at all Slightly Moderately Severe Very Severe

PAIN:

How much bodily pain have you had during the past 4 weeks?

None Very Mild Mild Moderate Severe Very Severe

During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)?

Not at all A little bit Moderately Quite a bit Extremely

ENERGY AND EMOTIONS:

These questions are about how you feel and how things have been with you during the last 4 weeks. For each

question, please give the answer that comes closest to the way you have been feeling.

Did you feel full of pep?

All of the time

Most of the time

A good Bit of the Time

Some of the time

A little bit of the time

None of the Time

Have you been a very nervous person?

All of the time

Most of the time

A good Bit of the Time

Some of the time

A little bit of the time

None of the Time

Have you felt so down in the dumps that nothing could cheer you up?

All of the time

Most of the time

A good Bit of the Time

Some of the time

A little bit of the time

None of the Time

Have you felt calm and peaceful?

All of the time

Most of the time

A good Bit of the Time

Some of the time

A little bit of the time

None of the Time

Did you have a lot of energy?

All of the time

Most of the time

A good Bit of the Time

Some of the time

A little bit of the time

None of the Time

Have you felt downhearted and blue?

All of the time

Most of the time

A good Bit of the Time

Some of the time

A little bit of the time

None of the Time

Did you feel worn out?

All of the time

Most of the time

A good Bit of the Time

Some of the time

A little bit of the time

None of the Time

Have you been a happy person?

All of the time

Most of the time

A good Bit of the Time

Some of the time

A little bit of the time

None of the Time

Did you feel tired?

All of the time

Most of the time

A good Bit of the Time

Some of the time

A little bit of the time

None of the Time

SOCIAL ACTIVITIES:

During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with

your social activities (like visiting with friends, relatives, etc.)?

All of the time

Most of the time

Some of the time

A little bit of the time

None of the Time

GENERAL HEALTH:

How true or false is each of the following statements for you?

I seem to get sick a little easier than other people

Definitely true Mostly true Don't know Mostly false Definitely false

I am as healthy as anybody I know

Definitely true Mostly true Don't know Mostly false Definitely false

I expect my health to get worse

Definitely true Mostly true Don't know Mostly false Definitely false

My health is excellent

Definitely true Mostly true Don't know Mostly false Definitely false

Appendix B

USPSTF Randomized Controlled Trials (RCTs) and Cohort Studies

Criteria

- Initial assembly of comparable groups: RCTs—adequate randomization, including concealment and whether potential confounders were distributed equally among groups; cohort studies—consideration of potential confounders with either restriction or measurement for adjustment in the analysis; consideration of inception cohorts
- Maintenance of comparable groups (includes attrition, cross-overs, adherence, contamination)
- Important differential loss to follow-up or overall high loss to follow-up
- Measurements: equal, reliable, and valid (includes masking of outcome assessment)
- Clear definition of interventions
- Important outcomes considered
- Analysis: adjustment for potential confounders for cohort studies, or intention-to-treat analysis for RCTs (i.e. analysis in which all participants in a trial are analyzed according to the intervention to which they were allocated, regardless of whether or not they completed the intervention)

Definition of ratings based on above criteria

Good:	Meets all criteria: Comparable groups are assembled initially and maintained throughout the study (follow-up at least 80 percent); reliable and valid measurement instruments are used and applied equally to the groups; <u>interventions</u> are spelled out clearly; important outcomes are considered; and appropriate attention to confounders in analysis.
Fair:	Studies will be graded “fair” if any or all of the following problems occur, without the important limitations noted in the “poor” category below: Generally comparable groups are assembled initially but some question remains whether some (although not major) differences occurred in follow-up; measurement instruments are acceptable (although not the best) and generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are accounted for.
Poor:	Studies will be graded “poor” if any of the following major limitations exists: Groups assembled initially are not close to being comparable or maintained throughout the study;

unreliable or invalid measurement instruments are used or not applied at all equally among groups (including not masking outcome assessment); and key confounders are given little or no attention.
