# Bariatric surgery in patients with body mass index < 35: a systematic review and meta analysis

Chris Gerard Smith

Memorial University
St. John's, NL

Submitted to the Memorial University School of Graduate Studies
In partial fulfillment for the degree of

Master of Science in Medicine (Clinical Epidemiology)

Memorial University
St. John's, NL
May 2019

## Abstract

Background

Bariatric surgery has been shown to be safe and effective for the treatment of morbid obesity (BMI > 40 $kg/m^2$) and some related comorbidities. The goal of the current study is to explore the role of surgery in patients with moderate obesity (BMI < 35 $kg/m^2$).

Methods

Systematic review and meta analysis was performed focusing solely on patients with BMI < 35 $kg/m^2$ who underwent laparoscopic roux en y gastric bypass (RYGB), sleeve gastrectomy (LSG), or adjustable gastric banding (AGB). Data were limited to randomized controlled trials (RCT) and prospective cohort studies. Primary outcome measure was fasting plasma glucose (FPG). Secondary outcome measures included hemoglobin A1c (HbA1c), and other obesity related comorbidities.

Results

13 studies were included in the analysis, 4 randomized controlled trials and 9 prospective cohort studies. Surgery was associated with significantly improved FPG compared to medical therapy (weighted mean difference (WMD) -3.24, 95% confidence interval (CI) -4.45; -2.02). Surgery was also associated with improved HbA1c, body weight, BMI loss, waist circumference, and resolution of hypertension and dyslipidemia. These results were consistent across each surgical procedure. 2 RCTs compared RYGB to LSG. There was no difference with respect to glucose metabolism however RYGB was associated with greater BMI loss (WMD -1.07, 95% CI -1.81; -0.33) and decreased waist circumference (WMD -3.51, 95% CI -6.99; -0.03). Complication rates were comparable to morbidly obese subjects.

Conclusion

RYGB, LSG and AGB appear to be safe and effective in the treatment of obesity and related comorbidities and should be offered to patients with BMI < 35. RYGB and LSG have similar effects on FPG and HbA1c however REYGB appears to have improved results with respect to waist circumference and BMI.

## Table of Contents

**List of Tables**

**List of Figures**

**List of Abbreviations**

AGB…….Adjustable gastric band

BMI……Body mass index

CI ……confidence interval

DM……..diabetes mellitus

EWL…….Excess weight loss

FPG……fasting plasma glucose

GLP-1…….Glucagon like polypeptide - 1

HbA1c…….hemoglobin A1c

HDL…….high density lipoprotein

IDF…….International diabetes federation

ILDSG……..Ileal interposition with diverted sleeve gastrectomy

ILSG……...Ileal interposition with sleeve gastrectomy

LDL……...low density lipoprotein

LSG……laparoscopic sleeve gastrectomy

MeSH……Medical subject headings

NIH……..National institute of health

MD……...Mean difference

OR……..odds ratio

PRISMA…….preferred reporting in systematic reviews and meta analyses

LAGB……..Laparoscopic adjustable gastric band

LRYGB…….Laparoscopic roux en y gastric bypass

LSG……...Laparoscopic sleeve gastrectomy

NG……...Nasogastric tube

QUORUM……...Quality of reporting in meta analysis

RCT……...Randomized controlled trial

RYGB……Roux en y gastric bypass

ROB……Risk of bias

ROBINS-I……..Risk of bias in non randomized studies index

RR……..Relative risk

T2DM……Type 2 diabetes mellitus

WMD……...Weighted mean difference

**List of Appendices**

## Introduction

Obesity is one of the most threatening health issues seen today and is defined as having abnormal or excessive fat accumulation. In 2016, 1.9 billion adults worldwide were overweight. Of these, 650 million persons were obese. Obesity is a risk factor for cardiovascular disease, musculoskeletal disorders, cancer and premature death[1]. Body Mass index (BMI) is a very commonly used measure of weight versus height and is used to classify individuals as overweight or obese. It is calculated as a persons weight in kilograms divided by the square of their height in meters ($kg/m^2$). Ideal BMI is between 18.5-25. Overweight individuals are defined as BMI between 25.1-30. Obesity is defined as a BMI between 30-40 and these patients are further classified into class I obesity (BMI 30.1-35) and class II obese (BMI 35.1-40). The term 'moderate obesity' is used to describe patients with BMI 30-35 and we use the term morbid obesity to describe individuals with BMI > 40. The advantage of using BMI as a measure of obesity is that it is easily calculated for most patients and is commonly accepted as a measure of obesity so it brings with it generalizability. However, BMI tends to overestimate for large frame individuals and underestimate for shorter individuals. BMI also does not take into account muscle mass versus fat mass, which also can also provide some limitations[2].

Surgical management of obesity was first conceived in the 1950's. This began with the observation that patients who lost various lengths of small intestine lost weight despite increased caloric intake. This was also noted in various experiments involving intestinal resection in dogs, which caused fat malabsorption and weight loss. A similar observation was made in patients who lost part or all of their stomach, which was

typically associated with a significant degree of weight loss. These observations led to the development of surgically induced weight loss for overweight and obese individuals[2].

The term 'bariatric' comes from the Greek 'baro' meaning pressure and is defined as the branch of medicine pertaining to the prevention and treatment of obesity, thus the term bariatric surgery or also metabolic surgery, as many of these procedures will have profound metabolic effects on the body. Excess body weight can be calculated using metropolitan life tables for ideal body weight for height[3]. We can therefore calculate excess weight loss post intervention and this is commonly used as a measure of success following bariatric surgery. Various procedures have been developed over time and rates of excess weight loss have varied from 25-85%[4].

Bariatric surgery has also been shown to be associated with resolution of obesity related comorbidities such as diabetes mellitus, hypertension, hypercholesterolemia, sleep apnea and arthritis[5]. Typically for any given procedure the rate of resolution of comorbidities tends to parallel the rates of excess weight loss. Procedures are generally categorized as having restrictive effects if the proximal gastrointestinal tract is modified to limit intake, or malabsorptive effects if a portion of the small intestine is excised or bypassed. Various procedures can also have a combination of restrictive and malabsorptive effects[6].

Surgery is associated with certain morbidity and mortality, and when considering any surgical intervention one must weigh the risk of surgery versus the potential benefit. With respect to obesity, surgery has been shown to be an effective tool in sustaining weight loss and in treating obesity related comorbidities such as Diabetes mellitus (DM), hypertension, hypercholesterolemia, obstructive sleep apnea and osteoarthritis. Obese

patients typically have higher than average risk for perioperative complications and some series report mortality rates as high as 2-3%, which is much higher than for most elective surgeries[7]. Most early studies explored the effects of bariatric surgery on patients who were morbidly obese (BMI >40) or class II obese (BMI >35) with significant comorbidities and in 1994 the National institute of health (NIH) developed consensus guidelines restricting access to bariatric surgery to these patients[8].

More recent evidence has revealed that patients who are overweight or class I obese (BMI 25-35) may also benefit from bariatric surgery, particularly those patients afflicted with diabetes. On the basis of some of these early studies the NIH published revised guidelines, which made allowances for these patients in the setting of a clinical trial[9]. This created some controversy in providing a surgical intervention with associated risks to a patient population where the absolute benefit is unclear. Much of the data to date is from retrospective studies with small numbers of patients and because of this we felt that a systematic review of the literature and meta analysis would be necessary to fully explore the effect of bariatric surgery in patients with BMI less than 35. On this basis the objectives of our study are (1) To determine the efficacy and safety of bariatric surgery in patients with a BMI < 35 $kg/m^2$ (2) To determine whether one surgical procedure will be superior in this patient population and (3) To fully explore the effect of surgery on less well studied obesity related comorbidities such as sleep apnea and osteoarthritis in patients with BMI < 35.

## Co Authorship Statement

The completion of this project would not be possible without the help and support of my Supervisors, Co Supervisors and everyone involved. Contributions were made in the following manner:

Project Design and Research Protocol

Dr. Chris Smith – Project inception and development of research question: Current role for bariatric surgery in patients with body mass index < 35.

Dr. Joseph Mamazza – University of Ottawa – Expert guidance in field of bariatric surgery, project design and development of research protocol.

Dr. Bryan Curtis – Expert guidance in field of Epidemiology, feedback regarding research protocol and project design.

Dr. Michael Hogan – Expert guidance in field of Epidemiology research, feedback regarding research protocol and project design.

Practical aspects of research

Dr. Chris Smith – Determination of study eligibility, article review for Inclusion/exclusion, database review, creation of funnel plot.

Dr. Dmitry Terterov – Independent review of selected studies to confirm Inclusion/exclusion (Independent verification by Dr. Priscille Cyr in the Event of disagreement).

Data analysis

Dr. Chris Smith – Determination of summary measures, synthesis of and interpretation of data.

Dr. Hensley Mariathas – Statistical analysis and creation of forest plots, Data extraction and interpretation of the data.

Manuscript Preparation

Dr. Chris Smith – Write up and review of final manuscript, editing and Correction based on feedback from co-authors.

Drs. Bryan Curtis, Michael Hogan, David Pace, Darrell Boone, Vanessa Falk, Aryan Modasi, Erin Mayo, Joe Mamazza and Lisa Bacque – feedback on format, grammar, etc. and final interpretation of the results.

The authors of this paper did not receive related funding during any period of the project.

Accepted as poster presentation at the Society of American Gastrointestinal and Endoscopic surgeons annual meeting in Baltimore, MD, USA April 2019.

# Chapter 1 - Background

## 1.1 Open surgery era

Many surgical procedures in the abdomen are performed through a very large incision in the abdominal wall through which surgeons can manipulate organs directly with their hands or surgical instruments. This is commonly referred to as 'open' surgery. This term is often used to differentiate from laparoscopic surgery, which involves intra abdominal surgery through very small incisions and using longer instruments. Up until the early 1980s abdominal surgery was done almost exclusively via the open approach. In 1963 Payne et. al. performed the first surgical procedure for obesity, which was the jejunocolic bypass. This involved division of the proximal small bowel approximately 50cm down stream and connecting the proximal small bowel segment to the mid transverse colon. The distal small bowel segment was closed leaving a long blind loop. Later, the procedure was modified by changing the site for the anastomosis to the ascending colon as the initial procedure was associated with significant diarrhea. Despite this modification the jejunocolic bypass was associated with significant steattorhea or loss of fat in the stool, dehydration, electrolyte imbalances, perianal complications, hepatic cirrhosis or scarring of the liver and subsequent hepatic failure. Despite reasonable weight loss and resolution of obesity related comorbidities the procedure was largely abandoned because of these complications[6].

In 1969 Payne and DeWind described the jejunoileal bypass. This involved a bypass of a large portion of the small intestine without proximal restriction. This approach was initially used to treat super obese patients or those with BMI > 60kg/$m^2$. This procedure was associated with significant weight loss and resolution of obesity related comorbidities but also led to significant long-term complications such as

Steatorrhea, kidney stones, abdominal distension, and hepatic failure secondary to protein malabsorption. Hepatic failure was one of the most serious complications of this procedure and initial mortality rates over the first 2 years were around 4%. These complications ultimately led to this operation being superseded by other procedures. In fact most patients who have undergone jejunoileal bypass subsequently went on to have this procedure reversed and converted to an alternative procedure in an attempt to avoid some of these long term complications[2].

Gastric bypass (Appendix 1 and 2) was developed out of the observation that patients post gastric surgery lost significant amounts of weight. Mason and colleagues conceived the procedure of gastric bypass in the 1967. This involved construction of a small approximately 30ml gastric pouch via gastric partition. This was initially performed using staplers that would occlude the stomach but not physically divide it. The procedure also involves division of the small bowel approximately 40cm down stream and creation of a 100cm alimentary limb which is approximated to the gastric pouch[10]. In 1977 Alden and colleagues altered the procedure by using alternate stapling devices, which would physically divide the stomach.[11]. Long term data on patients undergoing gastric bypass have been published over 25 years of follow up and have shown that excess weight loss is in the 60-70% range and has been sustainable over time. This makes gastric bypass the traditional gold standard procedure in bariatric surgery[12]. Over the years various forms of gastric partition were created in an attempt to simplify the operation. These procedures involved the use of various types of mesh and stapling devices. Although associated with significant weight loss initially, long term complications such as mesh erosion, stricture,

pouch dilation, esophageal dilation, staple line dehiscence and weight regain led to these modifications falling out of favor[13].

Biliopancreatic diversion (BPD) is an alternative procedure developed to maintain the restrictive effect of the gastric bypass while enhancing the malabsorptive effect and was initially described by Scopinaro in 1979. This procedure differs from the gastric bypass in that rather than dividing the stomach, a portion of the stomach is removed and similarly to gastric bypass the small bowel is divided and connected to the gastric pouch. This procedure provided excellent long-term weight loss with rates of excess weight loss of 85%, however was associated with very severe protein and micronutrient malabsorption[14].

In the 1993 Marceaux and colleagues refined the biliopancreatic diversion procedure creating the duodenal switch procedure. This procedure was initially described as a two-part procedure. The first part consisted of a sleeve gastrectomy, which is resecting about 70% of the stomach leaving a narrow sleeve (Appendix 3). The second part of the procedure would essentially involve bypass of the small bowel. This was in an attempt to maintain excellent weight loss but limit the severe protein and micronutrient malabsorption. Biliopancreatic diversion and similar procedures are performed less commonly, likely because of reports of mortality of greater than 2%, which is substantially higher than other metabolic surgeries[14].

## 1.2 Laparoscopic surgery era

Laparoscopic surgery refers to performing abdominal surgeries through small incisions around 5-10mm and using small instruments and a camera that pass through

these incisions. Initially developed in 1980 this approach has become the standard of care for many surgical procedures and has produced tremendous results including decreased pain and quicker recovery for most patients compared to traditional open approach. For many years open gastric bypass was the standard of care. In the early 1990's the laparoscopic approach simplified these operations and led to significant improvements in patient morbidity, hospital stay and return to normal activities[15]. Today the laparoscopic approach has become the standard of care in bariatric surgery. The 1990's also saw the invention of the adjustable gastric band (Appendix 4). This consists of a hollow tubing, which is placed around the stomach to create a restrictive effect where patients would only be able to consume small amounts of food. The tubing is connected to a subcutaneous port through which saline can be injected or withdrawn thereby adjusting the restrictive effect of the pouch. Excess weight loss is in the range of 25-35%. While the short term complication rate is favorable, excess weight loss can be substantially lower compared to some other procedures and that the re-operative rate for these patients can be as high as 40%[12].

An interesting observation in patients undergoing the duodenal switch operation was that many patients lost a significant amount of weight, and in fact approximately half did not require any further surgery. This led to the idea that the sleeve gastrectomy portion of the procedure could be a stand-alone bariatric procedure. Laparoscopic sleeve gastrectomy (LSG) is associated with excellent weight loss and resolution of obesity related comorbidities, and very low complication rates. In 2001, Sleeve gastrectomy was described as a stand-alone procedure. This procedure does carry a risk of gastroesophageal reflux and staple line leak however these appear to be uncommon.

Excess weight loss appears to be around 60%, which is comparable to gastric bypass however long term data are lacking. When performed laparoscopically this procedure has been shown to be a safe and effective treatment for obesity and related comorbidities[16].

There is much debate in the literature and in the bariatric community about which procedure is superior and certainly there are factors which may make one procedure more appropriate for any given patient. Varying degrees of excess weight loss, perioperative morbidity, complications, and long-term results have led to such controversies. In 1994 the National Institute of Health developed guidelines regarding which patients would be eligible for bariatric surgery. It was felt that patients with a 'BMI over 40' or 'greater then 35 with at least one obesity related comorbidity' would be eligible. These guidelines were developed based on the best available evidence at the time. Bariatric surgery in this patient population has been shown to be safe and effective in sustaining long-term weight loss and resolution of comorbidities[17].

**1.3 Surgery for moderate Obesity**

Diabetes Mellitus along with other obesity related comorbidities are serious conditions with potentially devastating complications that affects all age groups. In 2012 the International Diabetes Federation (IDF) estimated that 371 million people worldwide were afflicted with diabetes. That number is expected to rise to 552 million by 2030. Diabetes is the leading cause of blindness, end stage renal disease, and non-traumatic amputation in Canadian adults[18]. In 2017 the IDF recommended that overweight and moderately obese subjects (BMI between 25-35) should consume a low calorie diet between 800-1200 calories per day and lose approximately 10 pounds regardless of

starting body weight. This recommendation was made for patients with type II diabetes mellitus (T2DM). As part of this recommendation based on current evidence in the literature, a recommendation was also made by IDF to consider bariatric surgery for people with T2DM and BMI between 30 and 35 when the metabolic response to regular treatment has been poor[19]. There have been many studies looking at bariatric surgery in patients with lower BMI (20-35). Much of these studies were performed initially in Asian populations where there is extremely high prevalence of T2DM in normal weight or overweight individuals[20]. Much of the earlier studies were retrospective in nature involving lower numbers of patients however as evidence mounted these procedures were being applied to a wider range of patients. Today there are many prospective and randomized controlled trials that have been performed on overweight and obese patients, which is the subject of this analysis. Our goal is to review the highest levels of evidence to establish whether bariatric surgery is safe and effective in this patient population.

# Chapter 2 - Literature review

## Introduction

The main goal of any surgical intervention is to provide a high level of efficacy while maintaining a favorable safety profile. Bariatric surgery has evolved over time and a preliminary review of the literature is important to ensure that the intervention in question is indeed safe and efficacious. It is also important to review the current literature with respect to our research question. Our objectives with respect to literature review are: (1) Review evidence for bariatric surgery in patients with morbid and class II obesity. (2) Explore evidence for bariatric surgery in overweight and class I obese patients

(moderately obese). (3) Review previous systematic reviews of patients who are moderately obese.

## 2.1 Surgery for Morbid obesity

### Introduction

Initial data were presented with open gastric bypass, which had a very strong impact in morbidly obese patients. The 1980s saw the advent of laparoscopic surgery, which had dramatic effects on morbidity, mortality, length of stay in hospital, return to usual duties and overall faster recovery when compared to open surgery[13]. The 1990s saw the advent of newer somewhat novel procedures such as adjustable gastric banding (AGB) and sleeve gastrectomy. These procedures were shown to have favorable complication profiles while maintaining excellent excess weight loss and resolution of obesity related comorbidities in morbidly obese patients. Here we explore some of the evidence for bariatric surgery in morbidly obese and moderately obese subjects. The included studies were chosen to give broad representation of the evolution of bariatric surgery and because they are significant in their own right.

### Literature Review

### Mason et. al. 1969[10]

Mason and colleagues developed the open technique of gastric bypass after the observation that patients undergoing total gastric resection remained thin. They describe the 90 percent gastric bypass operation for obesity given that about 90 percent of the stomach is excluded from digestion. Upper midline laparotomy incision is described in

detail and the left lobe of the liver is mobilized towards the patients' right side. The stomach is then divided creating an approximately 10 percent by volume proximal gastric pouch. The ligament of treitz is divided and a short loop retro colic (behind and under the transverse colon) gastroenterostomy is performed. The mesocolon was secured to the gastric pouch by sutures to prevent internal hernia. The proximal gastric pouch is also sutured to the distal excluded portion of the stomach such to prevent torsion or intussusception of this segment. The idea was that this would help prevent gastric antral stasis and in turn help prevent development of jejunal ulcers. This procedure was initially performed in dogs to assess safety and feasibility before being trialed in humans.

Over a 3-year period this procedure was performed on 24 patients who averaged 222% of their estimated normal body weight. The range of preoperative weights was 80.9kilograms (kg) to 295.5kg. The patients were equally divided by sex and ranged in age from 22-68 years. The authors looked specifically at patients with duodenal ulcers and proposed that the gastric bypass procedure would be an effective treatment. Total of 10 patients had history of duodenal ulcer and all patients had relief of symptoms post operatively. There were two deaths in the cohort. One patient died from peritonitis in the absence of any suture line dehiscence and the other from pulmonary embolus. Several patients required monitoring in the intensive care unit with at least one patient requiring prolong ventilator support and tracheostomy. Two patients were readmitted postoperatively because of persistent vomiting and dehydration. Average weight loss during the follow up period was 39kg. Jejunal ulceration was initially a concern however only one patient developed jejunal ulcer. This was felt to be due to stasis in the excluded portion of the stomach and the patient underwent revisional surgery however this was not

detailed in the report. The authors conclude that gastric bypass can be recommended in the treatment of severe, intractable obesity with an acceptably low rate of jejunal ulceration. They go on to say that physically active, young or middle aged people weighing in the range of 200% of estimated normal body weight will respond better than others. Preoperative weight reduction may be required and Anesthesiologists experienced in care of obese patients and intensive care facilities will be necessary. Limitations of this study included non-randomized nature and of course much of the morbidity of the surgery associated with open technique. This was a landmark study and one of the first studies to explore the effects of gastric bypass on morbidly obese subjects, a surgical treatment that today remains one of the most commonly performed and efficacious procedures in this patient population.

**Pories et. al. 1982[21]**

Randomized controlled trial of 87 morbidly obese subjects to open gastric bypass versus gastric partition. Gastric bypass was performed by creating a small 50ml gastric pouch proximally using a non-cutting stapler that essentially occluded the stomach at this level. The small intestine was then brought through in a retro colic fashion and anastomosed to the gastric pouch using a hand-sewn technique. The size of the gastrojejunostomy was said to be standardized to 0.8cm by suturing over a number 18 nasogastric tube. Gastric partition was completed in a similar fashion however instead of a gastrojejunostomy anastomosis the gastric pouch was sutured to the inferior portion of the stomach that had been divided creating a gastro-gastrostomy. This was done in a similar fashion via hand-sewn approach over number 18 nasogastric tube. Patients were

monitored in the intensive care unit for 24 hours and then on the wards for 3 days with nasogastric (NG) tube in situ until they passed flatus. At that point the NG tube was removed and patients were started on a fluid diet and slowly advanced based on symptoms. There was no statistically significant difference between the groups in terms of age, weight, sex, and incidence of DM and hypertension. Follow up was performed at 3, 6, 9, 12 and 18 months. Gastric bypass patients had greater weight loss at each point in follow up which was statistically significant. This was sustained at all points and at 18 months follow up the gastric bypass group had a mean weight 79.2kg. compared to the gastric partition group at 110.1kg. Both procedures had a positive effect on diabetes and hypertension with all but one patient with diabetes being normoglycemic postoperatively. Of patients with hypertension only 1 of 19 in the gastric partition group had persistent hypertension and 2 of 16 in the gastric bypass group remained hypertensive. Complications were equally distributed amongst the groups. The authors conclude that gastric bypass is a superior operation to gastric partition. This study highlighted the reproducibility and sustainability of the gastric bypass procedure. It also highlighted the dramatic effect on weight loss that bypassing the proximal gastrointestinal tract would provide and that pure restriction could not provide the same effect. This result would be reproduced in years to come.

**Wittgrove et. al. 1996[15]**

Early study looking at effectiveness of laparoscopic roux en y gastric bypass (LRYGB). Cohort study looking at 27 patients who underwent LRYGB. Patients were selected using criteria set out by NIH consensus panel. The technique of open RYGB has

been described previously. Routine preoperative workup was completed and patients were operated via laparoscopic approach. Similar to technique described by Mason the stomach is completely divided creating small 15ml gastric pouch. Small bowel is divided and 75cm roux limb is brought in retro colic fashion and anastomosed to the gastric pouch using 21mm circular stapler. Standard enteroenterostomy is performed to complete the procedure. Upper gastrointestinal series was completed on first postoperative day using water-soluble contrast media and clear fluid diet was started on the same day. Average length of stay was 2.8 days. Complications were comparable to earlier studies with one anastomotic leak in this series. There was 1 intra abdominal abscess and 2 patients had stenosis of the gastrojejunostomy requiring dilation. Patients were followed up at 3, 6, 9, 12 and 18 months. Mean excess weight loss at 18 months was 80%. Hypertriglyceridemia, diabetes, stress incontinence, arthritis and sleep apnea were resolved in all patients post operatively. Gastroesophageal reflux persisted in 1 of 22 patients and hypercholesterolemia in 2 of 14 patients. Patients were seen to have shorter hospital stay, earlier return to usual activities and superior cosmetic effect as compared to studies looking at open surgery. The authors conclude that LRYGB is a safe, effective treatment for morbid obesity and because of significant reduction in length of stay potentially cost effective. The advent of laparoscopic surgery had a major impact on the outcome of surgical patients and this was also seen in bariatric patients undergoing surgery. This study was one of the first reports to show that laparoscopic RYGB was safe in morbidly obese patients with a favorable side effect profile.

**Dixon et. al. 2002[22]**

The 1990s and early 2000s saw further advancement in bariatric surgery as surgeons pushed to develop newer techniques that would decrease surgical morbidity and complications such as ulceration and internal hernias that were seen with gastric bypass. Dixon and colleagues published their data on laparoscopic adjustable gastric band (LAGB) for the treatment of morbid obesity. LAGB consists of a silicone elastomer with inflatable inner shell that is placed around the proximal stomach and secured using a buckle closure. The band is connected by tubing to an access port that is placed in the subcutaneous tissues anterior to the abdominal wall fascia. The port can then be accessed via syringe and the inner diameter of the band can be adjusted by injecting or withdrawing fluid as desired. This system was designed as an option for individuals who do not wish to have irreversible alteration to proximal gastrointestinal tract. This was a cohort study and 50 patients were included. Laparoscopic surgery was achieved in 47 while 3 patients had open approach for revisional surgery and band placement. Median hospital stay was 2 days. There were 2 patients with wound infections and 1 patient required postoperative respiratory support. All of these complications occurred in those partaking open surgery. While early postoperative outcomes are favorable with respect to complications, late issues such as prolapse of the stomach through the band (20%) and band erosion (6%) were seen not infrequently and most required operative intervention. At 1 year of follow up statistically significant changes were seen in fasting plasma glucose, hemoglobin A1c, fasting triglyceride, liver enzymes and need for oral hypoglycemics. Mean excess weight loss at 1 year was 38+/-14%. There were also significant improvements in qualitative markers such as physical function, pain, general

health, and energy. The authors conclude that LAGB is an effective treatment for T2DM and obesity related comorbidities in morbidly obese patients. The effects with AGB are somewhat less robust than what is seen with other surgical procedures such as LREYGB however the short-term complication rate is favorable.

**Nocca et. al. 2007[16]**

Laparoscopic sleeve gastrectomy (LSG) has been described as stand alone bariatric procedure after it's inception during duodenal switch procedure and has been widely adopted to morbidly obese patients. This is a multicenter prospective study. Over a 3-year period 163 patients underwent LSG by 5 different surgeons. Average body mass index (BMI) was 45.9 kg/m2. Surgery was standardized amongst all 5 surgeons. Left lobe of the liver was retracted anteriorly and the gastro colic ligament was divided and the bursa minor was entered using ultrasonic sheers. The dissection along the greater curvature started 10cm from the pylorus and progresses toward the diaphragm. 36french calibration tube was placed trans orally along the lesser curvature to perform a controlled vertical gastrectomy. The gastrectomy was completed using endoscopic linear stapler of appropriate cartridge and the staple line may have been buttressed with sutures or with absorbable material. The staple line was checked for leakage by injecting methylene blue through an orogastric tube. Suction drain was left in situ. All patients were optimized perioperatively in a standard fashion. Barium swallow was performed on postoperative day 2. Nasogastric tube was removed at this point and the patients were started on liquid diet.

There were no conversions to laparotomy in this group. LSG was primary procedure in most subjects however 22 persons were operated after failure of AGB. Perioperative complication rate was 7.4%. Patients were followed for total of 2 years post operatively. Excess weight loss (EWL) and BMI at 2 years were 61.5% and 31.6kg/m2. Weight regain was reported in 10 patients. This study is one of the first reports of sleeve gastrectomy as a primary bariatric procedure. The Authors conclude that sleeve gastrectomy is a safe and effective restrictive bariatric procedure to treat morbid obesity in select patients. They concede that weight regain, quality of life and evolution of morbidities due to obesity need to be evaluated in long term follow up.

**Summary**

Surgical management of morbid obesity has evolved and today surgery remains one of the most efficacious long-term treatments for severe obesity[14]. Moreover, Surgery has been shown to provide effective, durable treatment for T2DM and other obesity related comorbidities such as hypertension, hypercholesterolemia, sleep apnea, and osteoarthritis. The advent of laparoscopy and other minimally invasive techniques has decreased the morbidity of these procedures substantially where the risk benefit profile weighs heavily in favor of surgery. The American Diabetes association recommends consideration of bariatric surgery for any patient with T2DM and BMI > 40 where hyperglycemia persists despite adequate medical and lifestyle intervention. In fact, a substantial amount of literature exists to show that surgery when compared to medical and lifestyle modifications provides superior glycemic control and reduction of cardiovascular risk factors in obese patients with T2DM[14]. Surgery has also been shown

to be associated with long-term improvement in mortality[23]. I believe that the evidence weighs strongly in favor of bariatric surgery for obese individuals as highlighted by the previous studies. I believe that the true effect of these surgeries particularly with respect to patients with moderate obesity is largely unknown and will be the subject of much future research.

## 2.2 Surgery for Moderate Obesity

### Introduction

Over the years a multitude of data has arisen supporting the implementation of bariatric and metabolic surgery for patients with severe obesity (BMI > 35). The dramatic effect that has been seen with respect to T2DM has led investigators to consider whether surgery will be effective in patients with BMI < 35. Here we review some of the data regarding surgery in patients with BMI < 35. While this is not a comprehensive review we feel the following studies are representative of surgery in this patient population.

### Literature Review

#### Noya et. al. 1998[24]

Touted as being one of the first reports of metabolic surgery in patients with BMI < 35, Noya and colleagues performed biliopancreatic diversion (BPD) without gastric resection in 10 patients. Mean BMI was 33.2kg/$m^2$. Mean age and weight were 52.1 years and 85.2kg respectively. 5 males and 5 females in total. All patients had hypercholesterolemia and hypertriglyceridemia for over 5 years. All patients had T2DM with two patients taking insulin and two others taking oral hypoglycemic agents.

Remaining patients were diet controlled. Duodenum was transected and along with the small bowel at a point 50cm proximal to the ileocecal valve. Small bowel was connected to the duodenum using biodegradable ring in 3 cases and hand sewn in 7. Postoperative course was described as uneventful with patients resuming diet on postoperative day 10-12. All patients were discharged by postoperative day 15. Maximal follow up reported here was 12 months. All patients had cholesterol and triglyceride levels return to normal. 9 patients had normal glycemic values with one patient who had been taking 70 units of insulin per day decreased to 35 units of insulin with discontinuation of oral medications for same. All patients were normotensive post operatively. Mean BMI decreased to 30.5kg/$m^2$. One patient suffered pulmonary embolus that responded to medical treatment. One patient required re operation to treat obstruction secondary to the biodegradable ring. Weight loss was described as moderate with 10-15kg reduction in body weight in the first month after surgery. The authors conclude that this modification of the original BPD procedure without gastric resection is effective in controlling lipid metabolism and T2DM in patients who are moderately obese. Given that the amount of weight loss seen with this approach is lower than would be expected with original BPD, the modification described here is not recommended for morbidly obese patients. Limitations of the study were small number of patients, retrospective nature and limited generalizability due to higher complication rates compared to other bariatric procedures. This study has been touted as being one of the first series of patients with moderate obesity to undergo bariatric surgery[25].

**DePaula et. al. 2008[26]**

Prospective study of 39 patients undergoing either ileal interposition with sleeve gastrectomy (ILSG) or ileal interposition with diverted sleeve gastrectomy (ILDSG). The study population consisted of patients with BMI between 23.4-34.9 $kg/m^2$ with mean BMI of 30.1$kg/m^2$. Inclusion criteria also included type 2 diabetic patients whose disease had been diagnosed for at least 3 years, HbA1c > 7.5, stable weight defined as no significant change over previous 3 months, and evidence of stable treatment with oral hypoglycemic therapy or insulin for at least 12 months. Patients were not assigned randomly to either surgical group. Patients > 66 years old, previous major upper abdominal surgery, pregnancy, severe comorbidities, use of appetite suppressant, and presence of eating disorder or other endocrine disorder were all listed as exclusion criteria. IlSG started with division of the jejunum 50cm from the ligament of treitz. A 100cm segment of ileum is then isolated by division about 50cm proximal to the ileocecal valve. This 100cm segment of ileum is then anastomosed to the proximal jejunum at the point previously transected. Ileal interposition is proposed to improve glucose metabolism by different mechanisms including up regulation of glucagon like polypeptide-1 (GLP-1) by early food contact with ileal mucosa. This in turn helps regulate early phase insulin secretion and maintaining glucose homeostasis. Sleeve gastrectomy is standard part of the procedure. The greater curvature of the stomach is devascularized using ultrasonic scalpel. 30french orogastric tube is placed to calibrate the sleeve along the lesser curvature. The gastric resection is started at the antrum and carried to the diaphragm using linear stapler. Staple line is over sewn. ILDSG combines ileal interposition as described above with diverted sleeve gastrectomy. Once the gastric

sleeve is created the dissection is carried beyond the pylorus and the proximal duodenum is divided using a linear stapler. The proximal duodenal mucosa along with the gastric remnant is then attached to the distal end of the jejunum divided proximally. The proximal jejunum that had been divided 50cm from the ligament of treitz is then anastomosed to the interposed ileum. ILSG was performed for 23 patients and ILDSG in 16 patients. Median hospital stay was 4.3 days. Major complications including gastric leak and acute renal failure were experienced in 10.3% of patients. Minor complications including urinary tract infection and prolonged ileus were noted in 15.4% of patients. Total of 30 patients were followed for mean of 7 months (range 4-16 months). Significant improvements ($p < 0.001$) were seen in Hemoglobin A1c, fasting glucose, total cholesterol, and triglycerides (6.3 +/- 0.9 vs. 8.8 +/- 1.7; 116.7 +/- 33.1mg/dl vs. 210.7 +/- 66.6mg.dl; 165.7 +/- 27.6mg/dl vs. 215.1 +/- 49.9mg/dl; and 131.3 +/- 80.2mg/dl vs. 250.5 +/- 168.4mg/dl respectively). Mean percentage loss of initial weight was 22% with 5.2% of patients having BMI below 20kg/m2.

The authors conclude that IlSG and ILDSG may be considered for treatment of metabolic syndrome and type 2 diabetes mellitus in non-morbidly obese subjects. A profound effect was seen with respect to glucose metabolism and dyslipidemia. Weight loss as a secondary measure was also dramatic. The limitations of the study are small number of patients, brevity of the follow up period and lack of relevant control groups. One must also consider the staggering complication rate with more than 10% of patients suffering major complications in the early postoperative period. This is not surprising given previous studies of the ileal interposition procedure in obese subjects that also reported high rates of complications that ultimately led to these surgeries being largely

abandoned in favor of other surgeries that have more favorable perioperative outcomes such as gastric bypass and sleeve gastrectomy.

**Serrot et. al. 2011[4]**

Much of the strongest evidence for bariatric surgery in morbidly obese subjects came from data on patients undergoing open gastric bypass and later laparoscopic roux en y gastric bypass (LRYGB). This led investigators to explore the effects of LRYGB on moderately obese subjects, hence much of the literature to date on this patient population comes from patients who underwent gastric bypass. Serrot and colleagues performed a retrospective review of their data in patients undergoing LRYGB. Inclusion criteria were patients with BMI < 35 and T2DM. Interestingly these patients would have had BMI > 35 at some point in their life however at the time of surgery BMI was < 35. They also tracked patients undergoing routine medical management in their center and matched them to patients with similar BMI undergoing surgery. Patients underwent standard gastric bypass with 15-30ml gastric pouch and 75-150cm roux limb and 75-100cm biliopancreatic limb. Patients were followed up for a maximum period of 12 months. Patients in medical arm were followed similarly and had their medications adjusted by dedicated pharmacist and diabetic nurse. They also received counseling regarding nutrition, exercise and weight management. Total of 34 patients were included in the analysis. Participants at baseline were similar in almost every respect except the surgery group had higher proportion of females and higher HbA1c (13 vs. 6, P = 0.04; 8.2 vs. 7.0 p = 0.04). Significant differences were seen with respect to weight loss. Surgical group had change in BMI from 34.6+/-0.8 $kg/m^2$ to 25.8+/-2.5 $kg/m^2$ compared to essentially no

change in the medical group from 34.1+/-1.0 kg/$m^2$ to 34.3+/-2.1 kg/$m^2$ ( $p < 0.001$). HbA1c decreased in the REYGB group from 8.2 +/-2.0 to 6.1+/-2.7 but did not change in the medical group (7.0+/-0.7 to 7.1+/-1.8 $P < 0.001$). Blood pressure and LDL cholesterol did not significantly change in either group. Fewer patients in the surgical group were taking medications for glycemic control at the end of the study with 71% of patients taking less medications at one year compared to 6% in the non-surgical group ($p < 0.001$). Total readmission rate for the surgical group was 18%. At one year 2 patients had developed incisional hernias that required repair and 2 further patients developed marginal ulcers that were managed medically. No mortalities were reported in either group. Limitations of the study included the small numbers of patients and several differences that were present between the groups at baseline. Retrospective nature of the analysis was also identified as a limitation. The authors conclude that RYGB could be a safe and effective treatment in patients with moderate obesity, particularly in the setting of T2DM without risk of hypoglycemia.

**Lee et. al. 2011**[27]

Randomized controlled trial comparing moderately obese subjects undergoing gastric bypass versus sleeve gastrectomy (LSG). Patients were eligible if they were between 30-60 years of age and had BMI between 25-35 and had poorly controlled T2DM as defined as HbA1c > 7.5%. Patients were excluded if they had undergone previous bariatric procedure, drug or alcohol addiction, neoplasm or evidence of portal hypertension. Block randomization to either gastric bypass or LSG was performed with block size of 10. Randomization was performed in the operating room after

pneumoperitoneum was established. Sleeve gastrectomy was performed in standard fashion by resecting the greater curvature of the stomach from approximately 4cm proximal to the pylorus to the diaphragm using a stapler. Remnant stomach was approximately 2cm wide. Staple line was over sewn using running absorbable suture. For the gastric bypass group a long sleeved gastrectomy was created similarly and a loop gastroenterostomy was created with the small bowel approximately 120cm from the ligament of treitz. Postoperative care was standardized with patients being discharged on postoperative days 3-4. Primary endpoint of the study was glycemic control at 12 months post surgery. This was assessed as participants achieving remission of T2DM, defined as FPG < 126mg/dL and HbA1c values less than 6.5% without the use of oral hypoglycemics or insulin. Secondary outcome measures were weight, blood pressure, waist circumference and fasting lipids. Total of 60 patients were randomized with 30 patients in each group. Mean BMI and age were 30.3 and 45 years respectively. Both study groups were demographically similar. There were no deaths or major complications in each group. Surgical time was similar between groups and the mean hospital stay was 2.2 in the gastric bypass group and 2.1 days in the LSG group. Overall 70% of patients experienced T2DM resolution after 12 months. This was significantly better in the gastric bypass group than the LSG group (93% vs. 47% p = 0.02). Gastric bypass patients also had better weight loss and significantly lower FPG, waist circumference, HbA1c and lipids at 6 and 12 months post operatively. This was perhaps the first randomized controlled trial comparing surgical interventions in patients with moderate obesity (BMI < 35). Although both surgical groups appeared to have excellent results with low rates of complications, the gastric bypass group appeared to be superior

in almost every aspect. Exclusion of the duodenum and proximal foregut has been hypothesized as an important component of metabolic surgery and has been put forth as an explanation of why traditionally bypass procedures can achieve rates of resolution of T2DM of > 80% while purely restrictive procedures typically result in about 50% resolution. The results of this study seem to support that hypothesis however one main limitation is the lack of hormone data in this study. Sleeve gastrectomy was initially touted as being purely restrictive procedure however proponents of this procedure will argue that removal of 80% of the stomach has metabolic effects beyond the restriction that occurs. The lack of hormonal data in this study makes any direct comparisons in this regard virtually impossible. The authors conclude that gastric bypass surgery is more effective than sleeve gastrectomy for the surgical treatment of T2DM and control of metabolic syndrome. They go on to add that duodenal exclusion plays an important role in the mechanism of diabetes mellitus remission.

**Zhu et. al. 2012[28]**

Prospective study of 30 Chinese T2DM patients with BMI < 35 who underwent LRYGB. Patients were excluded if they had history of open abdominal surgery, unstable psychiatric illness, alcohol or drug dependence, active helicobacter pylori infection or age > 65 years. Patients underwent thorough preoperative workup and had LRYGB performed under general anesthesia with 4 trocars. Standard technique of RYGB was performed. Mean operative time was 2.85h. Clear fluid and liquid diets were started on the 4$^{th}$ and 7$^{th}$ postoperative day respectively. Solids were started in the 4$^{th}$ postoperative week. Patients were followed up at 3, 6 and 12 months after surgery. Of the 30 patients

who participated there were 22 males and 8 females. Mean age was 48.16. Mean FPG was 8.01 mmol/l and HbA1c was 8.02%. No significant surgery associated outcomes were identified except one patient who developed gastroparesis and had prolonged hospital stay. No mortalities were noted. There were significant changes noted in BMI (28.5+/-1.85 vs. 32.20 +/-1.56, P = 0.015), Waist – hip ratio (0.83+/-0.51 vs. 0.96+/-0.07, P = 0.010), Fasting plasma glucose (5.95+/-1.10mmol/L vs. 8.01+/-2.08mmol/L, P = 0.040), and HbA1c (5.59+/-1.02 vs. 8.02+/-1.77 P = 0.049). The authors discuss that Chinese patients with T2DM typically have BMI < 35 and central obesity. Bypassing the proximal GI tract seems to have a dramatic effect on glucose metabolism independent of weight loss although there was a significant change in BMI post operatively in this study. Limitation of this study is the relatively short term follow up of 12 months. The authors conclude that LRYGB is a safe and effective treatment for T2DM in non-morbidly obese patients with potential for complete remission of the disease. They also concede that further studies will be necessary to establish long-term efficacy.

**Summary**

Early results of bariatric surgery in moderately obese patients are promising. While early studies utilizing ileal interposition type procedure revealed dramatic effect on weight loss, glucose metabolism and other obesity related comorbidities, relatively high perioperative complication rate limits the transferability of these results. Further studies utilizing gastric bypass, sleeve gastrectomy and adjustable gastric banding via laparoscopy have shown promising data with a more favorable safety profile. While the early results are promising, majority of the data thus far has been retrospective in nature

and most studies involve relatively low numbers of patients. These limitations inherently weaken any conclusions drawn from this data. This opened the door for further prospective studies, randomized controlled trials and meta analyses.

## 2.3 Surgery for Moderate Obesity - Previous Meta Analyses

### Introduction

Systematic reviews and Meta analyses are some of the highest levels of scientific evidence that we have to base clinical practice. Outcomes from a meta-analysis may include a more precise estimate of the effect of treatment or risk factor for disease, or other outcomes, than any individual study contributing to the pooled analysis[29].

These analyses have some inherent biases such as selection bias, however the information provided is of the utmost importance in answering clinical questions. Combining several studies to perform review and if applicable, meta analysis has several clear advantages. Single studies may have unique characteristics that may limit the generalizability and combining studies can also increase the sample size and produce more precise estimates of the effect size. Many clinicians rarely have the time or resources to critically appraise the literature relevant to a particular clinical question and most systematic reviews will focus on a narrow, clearly defined topic that will allow the inclusion of all relevant articles[29]. Here we review some of the previous systematic reviews that have been performed relevant to our topic. While this review is not comprehensive, we feel that the chosen studies are representative of previous systematic reviews and meta analyses performed on patients with BMI < 35 undergoing bariatric surgery.

**Literature Review**

**Fried et. al 2010[25]**

Integrated review of patients operated with mean BMI < 35. Inclusion criteria were any form of bariatric/metabolic surgery in human subjects where mean BMI < 35. Primary outcome measure were data on treatment of type 2 diabetes mellitus. Major databases were employed such as MEDLINE, current contents, science citation index and Cochrane library. Total of 16 studies were included with total of 343 patients undergoing 8 different surgical procedures. Procedures were categorized as either restrictive, malabsorptive, or combination of restrictive/malabsorptive. 66% of patients were female and mean age at baseline was 46.2 years. Study design ranged from case report of 2 patients in larger prospective study of 50 patients, to retrospective case series and matched controlled trials. 11 of the studies were prospectively performed. 4 studies revealed statistically significant changes in BMI, 6 for fasting plasma glucose, and 6 for hemoglobin A1c. Overall the mortality in the entire analysis was 0.29%. Complication rate was low at 4% while one study reported an overall rate of major complications of 10.3%. The authors found that all categories of bariatric procedures were effective, by varying definitions of resolution, in reducing and in most cases resolving type 2 diabetes mellitus. The authors also noted that subset of patients with BMI 25.0-29.9 had lower percentage reduction in clinical and laboratory measures of T2DM compared to patients with BMI 30-35. Limitations of the review were the small numbers of patients in each of the studies and the fact that most of the studies were observational in nature. There were also a large proportion of retrospective studies included in this review. Reporting of

measures of T2DM was also highly varied with incomplete data in many studies. The authors conclude that bariatric surgery can be a safe and efficacious treatment for T2DM in non-morbidly obese population. They concede that the level of evidence based on these studies is weak and that their data represents preliminary evidence and that further research will be required to institute change in the inclusion criteria for patients undergoing bariatric and metabolic surgery.

**Meijer et. al. 2011[30]**

Systematic review looking specifically at reversal of T2DM in patients undergoing LRYGB or AGB. Typical PubMed search was performed using usual MESH terms. Other databases were not included in this review. Results with respect to improvement in T2DM and reversal of T2DM were not clearly defined. One study reported on patients requiring insulin and patients who were able to completely discontinue their insulin were said to have reversal of T2DM and others that were able to decrease their daily insulin use were said to have improvement in T2DM. Studies included patients with BMI < 35. Total of 9 studies were included in the analysis. All but 2 studies were retrospective case series with one RCT comparing AGB to medical therapy and one observational prospective study comparing AGB and LRYGB and as well vertical banded gastroplasty, a purely restrictive procedure that has been largely abandoned because of high complication and failure rates[6]. Reversal rates of T2DM reported in the included studies ranged from 43% of subjects to as high as 87%. Improvement was seen anywhere from 91-100% of subjects. More dramatic rates of resolution and improvement of T2DM were seen with LRYGB compared to other

surgical groups. Authors also noted improvement in hypertension in up to 80% of patients. Improvements in hypercholesterolemia, hypertriglyceridemia, cardiac function and obstructive sleep apnea were also noted to occur however less frequently. Limitations of the study were small numbers of patients and retrospective nature of the included studies. The lack of standardization of reporting, especially with respect to T2DM endpoints also creates heterogeneity in the data. This makes further analysis impossible beyond a descriptive analysis and in turn will make the results less generalizable. With respect to current review of patients with BMI < 35, these patients represented a lower proportion of the total number of patients studied. The authors conclude that bariatric surgery, especially LRYGB, leads to reversal of, or improvement in T2DM and should be considered in patients with poorly controlled T2DM and a BMI greater than 35. They also conclude that similar results have been shown in patients with BMI < 35 however this remains controversial and would require further study.

**Adegbola et. al. 2013[31]**

Systematic review of patients with BMI < 35 undergoing AGB. Articles published after 1990 and in the English literature were reviewed. Patients undergoing AGB with BMI < 35 were included. All procedures were performed laparoscopically unless otherwise specified. Medline and Embase databases were searched using standard MeSH terms. Appropriate reference lists and bibliographies of selected articles were also searched for relevant articles. Patient characteristics including BMI, comorbidities, duration of follow up and endpoints including complications were included. 6 studies ultimately met the inclusion criteria. Of these, 5 were retrospective and 1 RCT was

included. Weight loss was reported in 5 of the 6 studies. 85% of patients were followed up with mean percentage excess weight loss (%EWL) ranging from 52.5+/-13.2 to 78.6+/-9.4 and 57.6+/-29.3 to 87.2+/-9.5 at one and two years postoperatively respectively. At 5 years post operatively 72.4% of patients were followed up appropriately with %EWL of 71.9+/-10.7. With respect to obesity related comorbidities the data that was collected was somewhat heterogeneous. One study reported 89.1% of patients free of comorbidities at 1 year of follow up. Another study reported resolution of comorbidities in "most" patients. Resolution of metabolic syndrome was reported in 93.3% of patients in another study. One study did not report on resolution of comorbidities. Mortality ranged from 0-1% while other complications including wound infection, band slippage/migration, band erosion and port leaks were reported anywhere from 0-5.2%.

Limitations of this study include small number of studies included, short to medium term follow up, and exclusion of non-English language publications. The heterogeneous and vague nature of which obesity related comorbidities are evaluated also makes any firm conclusion in this regard difficult to make. The authors conclude that from the limited data available LAGB is well tolerated and effective with good short-term outcomes in obese patients with BMI < 35kg/m2. They also conclude that there appears to be a favorable impact on obesity related comorbidities however there remains a paucity of data on this group of patients and long term outcomes need to be further evaluated.

**Rao et. al. 2015[32]**

Systematic review and meta analysis of patients undergoing laparoscopic roux en y gastric bypass (LRYGB). Other surgery types were excluded from this review. Inclusion criteria were patients who underwent LRYGB and had BMI < 35kg/m2 and had T2DM. Embase, Ovid, PubMed, Cochrane and China national knowledge infrastructure databases were searched. Relevant journals including obesity surgery and surgical endoscopy were also reviewed. Where datasets were incomplete study authors were contacted however no further information was obtained. Total of 9 studies met the inclusion criteria and were included in the review. All studies included patients undergoing LRYGB however one study also included patients who underwent mini gastric bypass, a variation on LRYGB. The authors felt this was appropriate given likely similar mechanism of action of the procedure. Follow up ranged from 1-7years. 6 of the studies included were prospective. Complete remission of T2DM was defined as HbA1c < 6. Partial remission of T2DM was defined as HbA1c between 6-7 and improvement was defined as HbA1c > 7. BMI was significantly lower in the postoperative group ($p < 0.00001$, WMD -7.42, 95% CI -8.87 to -5.97). Values for glucose and HbA1c were also lower postoperatively (glucose $p < 0.00001$, WMD -59.87, 95% CI -67.74 to -52.01; HbA1c $p < 0.00001$, WMD -2.76, 95% CI -3.41 to -2.11). Funnel plot was symmetrical indicating no significant bias in this analysis. No deaths were reported in any of the trials with complication rates varying from 6.7-25.9%. The authors conclude that LREYGB can offer a substantial improvement and cure of T2DM for many patients although in this analysis not all patients achieved a complete remission (HbA1c < 6). They offer that

further clinical studies with larger sample sizes and longer follow up will help elucidate this issue.

**Muller-Stitch et. al. 2015[33]**

Systematic review and meta analysis of patients undergoing multiple surgical types in population with BMI < 35kg/$m^2$. Medline, Embase and Cochrane databases were searched using key word algorithm. Studies evaluating metabolic surgery effect on T2DM in patients with BMI < 35kg/$m2^2$. Studies were also included if mean BMI of treatment group was < 40kg/$m^2$. Abstracts were reviewed and full text review was performed for studies meeting the inclusion criteria. Further cross referencing was carried out for all included studies. T2DM remission was primary endpoint. This was defined as achievement of HbA1c of < 7 and fasting plasma glucose of < 7.2mmol/L and discontinuation of diabetic medications. It is noteworthy that studies included in this analysis had different cutoffs to define remission of T2DM. The study authors state that this was accounted for in their analysis. Secondary outcome measures were HbA1c levels, BMI, presence of hypertension and dyslipidemia. Total of 13 studies met the inclusion criteria and were included in the analysis. There were 7 randomized controlled trials and 6 observational cohort trials included comprising a total of 818 diabetic patients. Follow up ranged from 12-36 months. Remission rates of T2DM were significantly higher in the surgery group compared to medical treatment (OR 14.1, 95% CI 6.7-29.9, P < 0.001). Glycemic control was significantly better in the surgical group (OR 8.0; 95% CI 4.2-15.2, P < 0.001). Using trim and fill method to adjust for potential publication bias this effect the effect on glycemic control remained strong (OR 7.2; 95%

CI 5.0-1-.3, P < 0.001). HbA1C values were lower in the surgical group (MD -1.4% CI -1.8% to -0.9%, P < 0.001). Similar results were seen for body mass index, arterial hypertension and dyslipidemia (MD -5.5kg/m2, 95% CI -6.7 to -4.3kg/m2, P < 0.001; OR 0.25, 95% CI 0.12 – 0.50, P < 0.001; OR 0.21, 95% CI 0.10 – 0.44, P < 0.001). Moderate heterogeneity was seen with respect to hypertension with $I^2$ = 64%, P =0.010. Sensitivity network meta analyses were performed to compare treatment effects across different surgical procedures. AGB, biliopancreatic diversion (BPD), LSG, and LRYGB were all proven to be significantly effective for T2DM remission. No significant heterogeneity was observed. Similar result was seen for glycemic control although LSG failed to reach level of significance. Analysis of HbA1c levels was compromised by relevant heterogeneity (Q = 72.2, $I^2$ = 86%, P < 0.001). Serum HbA1c levels were significantly decreased after BPD, LSG, LRYGB but not by AGB. This study was one of the first to explore T2DM control in non-severely obese patients undergoing surgery (BMI < 35). Limitations included the heterogeneity in which T2DM remission was defined as well as other outcome parameters. This study clearly demonstrated the superior effect of surgery compared to medical therapy on T2DM and other obesity related comorbidities in patients with BMI < 35. Because of a paucity of studies looking at this patient population the authors included studies where patients with BMI between 35 and 40 were included which also make any conclusions drawn from this analysis somewhat questionable and limits the generalizability. The authors concede this point however their conclusion was that surgery should be offered to non severely obese patients with T2DM and other elements of the metabolic syndrome. They also conclude that longer follow up and well

designed RCTs with standardized definitions of T2DM remission and glycemic control will be helpful to further explore surgery in this patient population.

**Summary**

Bariatric and metabolic surgery for moderately obese patients is a relatively new area of investigation however the early results seem very promising. Much of the data had been extrapolated from morbidly obese subjects however more studies are being completed and the relative effectiveness of surgery for these patients is becoming clearer. Previous studies and meta analyses have shown very strong results however there are some limitations within these studies. Firstly, most studies to date have been retrospective in nature that brings inherent limitations such as a high risk of bias due to use of inappropriate control groups[34]. Lack of prospective and randomized data leaves the validity of these results in question. Secondly, most studies look primarily at T2DM and glucose metabolism as this seems to be highly prevalent not only in morbidly obese patients but also patients who are overweight or moderately obese. The effectiveness of bariatric and metabolic surgery on T2DM in this patient population seems to be well established however the treatment of other obesity related comorbidities such as hypertension and hypercholesterolemia seems to be less well defined. Minimal information regarding perioperative complications and lack of longterm follow up are also limitations of previous studies. Obesity related comorbidities are common even in moderately obese patients and the effectiveness of surgery on these factors remains largely in question. The results of previous meta analyses provide valuable information with regards to surgery in patients with moderate obesity however many studies which

were included in these analyses included patients with BMI > 35. This limits the generalizability of the results to moderately obese patients. Another important factor is that previous meta analyses have not included any randomized trials comparing surgical groups. As we have seen from the literature there are many bariatric and metabolic procedures that have been, and are being performed for obesity. Without any direct comparison it is extremely difficult to determine which procedure, if any will have better efficacy and safety for morbidly obese subjects. Meta-analysis is a quantitative, formal, epidemiological study design used to systematically assess previous research studies to derive conclusions about that body of research. Based on the key points highlighted above we felt that a meta analysis using the highest level of evidence available in patients with BMI < 35 would be necessary to fully answer our research question and to achieve our objectives which are: (1) To determine the efficacy and safety of bariatric surgery in patients with a BMI < 35 $kg/m^2$ (2) To determine whether one surgical procedure will be superior in this patient population. (3) To fully explore the effect of surgery on less well studied obesity related comorbidities such as sleep apnea and osteoarthritis in patients with BMI < 35.

# Chapter 3 – Methods

## 3.1 – PRISMA Statement

This systematic review and meta analysis was constructed according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses guidelines (PRISMA). These guidelines consist of a 27 point checklist that includes items deemed essential for transparent reporting of a systematic review. These guidelines were

developed in 2005 by a group of experts to ensure transparent reporting and expanded on previous guidelines developed in 1996 that led to the development of the Quality Of Reporting Of Meta-analyses statement (QUORUM) that came from the realization that key information was often omitted in systematic reviews diminishing their usefulness[35].

## 3.2 Protocol

Prior to beginning the study a detailed protocol was developed by the initial research team including members from Memorial University and The University of Ottawa (see co authorship statement). The protocol included primary area of interest, and how the data would be collected and analyzed. Once the protocol was finalized it was submitted to the Memorial University Department of Research and Graduate Studies as part of the final application. The details of the protocol will be discussed in the following sections. The protocol was not published or registered prior to commencement of the study.

## 3.3 PICOS

**P** – Patients with BMI < 35 undergoing bariatric surgery.

**I** - Either Roux-en-Y gastric bypass, sleeve gastrectomy, or adjustable gastric band.

**C** - Medical therapy or in case of comparison between surgical groups LREYGB would be considered experimental group.

**O** - Primary outcome was fasting plasma glucose (FPG). Secondary outcomes were obesity related comorbidities, quality of life, mortality.

**S** - Prospective studies including cohort studies and randomized controlled trials.

**3.4 Eligibility criteria**

We only included studies that were prospective in nature including randomized controlled trials, prospective cohort studies and case control studies where data were collected prospectively. Presence of a control group was not a prerequisite for inclusion. This has not been done in previous systematic reviews and meta analyses. Studies from English language literature were included along with studies that were specific to patients with BMI < 35. We also did not place any limits on language and felt that any relevant studies that were not printed in English would be translated and included in our study. While there are many procedures described for obesity and related comorbidities, we included only data on patients undergoing laparoscopic surgery, as this has become the standard of care for many abdominal procedures. We also felt that to create generalizability we would limit our review to patients who underwent adjustable gastric band procedure, gastric bypass, and sleeve gastrectomy. These three procedures are by far the most commonly performed procedures throughout the world and have proven safety profiles and have been shown to be efficacious. We excluded other procedures such as biliopancreatic diversion and duodenal switch as in some series these procedures have been shown to be associated with 2% risk of mortality, which we felt was unacceptably high to be translated to a patient population of moderately obese individuals until further data becomes available[5]. This decision was made on the basis of the current state of bariatric surgery throughout the world and expert opinion from our committee members.

Studies were excluded if they were retrospective in nature or if they included any patients who had BMI > 35. We also excluded studies where data was insufficient to complete any meaningful analysis. The analysis was limited to adult patients as the

implementation of bariatric surgery in the pediatric population is controversial and adds another confounding variable, which we felt would skew the results.

Studies were assessed for level of bias and this is detailed in section 3.10. Non randomized studies that were assessed critical level of bias and randomized studies that were assessed high level of bias were excluded from the analysis.

**3.5 Information sources**

Electronic databases were searched using our pre specified protocol. Main databases included PubMed, EMBASE, Cochrane, and CINAHL. Relevant abstracts were collected and reviewed to ensure all relevant studies were captured. The search was started on may 1, 2012 and the final search was performed on December 1, 2017. PubMed is one of the most comprehensive databases throughout the world and we felt that this would be one of the most important databases to include. EMBASE has a special focus on drugs and pharmacology and we felt this would be especially relevant given that bariatric surgery is often compared to medical therapy especially with respect to T2DM. Cochrane database of systematic reviews and meta analyses is generally accepted as one of the most comprehensive databases of systematic reviews and was therefore included in the search. CINAHL was included to include all relevant allied health studies.

**3.6 – Search**

We felt it was important to encompass all relevant studies and as such started with a very broad search. Of course we wanted to capture all relevant articles without having an excess of extraneous citations to review which would make the process much more

difficult and time consuming. Our main focus was patients with BMI < 35 so part of our search strategy was to ensure that all abbreviations and written forms were captured. For example we used the search terms "bmi < 35", "body mass index < 35", "body mass index less than 35", etc. This strategy was also applied to other aspects of our search such as for gastric bypass where many abbreviations and variations on how the procedure is labeled are used in the literature. We used this strategy for each of the databases that were searched. For a complete list of search terms see appendix 5. We wanted to evaluate whether our search strategy would be too broad or too narrow and to do this we used our search strategy first in MEDLINE and reviewed the first 100 citations that were listed. These were reviewed and it was felt that approximately 20% of these citations would be potentially relevant to our topic of interest. Based on expert opinion within our group (JM) it was felt that this search strategy would be sufficiently broad to capture all relevant citations while sufficiently narrow such that the amount of time to review these citations would not be excessive.

### 3.7 – Study Selection

All relevant citations were reviewed in an unblinded standardized manner by two independent reviewers (CS and DT). Any articles in question were reviewed by an independent third reviewer (PC). To ensure transparency amongst reviewers data on each of the relevant citations were extracted and placed in an excel spreadsheet so that it could be examined by all of the reviewers. All reviewers were physicians associated with Eastern Health and Memorial University. Any disagreement was resolved by consensus between all three reviewers. In general we felt that inter-rater agreement was excellent.

Ultimately there were 92 articles that were felt to be potentially relevant and upon initial review there was only disagreement regarding 4 of these citations (4.3%).

### 3.8 – Data Collection Process

As a guide we used the Cochrane Consumers and Communication Review Group's data extraction template to ensure all relevant data were extracted appropriately[36]. For ease of data extraction and translation to statistical software, initial data extracted was input into a Microsoft Excel file, which could then be easily modified as appropriate. Only citations where complete consensus between all three reviewers was reached were reviewed in detail. Data extraction and compilation was done by two primary researchers (CS and DT) and where disagreements occurred it was planned that a third author (PC) would review the data and consensus would be used to resolve such disagreements. Where data sets were felt to be incomplete we contacted study authors directly via email to inquire as to whether any further data could be supplied for the analysis. Of the articles included in the study, four of these articles were felt to have incomplete data. Of the four study authors who were contacted three of these responded promptly however unfortunately they were unable to provide us with any further data or clarification. We were careful to exclude any duplicate reports from the data collection process.

### 3.9 – Data Items

Information was extracted from each included study on (1) characteristics of study participants (including, surgery type, duration of follow up and number of participants); (2) Primary outcome measure FPG (which was converted to mmol/l when this variable

was given in different units); (3) Secondary outcome measures HbA1c, body weight, BMI, waist circumference, hypertension, dyslipidemia, and other obesity related comorbidities including osteoarthritis, obstructive sleep apnea, infertility, gastroesophageal reflux and venous stasis. For categorical variables data was converted to standard units. Where continuous variables were analyzed, we collected mean and standard deviation for study groups and control groups and these were identified appropriately. For cases where standard deviation was not given we used the standard formula using range divided by four to provide an accurate estimate of standard deviation to ensure that our statistical analysis could be complete.

### 3.10 - Risk of Bias In Individual Studies

As per Cochrane and PRISMA guidelines a risk of bias assessment was performed for each study[37]. For non-randomized studies we used the Risk of Bias in Non Randomized Studies of Interventions (ROBINS-I) tool. Studies were analyzed and graded based on a number of variables including: confounding variables, selection of study participants, classification of interventions, deviations from intended interventions, missing data, measurement of outcomes, and selection of the reported result. Level of bias assigned overall was based on the most serious levels of bias that was given using the signaling questions. Once each individual study was graded it was then placed into one of the following four categories:

1. **Low risk of bias –** The study is comparable to a well designed randomized controlled trial.

2. **Moderate risk of bias –** The study is significant but not comparable to well designed RCT.
3. **Serious risk of bias –** The study has some significant problems.
4. **Critical risk of bias –** The study is too problematic and should not be included in the analysis.

For non randomized studies this tool was applied and any studies deemed to have critical risk of bias were excluded.

For randomized controlled trials we used the Risk of Bias (ROB) 2.0 tool for individually randomized parallel group trials. The two scoring systems are similar, however with the ROB 2.0 tool signaling questions could also be answered with 'not included' if there were insufficient data in the study to sufficiently answer the questions. Once the studies were graded they were then assigned to one of the following levels of risk of bias:

1. **Low risk of bias –** The study is judged to be at low risk of bias for all domains.
2. **Some concerns –** The study is judged to be at some concerns in at least one domain.
3. **High risk of bias –** The study is judged to be at high risk of bias in at least one domain or to have some concerns for multiple domains.

For randomized controlled trials that were being considered for the study, those with high risk of bias judgments were excluded. If there was no information available for either randomized or non randomized studies on which to make a judgment these were also excluded.

## 3.11 – Summary Measures

Meta analysis was planned to be performed using weighted mean difference for continuous outcomes with 95% confidence interval and $p < 0.05$. Means were weighted according to inverse variance method. Negative mean difference was indicative of positive associative effect with experimental variable. For categorical variables we used the Cochrane –maentel- hantzel method with 95% confidence interval to calculate the odds ratio (OR). Odds ratio less than 1 was indicative of association of individual variable with intervention in question. Unless otherwise specified a p value of $< 0.05$ was considered significant. Odds ratios are often reported in the literature as a measure of association between exposure and outcome. This statistical method does have some disadvantages. For example, in the event that the outcome measure is common in a research study (>10% when compared to the control group) the OR may overestimate the risk ratio or relative risk, which may be more intuitively understood as a measure of association. Risk ratios (RR) can only be calculated for cohort studies, which makes this statistic less versatile[38]. In our study we included results from cohort studies, randomized controlled trials and case control studies, which meant OR was a more practical statistic to analyze the data. This would ensure that the analysis of our results would be more homogenous.

## 3.12 – Synthesis of Results

Statistical analysis and graphical representation was completed using R version 3.3.1 statistical software using random effects model. We used forest plots to graphically display the overall results. Subgroup analysis was performed by looking at each surgery

type separately and by comparing randomized data between surgery groups. We used $I^2$ statistic as a measure of heterogeneity. This method is generally considered superior to Cochrane Q statistic, which has been shown to be poor at detecting true heterogeneity among studies as significant[39] .

Given that we are exploring surgical interventions and due to the nature of the studies we chose random effects model for the meta analysis. As opposed to fixed effects model, this technique relies on the assumption that treatment effect across studies will be variable[38]. With differences in surgical technique, etc. we felt that this would be most appropriate for our analysis.

**3.13 – Risk of Bias Within Studies**

Risk of bias assessment was completed using the risk of bias in non randomized studies instrument (ROBINS-I) as per PRISMA and Cochrane guidelines. Similarly for RCTs we used the risk of bias tool (ROB 2.0) for individually randomized parallel group trials again as per Cochrane and PRISMA. For each of the signaling questions each study was given an assessment of bias for that question and ultimately a final judgment regarding bias that was essentially the highest level of bias assigned during each of the signaling question assessments. Kakoulidis et. al[40] was excluded at this stage given it had been assigned critical level of bias as outlined previously. Data from this study was excluded from meta analysis and included only for descriptive purposes.

## 3.14 – Risk of Bias Across Studies

We used funnel plots to represent publication bias. Funnel plot is a simple scatter plot of the intervention effect estimates of individual studies against some measure of each studies size or precision. Generally the effect estimates will be plotted on the horizontal scale and the measure of study size on the vertical axis. The precision of the estimated intervention effect increases as the size of the study increases. Effect estimates from small studies will therefore scatter more widely at the bottom of the graph with the spread narrowing among larger studies. In the absence of bias the plot should approximately resemble a symmetrical inverted funnel. The presence of bias will lead to an asymmetrical appearance of the funnel plot. The more pronounced the asymmetry the more likely it is that the amount of bias will be substantial. The statistical power of a study is determined by factors in addition to sample size such as the number of participants experiencing the event for dichotomous outcomes and the standard deviation of responses for continuous outcomes. In other words there are more factors beyond strictly sample size that will determine the power of an individual study. Because of this it is usually recommended to use the standard error of the intervention effect estimate rather than sample size to create the funnel plot. Here we used the effect on FPG as this was our primary outcome variable and because the majority of the studies included reported this information in an accurate fashion[37].

## 3.15 – Additional Analysis

We wanted to assess whether bariatric surgery would be effective for moderately obese individuals. This would include assessment of weight loss or change in BMI and

also whether it would be effective in resolution of obesity related comorbidities in this patient population. We also wanted to assess whether the treatment effect would be consistent between different procedure types which lead to further sub group analysis based on our literature review. An important question was also whether one surgical procedure would be superior to another. On this basis we aimed to answer the following questions:

1) Is bariatric or metabolic surgery safe and effective for weight loss and resolution of obesity related comorbidities in patients with BMI < 35?

(2) Are the effects consistent between surgical procedures?

(3) Is there one surgical procedure that is superior in this patient population?

# Chapter 4 – Results

## 4.1 – Study Selection

Review of Cochrane, Medline, Embase and CINAHL databases yielded total of 1966 potentially relevant citations. All of the potentially relevant articles returned were from English language literature. To ensure that our search strategy was appropriate we reviewed 100 of these citations to see what percentage were potentially relevant. Once these were reviewed we found that 20% of these references were potentially valid. Based on expert opinion (JM) we felt that this would be appropriate and indicative that our search strategy was sufficiently broad to capture all relevant articles but also focused enough as to minimize the amount of time spent reviewing non relevant articles. Once we removed duplicates and screened for articles that were clearly not relevant based on review of the abstracts we were left with 546 potentially relevant articles. These were

further screened which left us with 92 articles for full text review and application of inclusion/exclusion criteria. 8 studies were excluded because of retrospective nature of study design. 22 were excluded because patients with BMI > 35 were also included. 13 were excluded because outcome measures of interest were not reported. Remaining articles were excluded for various reasons including: lack of control group, systematic review, alternative surgical procedure or variation included, etc. There were total of 4 articles on which the reviewers disagreed regarding inclusion/exclusion. A third independent reviewer reviewed these and consensus was reached on whether they should be included. This left us with a total of 14 articles to be included in the final systematic review and meta analysis. During the assessment of bias one study was assessed critical risk of bias and was therefore excluded from the analysis. Kakoulidis et. al was ultimately excluded because of lack of control for confounding variables, inconsistent reporting of data and because ultimately this was a subgroup analysis of a larger study. This left 13 articles to be included in the final synthesis (see figure 4.1 - PRISMA diagram).

## 4.2 – Study Characteristics

Total of 505 patients were included in the analysis. All studies were published in English language literature. There were total of 4RCTs and 9 prospective cohort studies. Of the 4 randomized controlled trials, 2 compared LSG to LRYGB. One compared LSG to medical therapy and one compared AGB to medical therapy. Of the 9 cohort studies 6 involved patients undergoing LRYGB. One study looked at patients undergoing LSG and 2 involved patients who received AGB. Control group for all included prospective studies was pre-surgical group. To date this is the only body of work utilizing prospective data

and limited to patients with BMI < 35. Follow up ranged from 1 to 5 years. All included studies were single centered. 10 studies gave complete data on our primary outcome measure of FPG, 9 studies included data on HbA1c. With respect to glucose metabolism remaining studies reported change in medications and resolution of diabetes, etc. 11 studies gave data on BMI. 6 studies gave accurate information regarding hypertension and 4 reported on waist circumference. 5 studies gave data regarding dyslipidemia and body weight. Some studies reported on other variables including: osteoarthritis, infertility, insulin homeostasis, venous stasis, and urinary stress incontinence however the data was too heterogeneous and inconsistent to provide any meaningful analysis. See tables section for complete summaries of included studies and table 22 for overall summary of included studies.

## 4.3 – Results of Individual Studies and Synthesis of Results

### Introduction

Here we summarize the main results of the study. First we will look at measures of glucose metabolism namely fasting plasma glucose and hemoglobin A1c. Then we will look at other measures of obesity including body weight, body mass index, and waist circumference. We will then look at obesity related comorbidities including hypertension and hyperlipidemia. Lastly we will summarize data from other obesity related comorbidities, which were not included in the meta analysis due to the heterogeneous nature of the data. These will include obstructive sleep apnea, osteoarthritis, infertility, gastroesophageal reflux, urinary stress incontinence and venous stasis. Where possible we looked at results from each individual procedure type and also compared results

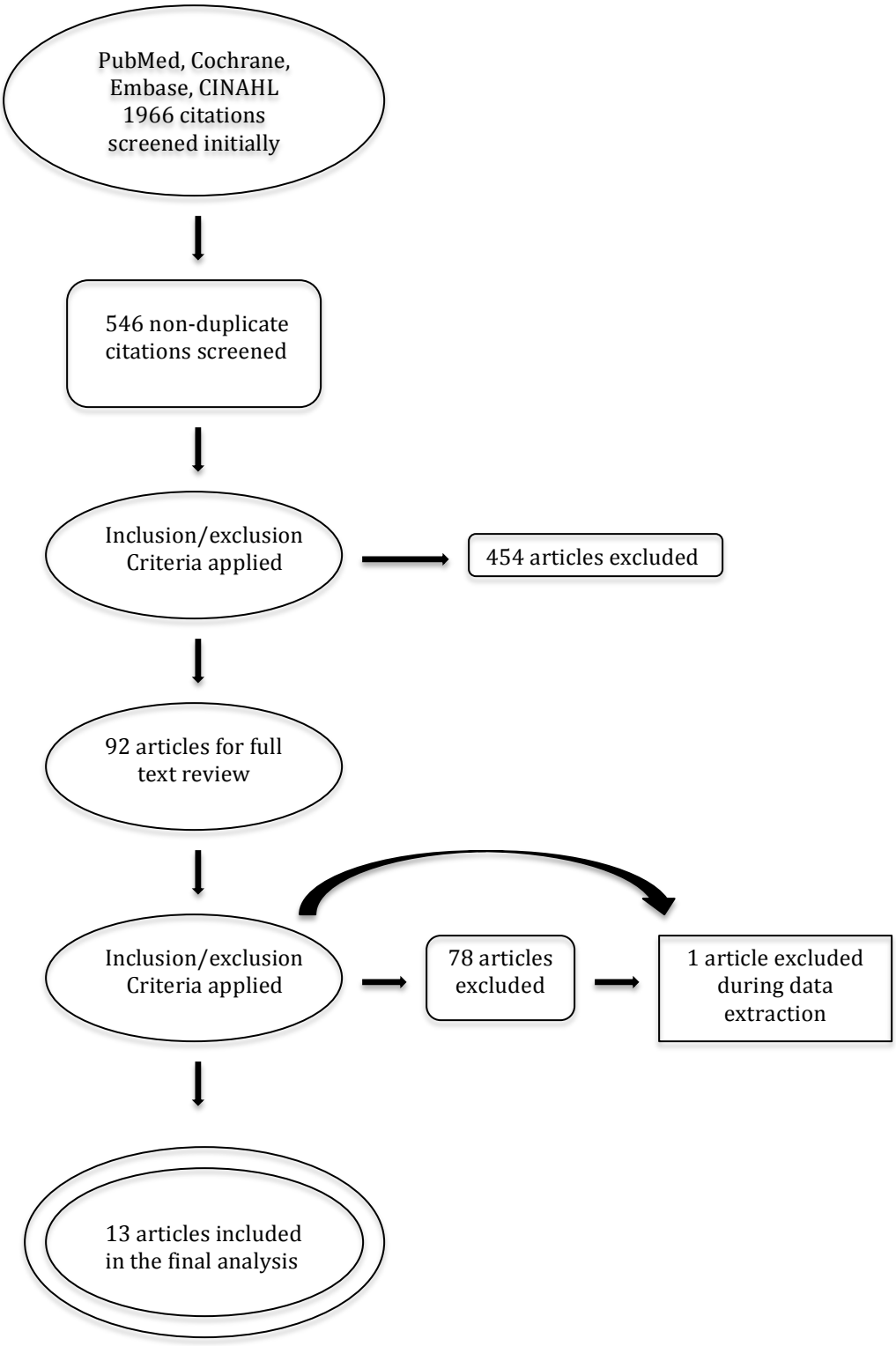


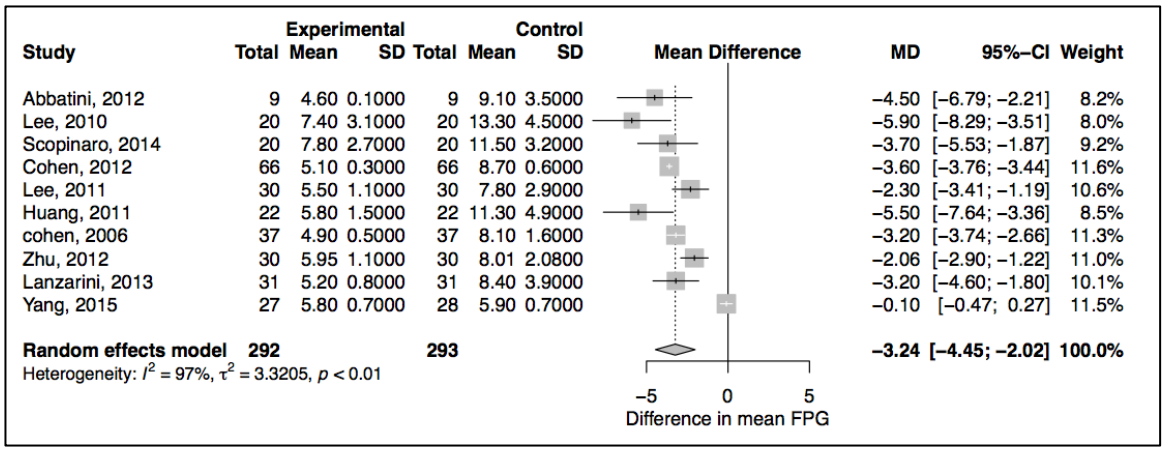

Figure 4.1 – PRISMA Diagram

between individual procedures to determine if treatment effects were true for each surgery type and to help answer the question: is one surgery superior for the treatment of obesity and comorbidities in patients with BMI < 35? For cohort studies the experimental group was assigned as the surgical group and the control group was assigned as the pre surgical group. For RCTs comparing REYGB to LSG gastric bypass was assigned as the experimental group and LSG was assigned as the control group. For RCTs comparing surgery to medical therapy surgery was considered the experimental group. 95% CIs to the left of 0 for continuous variables and 1 for categorical variables was considered significant in favor of the surgical or experimental group.

**Fasting Plasma Glucose**

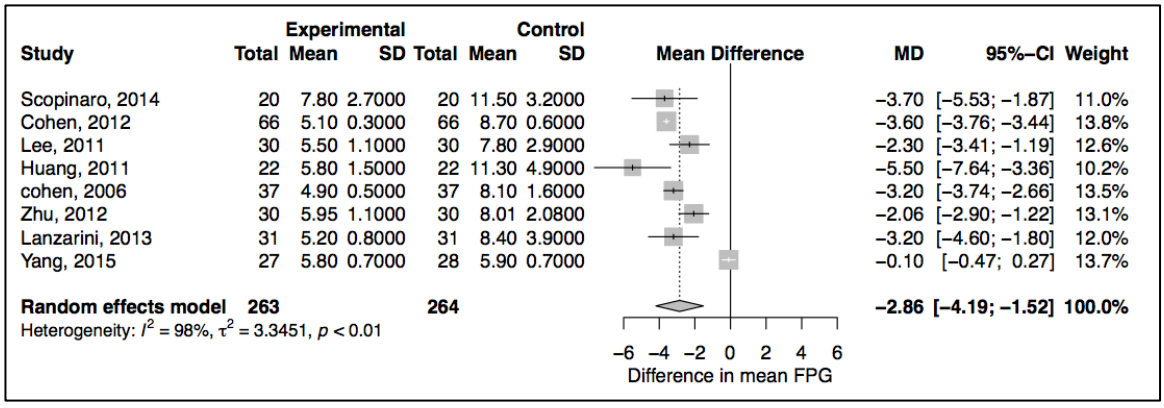
Results of Fasting plasma glucose data are summarized in figure 4.2. Total of 10 studies included well defined results of glucose metabolism that could be used for meta analysis. Remaining 3 studies reported data in terms of improvement of hypoglycemic medications or other non standardized outcomes. All data was reported in units of mmol/l. Surgical groups had a total of 292 patients versus 293 in the control groups. According to predefined analysis using weighted mean difference and random effects model there was a significant improvement in glycemic control in the Surgical group compared to the control group (WMD -3.24, 95% CI -4.45; -2.02). Heterogeneity as defined by $I^2$ statistic was high at 97% ($p < 0.01$).

Figure 4.2 – Fasting plasma glucose overall results




The effect seen on FPG in the pooled data held true in the Roux en y gastric bypass group. A total of 263 patients were in the surgery group compared to 264 in the control group. Weighted mean difference was -2.86 with 95% CI -4.19; -1.52. Heterogeneity was high with $I^2$ = 98% (p<0.01).

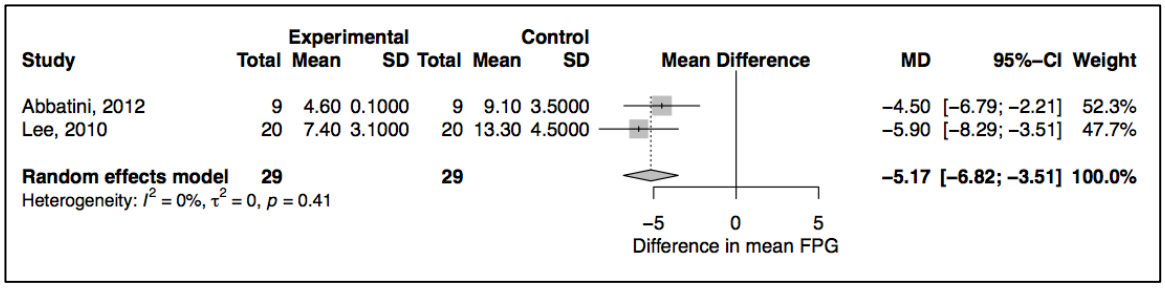
Figure 4.3 – Fasting plasma glucose roux en y gastric bypass




Total of 2 studies compared sleeve gastrectomy to medical therapy with respect to FPG. 29 patients were in each of the surgical and medical groups. Statistically significant

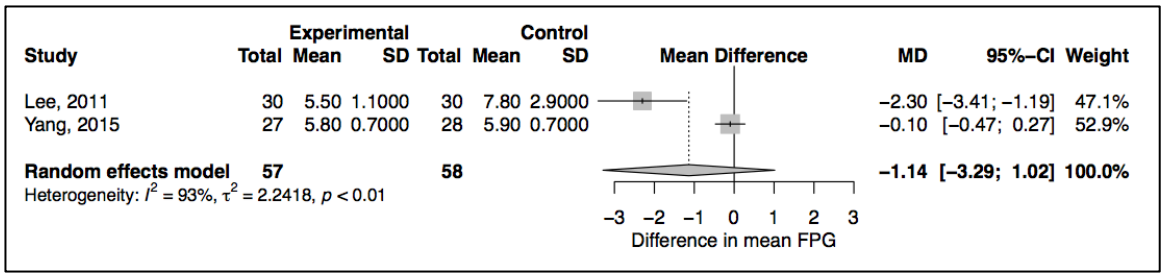
difference in FPG between the groups was noted (WMD -5.17, 95% CI -6.82; -3.51). Heterogeneity was low with $I^2 = 0\%$ ($p = 0.41$). Insufficient data were obtained to perform subgroup analysis on patients who underwent adjustable gastric banding however these were included in the overall analysis.

Figure 4.4 – Fasting plasma glucose sleeve gastrectomy




Only two surgical groups were available for direct comparison by RCT. 2 studies compared roux en y gastric bypass to sleeve gastrectomy. Roux en y gastric bypass group was considered the experimental group and sleeve gastrectomy was taken as the control group. Total of 57 patients were in the gastric bypass group versus 58 patients in the sleeve gastrectomy group. No statistically significant difference was seen between the two surgical groups with respect to change in FPG (WMD -1.14, 95% CI -3.29; 1.02). Heterogeneity was high with $I^2 = 93\%$ ($p < 0.01$).

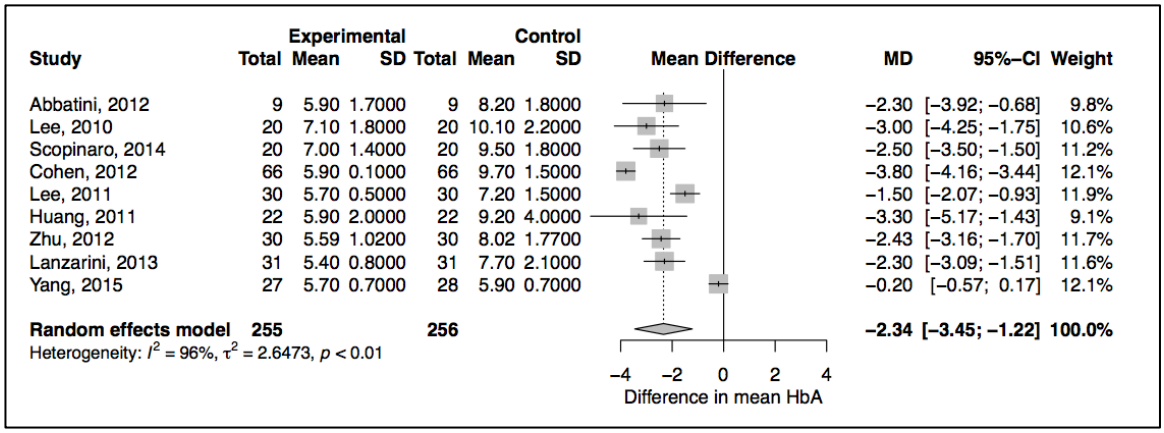
Figure 4.5 – Fasting plasma glucose REYGB versus LSG




## Hemoglobin A1c

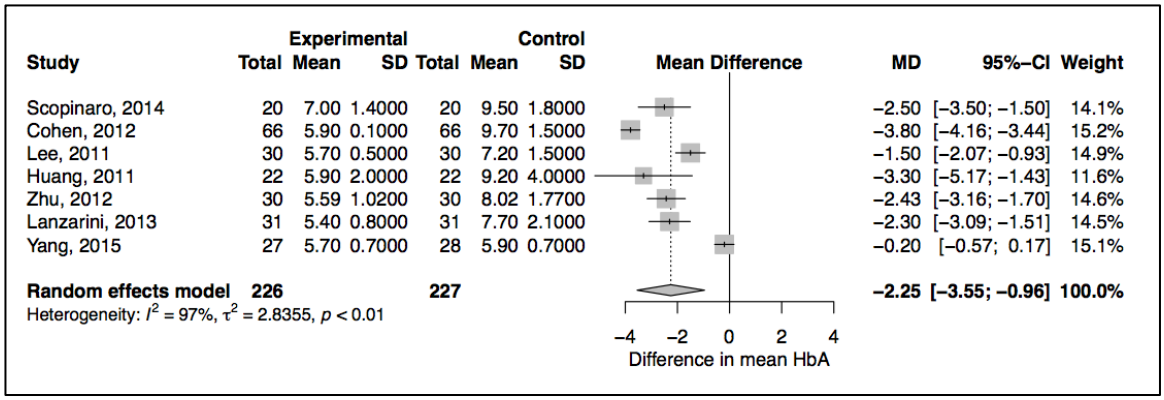
Total of 9 studies included data where hemoglobin A1c was well defined. Others gave vague descriptions such as 'improved' without meaningful descriptions and were excluded here. Weighted mean difference and 95% confidence intervals were obtained for the overall group looking at all surgical interventions versus control groups. The overall results show that there is a statistically significant improvement in hemoglobin A1c in the surgical groups versus control (WMD -2.34, 95% CI -3.45; -1.22). Heterogeneity as calculated using $I^2$ statistic was high at 96% ($p < 0.01$).

Figure 4.6 – HbA1c overall results



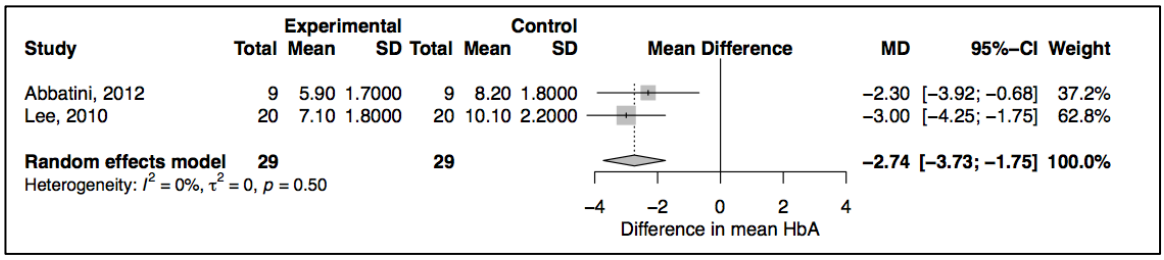
7 of the 9 included studies involved gastric bypass. Total of 226 patients were included in the experimental group versus 227 in the control group. A statistically significant improvement was seen in the REYGB group compared to control group (WMD -2.25, 95% CI -3.55; -0.96). Heterogeneity was high with $I^2 = 97\%$ ($p<0.01$).

Figure 4.7 – HbA1c Roux en Y gastric bypass



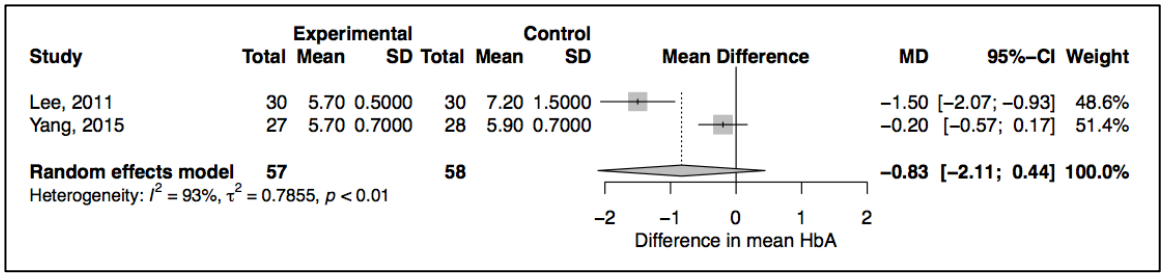

2 studies compared sleeve gastrectomy to medical therapy. Total of 29 patients were included in each group. Statistically significant improvement in HbA1c was maintained across these studies (WMD -2.74, 95% CI -3.73; -1.75). Data was insufficient to perform a meaningful comparison amongst patients who underwent AGB.

Figure 4.8 – HbA1c Sleeve gastrectomy group



Comparison of REYGB versus LSG yielded 2 RCTs with a total of 57 patients in the gastric bypass group and 58 in the LSG group. No statistical difference was seen with respect to HbA1c between the two surgical procedures (WMD -0.83, 95%CI -2.11; 0.44). Heterogeneity was high with $I^2$ =93%.

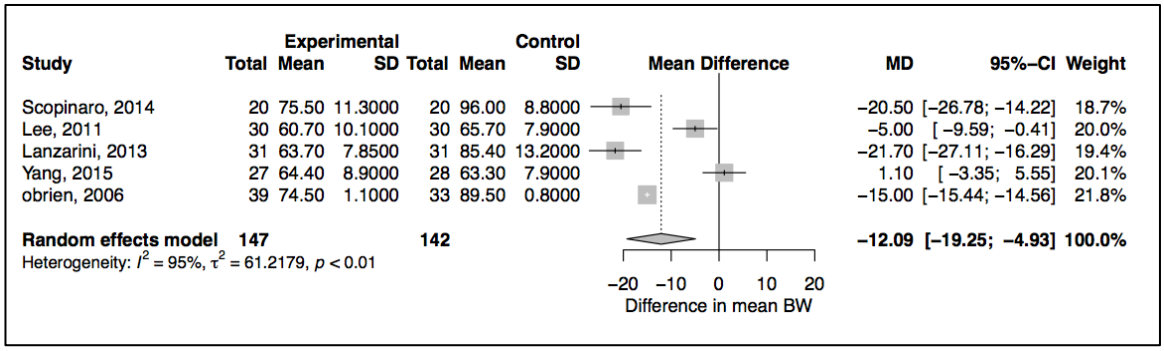
Figure 4.9 – HbA1c REYGB versus LSG




## Body Weight

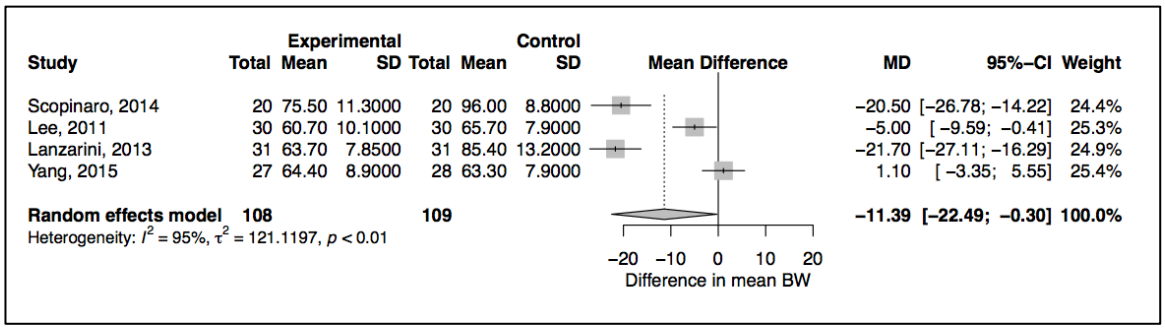
Total of 5 Studies included body weight measured in kilograms. Total of 147 patients were included in the treatment group compared to 142 studies in the control group. Results demonstrated a significant decrease in body weight favoring the treatment group (WMD -12.09, 95% CI -19.25; -4.93). Heterogeneity was calculated as high with $I^2$ = 95% ($p < 0.01$).

Figure 4.10 – Body Weight Overall Results



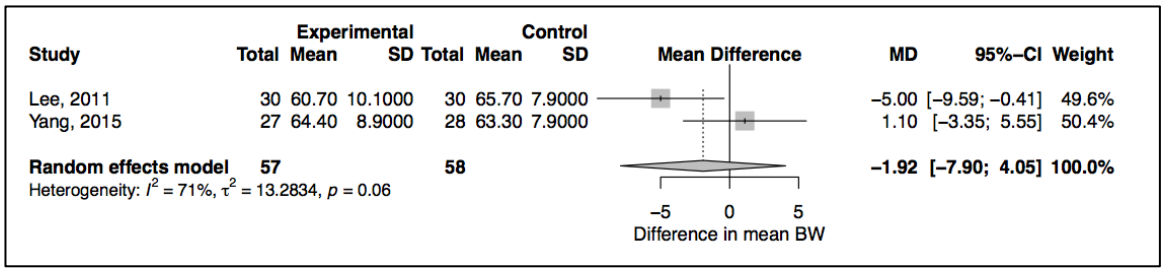
4 out of 5 studies that included accurate body weight analysis were done on patients undergoing laparoscopic roux en y gastric bypass. One study included patients who underwent adjustable gastric banding. Analysis of this group revealed that there was a statistically significant difference compared to the control group (WMD -11.39, 95% CI -22.49; -0.30). Heterogeneity was high with $I^2 = 95\%$ ($p < 0.01$). No further subgroup analysis was possible given the paucity of data reported in the included studies.

Figure 4.11 – Body weight REYGB




Two RCTS included data on body weight. Both compared REYGB to LSG. Total of 57 patients in the experimental group versus 58 in the control group. No statistically significant difference was seen with respect to body weight amongst the study participants (WMD -1.92, 95% CI -7.90; 4.05). Heterogeneity was high with $I^2 = 71\%$ however this result was not statistically significant ($p = 0.06$).

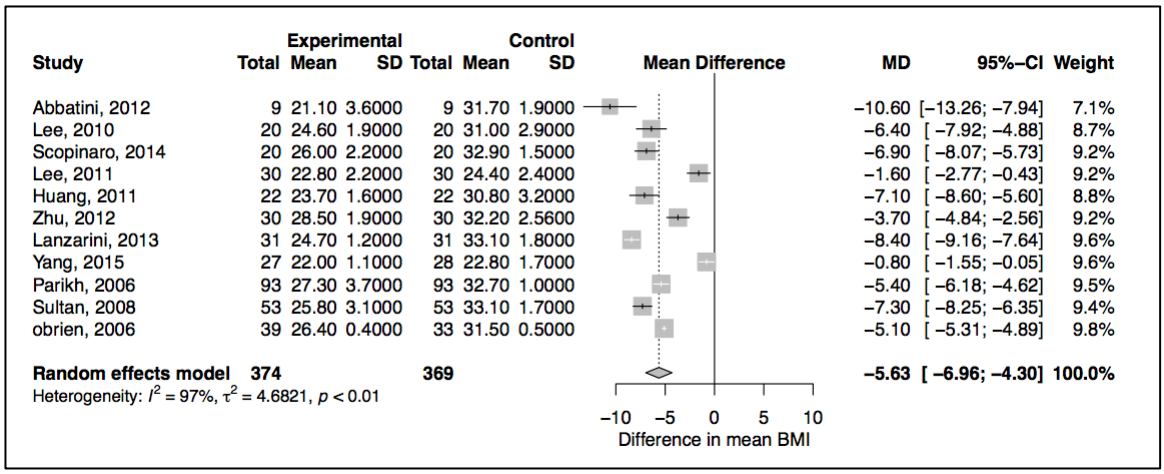
Figure 4.12 – Body weight REYGB versus LSG




## Body Mass Index

A total of 11 studies included data on body mass index. There were a total of 374 patients in the treatment group versus 369 in the control group. Results demonstrate a statistically significant decrease in BMI for surgery compared to the control group (WMD -5.63, 95% CI -6.96; -4.30). Heterogeneity was high with $I^2$ calculated at 97% ($p < 0.01$).

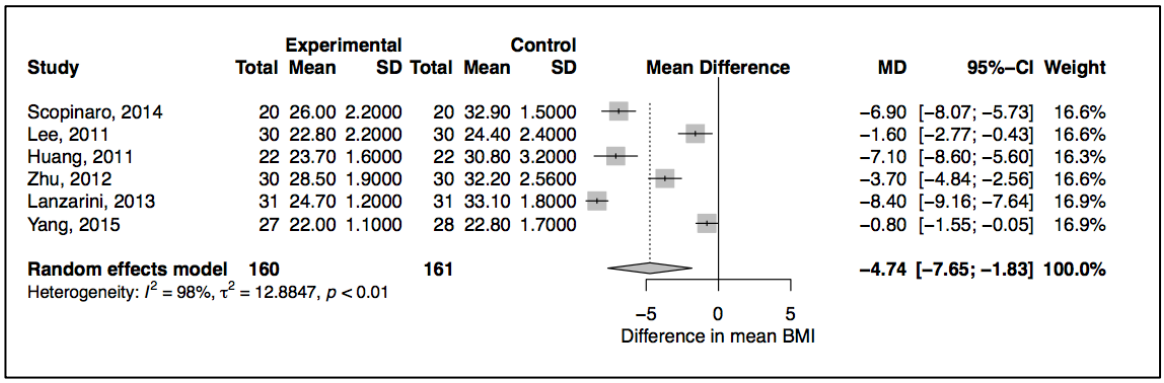
Figure 4.13 – Body mass index overall results




6 studies included roux en y gastric bypass in the analysis. There were 160 patients in the treatment group versus 161 in the control group. The results demonstrate a

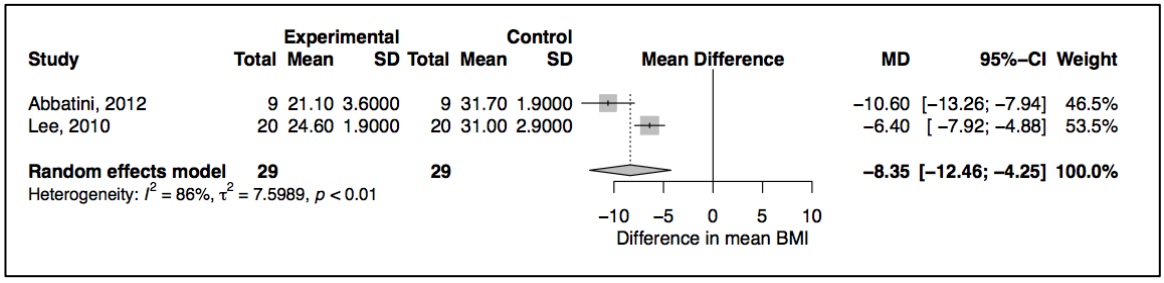
significant decrease in BMI for patients undergoing REYGB (WMD -4.74, 95% CI -7.65; -1.83). Heterogeneity was high with $I^2 = 98\%$ ($p < 0.01$).

Figure 4.14 – Body mass index REYGB



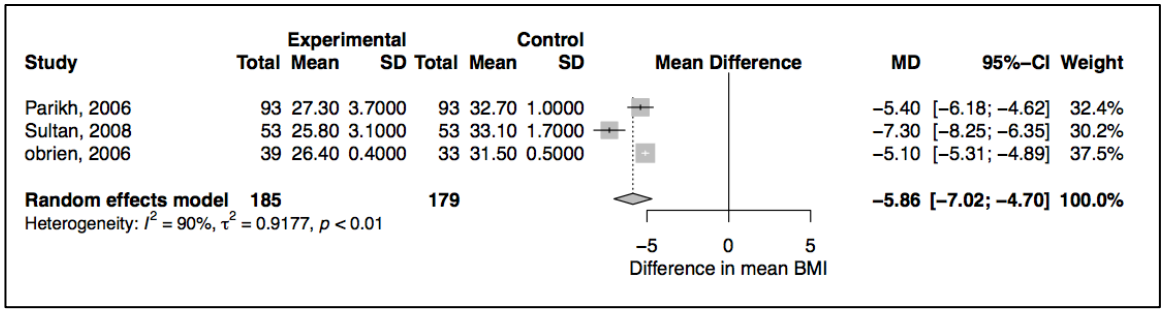

2 studies in patients undergoing laparoscopic sleeve gastrectomy included data on body mass index. There was a total of 29 patients in each of the experimental and control groups. A statistically significant decrease in BMI was seen in these patients (WMD -8.35, 95% CI -12.46; -4.25). Heterogeneity was high with $I^2 = 86\%$ ($p<0.01$).

Figure 4.15 – Body mas index LSG



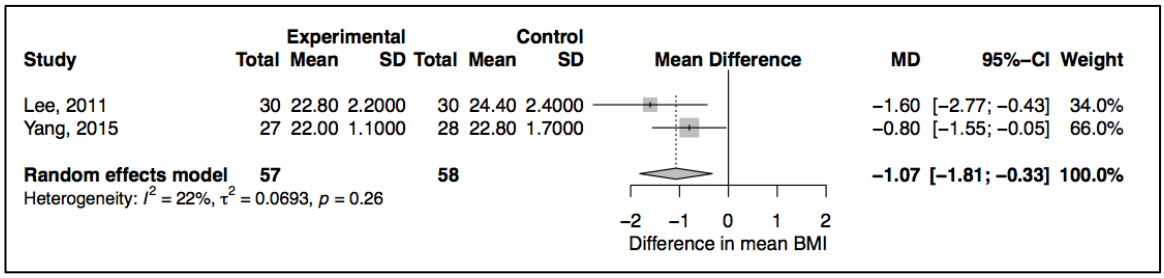
3 studies in patients undergoing adjustable gastric banding reported data on body mass index and were included here. There were 185 patients in the surgical group versus 179 patients in the medical therapy group. Overall a statistically significant difference between the two groups was seen (WMD -5.86, 95% CI -7.02; -4.70). Heterogeneity was high as $I^2$ statistic was calculated at 90% ($p< 0.01$).

Figure 4.16 – Body mass index AGB




Lee et. al 2011 and Yang et. al compared Laparoscopic Roux en y gastric bypass to sleeve gastrectomy and data on BMI are included here. Total of 57 patients in the gastric bypass group versus 58 in the sleeve gastrectomy group. Overall there was a statistically significant difference in BMI between the two surgical groups favoring gastric bypass (WMD -1.07, 95% CI -1.81; -0.33). Heterogeneity was low with $I^2$ = 22% however this result was not statistically significant (p = 0.26).

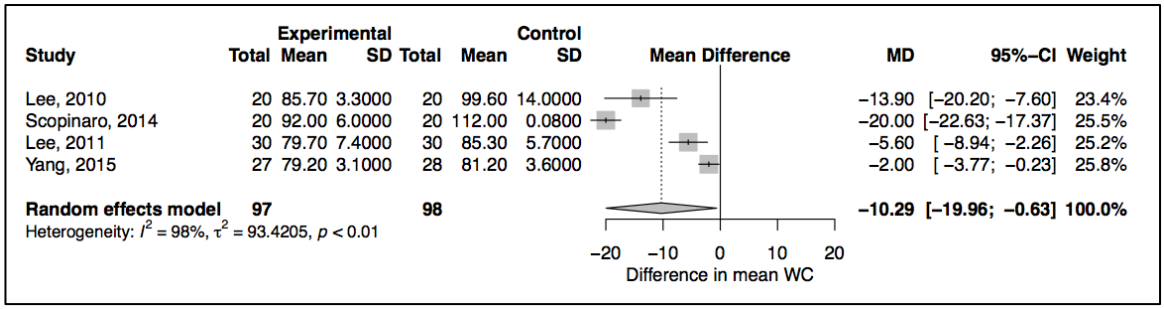
Figure 4.17 – Body mass index RYGB versus LSG



**Waist Circumference**

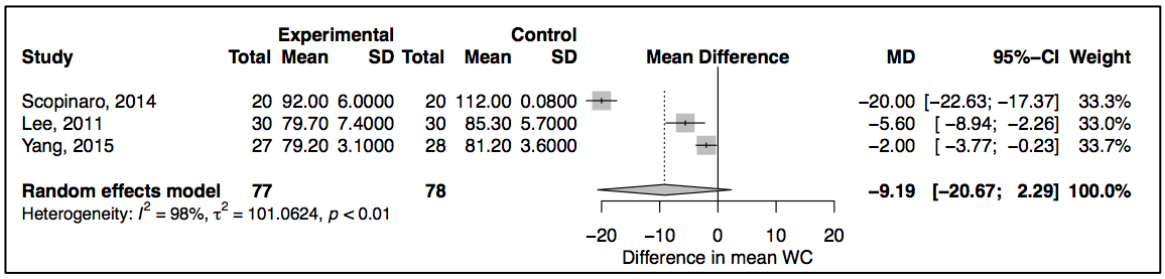
A total of 4 studies included data on waist circumference. 97 patients were in the experimental group with 98 patients in the control group. Overall the results were in favor of the experimental group with a significant decrease in waist circumference (WMD - 10.29, 95% CI -19.96; -0.63). Heterogeneity was high with $I^2 = 98\%$ ($p < 0.01$).

Figure 4.18 – Waist circumference overall results



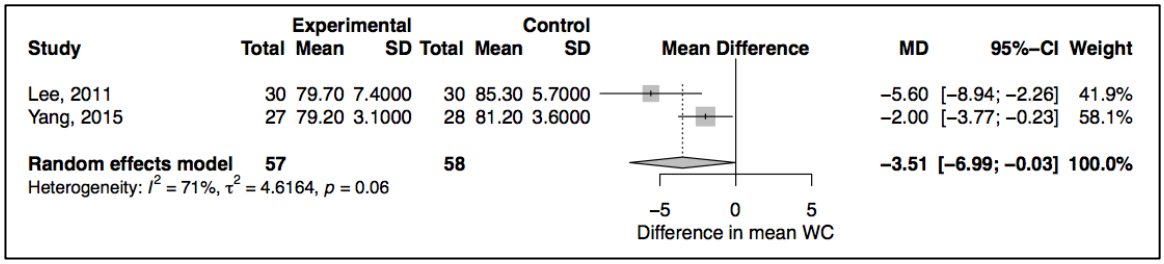

3 out of 4 studies that included data on waist circumference were done on patients undergoing roux en y gastric bypass and these are included here. 77 patients were included in the treatment group while 78 were included in the control group. Interestingly the results show that there was no significant difference in the experimental group versus the control group (WMD -9.19, 95% CI -20.67; 2.29). Heterogeneity was high with $I^2 = 98\%$ ($p < 0.01$).

Figure 4.19 – Waist circumference RYGB




2 RCTS compared RYGB to LSG and data on waist circumference are presented here. Total of 57 patients in the RYGB group versus 58 in the LSG group. Overall a statistically significant decrease in waist circumference was seen in favor of gastric bypass (WMD -3.51, -6.99; -0.03). Heterogeneity was high with $I^2$ = 71%, however this result was not significant (p = 0.06).

Figure 4.20 – Waist circumference RYGB versus LSG

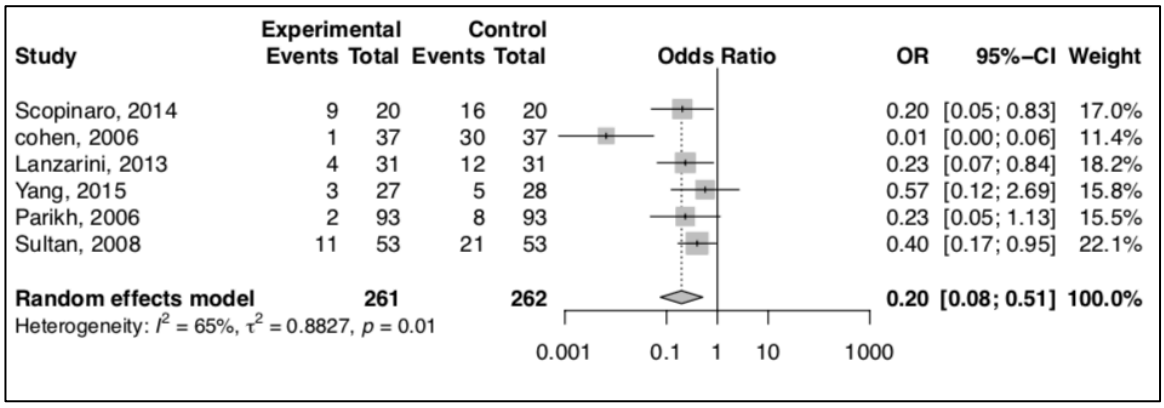



## Hypertension

We used rate of resolution of hypertension as a measure of success of surgery versus control. Blood pressure measurements were given in mmHg unless otherwise specified. Resolution was defined as normalized blood pressure and off antihypertensive medications. Odds ratios were calculated as per Cochrane – mantel – hantzel method. A

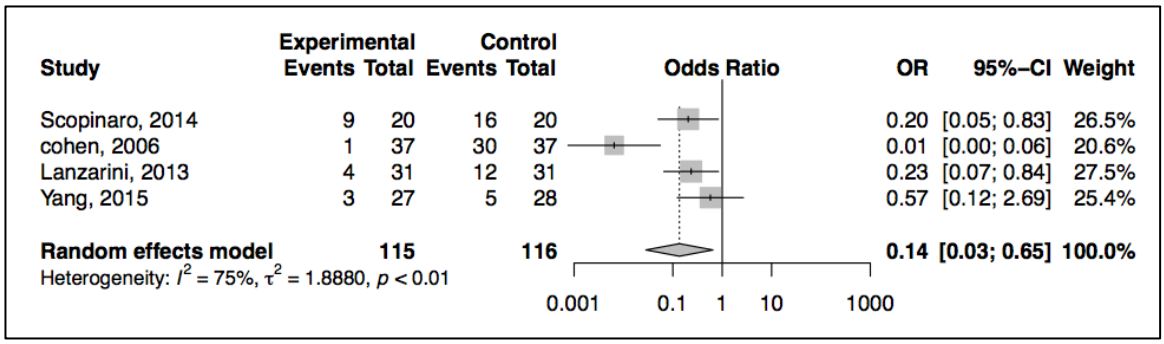
total of 6 studies included accurate data on hypertension and rates of resolution. There were 261 patients in the experimental group and 262 in the control group. According to pre defined statistical analysis there was a significantly higher rate of resolution of hypertension in the experimental group (OR 0.20, 95% CI 0.08; 0.51). Heterogeneity was high with $I^2 = 65\%$ ($p = 0.01$). No randomized controlled trials were available for comparison here as insufficient data was reported in these studies with respect to hypertension.

Figure 4.21 – Resolution of Hypertension Overall Results



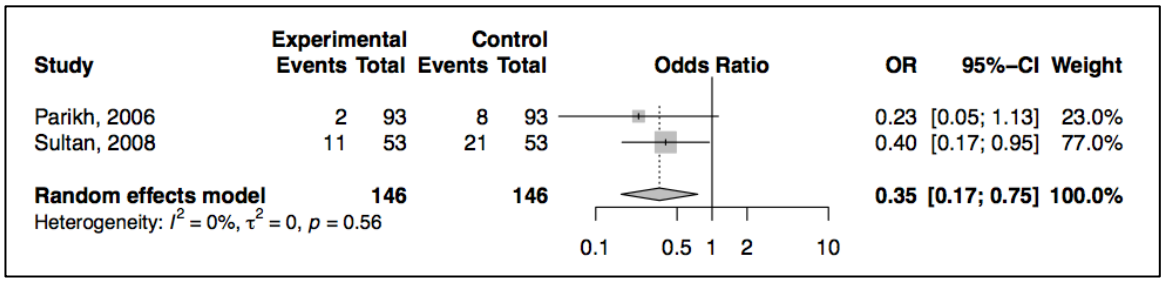

4 studies involving patients who underwent roux en y gastric bypass included data on resolution of hypertension. There were a total of 115 patients in the experimental group and 116 in the control group. Overall there was a statistically significant rate of improvement in the surgical group compared to control (OR 0.14, 95% CI 0.03; 0.65). Heterogeneity was high with $I^2 = 75\%$ ($p < 0.01$).

Figure 4.22 – Resolution of Hypertension RYGB




2 remaining studies involved patients undergoing adjustable gastric banding. There were a total of 146 patients in the surgical group and 146 in the group who underwent medical therapy. Overall there was a significant difference in the rate of resolution of hypertension favoring surgery (OR 0.35, 95% CI 0.17; 0.75). Heterogeneity was low in this comparison with $I^2$ = 0%, however this result was not significant (p = 0.56).

Figure 4.23 – Resolution of Hypertension AGB

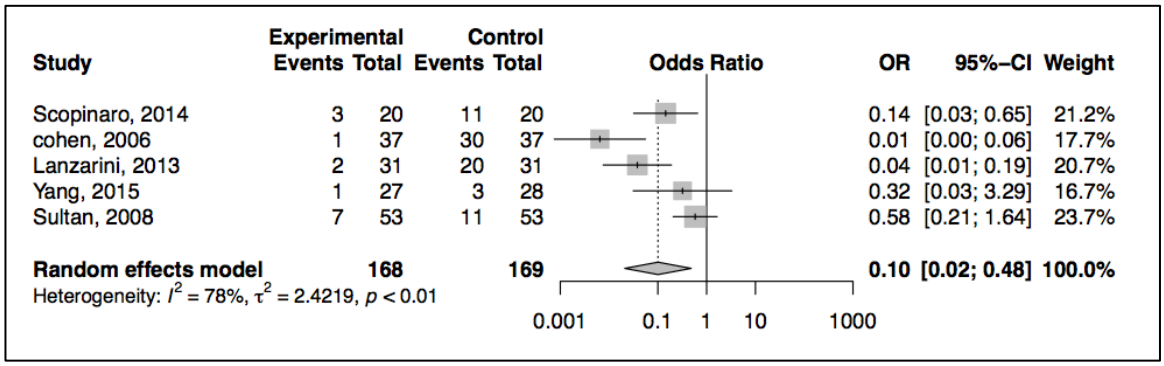



## Dyslipidemia

We looked at rates of resolution of dyslipidemia that was defined as normal cholesterol and triglyceride panel and discontinuation of any lipid lowering medications. Absolute values of triglycerides, total cholesterol, low density lipoprotein (LDL) and high

density lipoprotein (HDL) were sometimes provided but not in all studies. Normal values for lipid parameters were also inconsistently reported. A total of 5 studies included complete data on resolution of dyslipidemia. All of these studies were done on patients who underwent roux en y gastric bypass therefore no other surgical groups were available for sub group analysis. A total of 168 patients were in the experimental group and 169 were in the control group. Overall the results were in favor of surgery with a statistically significant difference in the 2 groups (OR 0.10, 95% CI 0.02; 0.48). Heterogeneity was high with $I^2 = 78\%$ ($p < 0.01$).

Figure 4.24 – Resolution of Dyslipidemia Overall Results




## Other Obesity related comorbidities

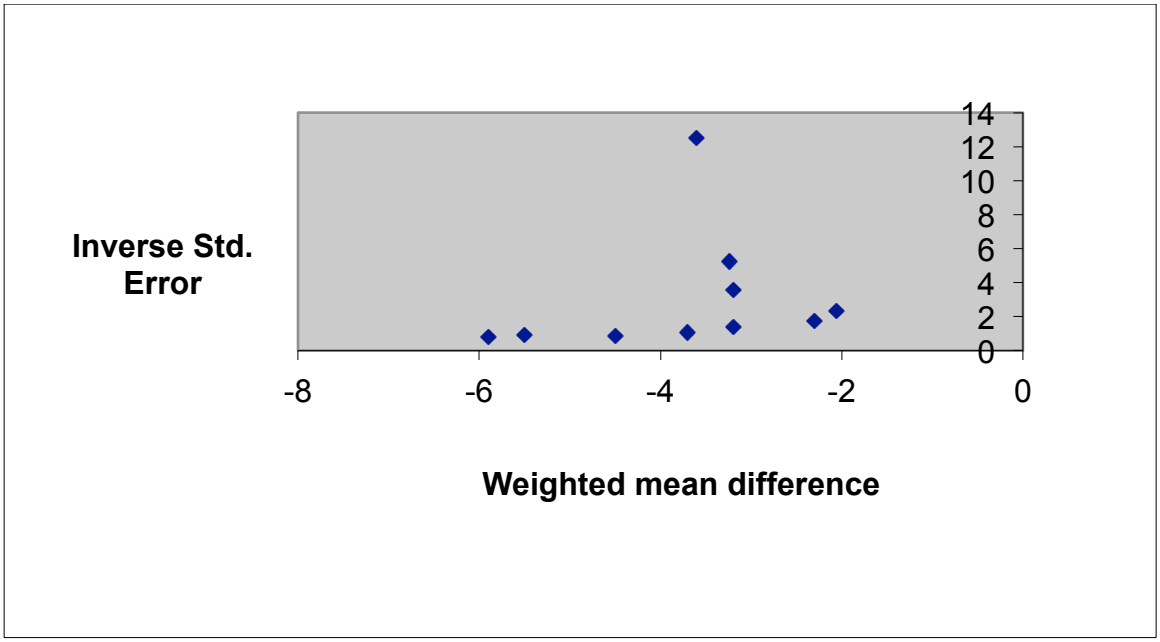
Some authors studied other variables related to obesity in this patient population however there was insufficient data across studies to perform a meta analysis. Obstructive sleep apnea was seen to be resolved in anywhere from 7%-100% of patients. Osteoarthritis was improved in 33-47% of patients. Infertility has been shown to be associated with obesity and the included studies here reported a range of 7-50% improvement in fertility after surgery. Gastroesophageal reflux was improved in 7-75%

of patients. Urinary stress incontinence improved in 20% and venous stasis improved in 7%. These results indicate that although there was a wide range of results across studies and that bariatric surgery in moderately obese patients can be an effective treatment for a number of less commonly explored obesity related comorbidities.

## 4.5 - Risk of Bias Across Studies

Significant heterogeneity was seen across included studies with $I^2$ values ranging from 78% to 98%. For several of the variables studied there was slight to moderate heterogeneity reported with $I^2$ values ranging from 0% to 71% however for these analyses the calculated value of $I^2$ was not statistically significant. A funnel plot was constructed to assess publication bias and this revealed a fairly symmetrical graph, especially given the relatively low number of studies included in the analysis. Several studies did cause the graph to skew somewhat to the left that may indicate that studies with smaller numbers of patients may have had a slightly exaggerated treatment effect. Based on this visual representation we conclude that there is a slight possibility of bias in the presented studies.

Figure 4.25 – Funnel Plot (FPG Surgery versus control)




## 4.6 - Safety profile of surgery

In order to make a truly strong recommendation regarding surgery for moderately obese individuals, one would have to ensure that the safety profile is comparable to that seen in morbidly obese subjects. It would be very difficult to make a recommendation if the safety profile was less favorable than seen with morbidly obese patients. Encinosa et. al. examined insurance claims for almost 10,000 patients undergoing bariatric surgery in the United States between 2001 and 2006 and found that the risk adjusted rates of readmissions from complications during bariatric surgery was 6.8%[41]. With increasing procedures being performed laparoscopically the mortality rate has declined to 0.1%[17]. Most studies included in this meta analysis had small numbers of patients however there were no mortalities reported in any of the studies included here. Yang et. al. reported no

major complications or deaths in their group of patients undergoing LREYGB and LSG. There were minor complications in 5.5% of patients including GERD and minor anemia not requiring blood transfusion[42]. Parikh et. al. reported no perioperative complications in their patients undergoing AGB. All patients had their surgery performed laparoscopically and were discharged home from hospital within 24 hours[43]. Cohen et. al. 2012 reported no major surgical complications or mortality in their cohort of patients undergoing LREYGB[9]. These results were comparable across studies and certainly comparable to that seen in morbidly obese subjects undergoing bariatric surgery, in some cases the safety profile was more favorable compared to morbidly obese subjects. The way in which complications is reported amongst studies is quite heterogeneous making a pooled estimate of complications of limited value.

## Chapter 5 – Discussion

To our knowledge this is the first systematic review and meta analysis to exclusively explore patients with BMI < 35. We also limited our analysis to the highest levels of evidence available, that being randomized controlled trials and prospective cohort studies. We considered commonly measured variables including FPG, HbA1c, body weight, BMI, waist circumference and resolution of hypertension and dyslipidemia. We also considered less commonly studied variables including obstructive sleep apnea, osteoarthritis, infertility, gastroesophageal reflux, urinary stress incontinence and venous stasis.

Our results indicate that for patients with moderate obesity (BMI < 35), surgery appears to be superior to medical therapy in terms of weight loss and treatment and

resolution of obesity related comorbidities. Specifically, we found that surgery is superior in terms of treating diabetes mellitus as fasting plasma glucose and HbA1c were significantly lower in the surgical group compared to medical therapy. Other sub group analysis confirmed that surgery is superior to medical therapy in terms of weight loss, BMI loss, decreased waist circumference, and resolution of hypertension and dyslipidemia. This result was mostly consistent across each of the surgical procedures that were reviewed, namely roux en y gastric bypass, adjustable gastric banding and sleeve gastrectomy. This finding was also consistent between cohort studies and RCTs that compared surgery to medical therapy.

There is also much debate in the literature regarding which if any bariatric procedure is superior and this remains controversial particularly in the setting of morbidly obese patients. The question of superiority amongst bariatric surgical procedures is even more crucial when considering their application to a patient population where its use is even more controversial. We were able to make a direct comparison in this meta analysis between roux en y gastric bypass and laparoscopic sleeve gastrectomy as there were 2 randomized controlled trials that met our inclusion criteria and both randomized patients to either RYGB or LSG. We found that there was no significant statistical difference in the effect on FPG or HbA1c between either procedure meaning that both of these surgical procedures will be equally effective in the treatment of diabetes mellitus in patients with BMI < 35. The confidence intervals in this part of the analysis were very wide which certainly is a limitation in interpreting these results. We did however find that there was a significant decrease in waist circumference and BMI with RYGB compared to LSG. There was insufficient data to determine if either procedure would be superior at treating

hypertension, dyslipidemia or other obesity related comorbidities. An interesting finding was the fact that despite improved BMI and waist circumference with RYGB, there did not seem to be any difference in body weight between the two surgical procedures. Yang et. al. demonstrated a significant difference in the starting body weight between the two surgical groups. Patients who underwent RYGB had a starting body weight of 94.3kg compared to 88.4kg in the sleeve gastrectomy group. Therefore it stands to reason that the LSG group could plausibly have a lower weight after surgery compared to the RYGB group. This systematic flaw could have been circumvented by calculating change in body weight and making that comparison however the data was too heterogeneous to allow this data to be collected and analyzed accurately for each of the included studies. For example, in some studies only the post surgical body weight was given making the change in body weight impossible to calculate. We attempted to get this information from study authors however we were unable to obtain any further data.

It is also a curious result that BMI in the RYGB group was significantly lower given that there was no significant difference in body weight. Given that the pre surgical BMI was not significantly different (32.3 +/- 2.4 RYGB versus 31.8 +/- 3.0 LSG) one can deduce that the heights of the study participants was substantially different amongst both groups, however this data was not included in any of the studies. In other words absolute weight loss was higher in the RYGB group however the starting weights were different so that the absolute difference in BMI was greater in the RYGB without any significant difference in the post surgical weights between the two groups.

A similar discrepancy was seen with respect to waist circumference. Comparison of all studies revealed a significant reduction in waist circumference favoring the surgical

group. There was also a significant difference when RYGB was compared to LSG favoring the bypass group. However when RYGB was compared on its own during subgroup analysis no significant difference was seen. The data seen in Scopinaro et. al., which was included in this subgroup analysis, appears to be an outlier and also appears to skew the data into non significance. On this basis one can conclude that RYGB is more effective than either LSG or medical therapy in reducing waist circumference. The end result is that based on our analysis RYGB appears to have better results with respect to BMI, waist circumference and possibly weight loss when compared to LSG. There was no difference in parameters of glucose metabolism (FPG and HbA1c).

Despite using the highest level of evidence available this study did have several limitations. There was significant heterogeneity across studies included in this analysis. Although several studies did have $I^2$ values of zero, which would indicate essentially no heterogeneity, these findings were not statistically significant. Majority of studies had $I^2$ values >90% which would indicate a high level of heterogeneity. More accurate analysis would potentially be performed using studies with less heterogeneity, however we were limited in this regard by the relatively small number of studies included in the analysis. There are a number of factors which could explain high levels of heterogeneity between studies including: sample size, inclusion criteria of individual studies and confounding variables such as variations in surgical technique. Excluding studies with high levels of heterogeneity would have left us with insufficient number of studies to perform meaningful analysis. Data extraction was also somewhat heterogeneous across the studies which potentially led to this phenomenon. We also grouped randomized and non randomized data during the analysis. This approach does have some inherent statistical

limitations which no doubt contributed in part to the anomalies that were seen in some of the results with respect to body weight and waist circumference. We felt this approach would be most appropriate for this study given that we were comparing several different obesity related comorbidities and surgical procedures and also given that we have relatively few studies for comparison. We considered performing subgroup analysis of only randomized controlled trials however we felt that comparing 2 studies comparing surgical procedures and 2 studies looking at surgery versus medical therapy would lead to significant heterogeneity given that comparison of 2 RCTs comparing surgical groups already had a significant level of heterogeneity. We felt that this further subgroup analysis would not lend any meaningful results to our analysis. There was also heterogeneity with respect to risk of bias. 7 of the included studies were assessed moderate risk of bias, 3 studies were found to have some concerns, 2 studies had serious concerns and one study had low risk of bias. This was also heterogeneous amongst cohort studies and RCTs. 6 of the 7 studies assessed moderate level of bias were cohort studies assessing RYGB. Both studies assessed serious level of bias were cohort studies looking at AGB. As we had analyzed the data looking at each surgery type we felt this subgroup analysis would not be meaningful. Similarly because all 3 RCTs that were found to have some concerns with respect to bias were analyzing different surgery types we felt this subgroup analysis would not be helpful and these were omitted.

Our study was limited to three different surgical procedures. We felt this was prudent given that these procedures are most commonly being performed throughout the world and have very good safety profiles and proven efficacy in obese patients. However, this does decrease the generalizability of the results and also potentially exclude other

surgical procedures that may have favorable results in this patient population. The heterogeneity of data reporting amongst studies was a serious limitation. The strength of performing a meta analysis comes from the ability to pool data and increase numbers of patients for analysis. In many instances data was just too heterogeneous for any meaningful analysis. Many studies would have for example 'improved' as their endpoint of interest instead of variables that are more easily measured. This ultimately limited our ability to analyze other obesity related comorbidities such as sleep apnea, infertility, urinary incontinence, etc.

Our review also had a paucity of randomized data. Only 4 of 13 studies included data that was collected in a randomized fashion. Future studies looking at the role of surgery for moderately obese patients should include well designed randomized controlled trials ideally with large numbers of patients. These studies should also have a standardized method of reporting outcomes to improve generalizability and interpretation of the data. This would also potentially make the studies less heterogeneous and allow researchers to explore whether other less commonly reported obesity related comorbidities such as sleep apnea can be effectively treated by surgery in this patient population. Inclusion of other surgical procedures such as duodenal switch or mini gastric bypass will also be essential to see if any significant difference exists with respect to moderately obese patients. Other meaningful analysis would be to consider if there is a low point in BMI where patients may or may not benefit from surgery. For example patients with low BMI (<25) and diabetes have been shown in previous studies to have less benefit from surgery as these patients typically have more pancreatic beta cell

dysfunction and less peripheral insulin resistance[35]. Analyzing patients based on class of obesity or subgroup of BMI will be useful in this regard.

## 5.1 – Conclusion

Bariatric surgery appears to be effective for the treatment of obesity related comorbidities in patients with BMI < 35. Safety profile is comparble to that seen in morbidly obese subjects with a relatively low risk of morbidity and mortality. Surgery compared to medical therapy offers improved results with respect to fasting plasma glucose, Hemoglobin A1c, weight loss, BMI loss, waist circumference, and resolution of hypertension and dyslipidemia. Variable rates of improvement are also seen with respect to obstructive sleep apnea, osteoarthritis, infertility, gastroesophageal reflux, urinary stress incontinence and venous stasis. Roux en y gastric bypass, sleeve gastrectomy and adjustable gastric banding when performed laparoscopically are safe and effective treatments for moderately obese individuals. While there was no difference between RYGB and LSG with respect to FPG and HbA1c, RYGB was associated with improved BMI loss and decreased waist circumference. These surgical procedures should be offered to patients with moderate obesity, especially those with related comorbidities and should be part of current treatment guidelines for obese individuals. With further studies these guidelines can be refined to ensure that the most appropriate surgical procedure can be offered to an individual patient. Future studies should include randomized controlled trials with large numbers of patients. These studies should aim to standardize reporting of variables such as FPG and HbA1c and where possible should standardize surgical procedures.

## Bibliography


1. Executive summary: Clinical guidelines on the identification, evaluation and treatment of overweight and obesity in adults. Accessed from: World Health Organization fact sheet on overweight and obese individuals. Accessed from: http://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight

2. Bagchi, D, Preuss, H.G. (2007). *Obesity: Epidemiology, pathophysiology, and prevention.* Boca Raton, FL, USA. CRC Press.

3. Ideal weight for height tables metropolitan life. Accessed from: https://somepomed.org/articulos/contents/mobipreview.htm?8/44/8905

4. Serrot, F.J, Dorman, R.B, Miller, C.J, Slusarek, B, Sampson, B, Sick, B.T, Leslie, D.B, Buchwald, H, Ikrammuddin, S. Comparative effectiveness of bariatric surgery and nonsurgical therapy in adults with type 2 diabetes mellitus and body mass index < 35 $kg/m^2$. Surgery. 2011;150(4):684-691.

5. Lockwood, D.H, Heffner, T.G. (2000). *Obesity: Pathology and Therapy*. Berlin: Springer

6. Bray, G.A, Bouchard, C. (2004). *Handbook of Obesity: Clinical applications, 2nd edition.* New York, NY: Marcel Dekker, Inc.

7. Sugerman, H.J, Nguyen, N.T. (2006). *Management of Morbid Obesity.* New York, NY: Taylor and Francis Group.

8. Consensus development conference panel, National Institutes of Health. Gastrointestinal surgery for severe obesity. Ann Int Med 1991; 115:956-61

9. Cohen, R.V, Pinheiro, J.C, Schiavon, C.A, Salles, J.E, Wajchenberg, B.L, Cummings, D.E. Effects of gastric bypass surgery in patients with type 2 diabetes and only mild obesity. Diabetes care. 2012; 35:1420-1428.

10. Mason, E.E, Ito, C. Gastric bypass. 1969; 170(3):329-339

11. Martin, L. Type 2 diabetes mellitus in patients with mild obesity: preliminary results of surgical treatment. Obes Surg. 2013; 23(2): 234-240.

12. Gumbiner, B. (2001). *Obesity.* USA: Versa Press.

13. Wadden, T.A, Stunkard, A.J. (2002). *Handbook of Obesity treatment.* New York, NY: The Guilford Press.

14. Rubino F, Marescaux J. Effect of duodenal-jejunal exclusion in a non-obese animal model of type 2 diabetes: a new perspective for an old disease. Ann Surg 2004; 239:1–11

15. Wittgrove, A.C, Clark, G.W. Laparoscopic gastric bypass, roux en y: experience of 27 cases, with 3-18 months follow up. Obes Surg. 1996; 6:54-57.

16. Nocca, D, Krawczykowsky, D, Bomans, B, Noel, P, Picot, M.C, Blanc, P.M, de Seguin de Hons, C, Millat, B, Gagner, M, Monnier, L, Fabre, J.M. A prospective multicenter study of 163 sleeve gastrectomies: results at 1 and 2 years. Obes Surg. 2008; 18:560-565.

17. Agency for healthcare research and quality. (2007). Bariatric surgery utilization and outcomes in 1998 and 2004. Accessed from http://www.hcup-us.ahrq.gov/reports/statbriefs/sb.jsp.

18. American Diabetes Association Position statement 7: Obesity management for the treatment of type 2 diabetes: Standards of medical care in diabetes. Diabetes Care. 2018; 41(1):565-572.

19. International Diabetes Federation. Recommendations For Managing Type 2 Diabetes In Primary Care, 2017. Access from: www.idf.org/managing-type2-diabetes

20. Canadian diabetes association clinical practice guideline expert committee. Canadian diabetes association 2008 clinical practice guidelines for the prevention and management of diabetes in Canada. Can J Diabetes 2008;32-(suppl 1):51-201

21. Pories, W.J, Flickinger, E.G, Meelheim, D, Van Rij, A.M, Thomas, F.T. The effectiveness of gastric bypass over gastric partition in morbid obesity. Ann Surg. 1982;196(4):389-399.

22. Dixon, J.B, O'Brien, P.E. Health outcomes of severely obese type 2 diabetic subjects 1 year after laparoscopic adjustable gastric banding. Diabetes Care. 2002; 25(2):358-363.

23. Sjöström L, Peltonen M, Jacobson P, et al. Bariatric surgery and long-term cardiovascular events. JAMA 2012; 305:56–65.

24. Noya, G, Cossu, M.L, Coppola, M, Tonolo, G, Angius, M.F, Fais, E, Ruggiu, M. Biliopancreatic diversion preserving the stomach and pylorus in the treatment of hypercholesterolemia and diabetes type 2: results in the first 10 cases. Obes Surg. 1998; 8:67-72.

25. Fried, M, Ribaric, G, Buchwald, J.N, Svacina, S, Dolezalova, K, Scopinaro, N. Metabolic surgery for the treatment of type 2 diabetes in patients with BMI < 35 kg/$m^2$: an integrative review of early studies. Obes Surg. 2010; 20:776-790.

26. Depaula, A.L, Macedo, L.V, Rassi, N, Machado, C.A, Schraibman, V, Silva, L.Q, Halpern, A. Laparoscopic treatment of type 2 diabetes mellitus for patients with a body mass index less than 35. Surg Endosc. 2008; 22:706-716.

27. Lee, W.J, Chong, K, Ser, K.H, Lee, Y.C, Chen, S.C, Chen, J.C, Tsai, M.H, Chuang, L.M. Gastric bypass vs. sleeve gastrectomy for type 2 diabetes mellitus. Arch Surg. 2011; 146(2):143-148.

28. Zhu, L, Mo, Z, Yang, X, Liu, S, Wang, G, Li, P, Tan, J, Strain, J, Im, I, Zhu, S. Effect of laparoscopic roux en y gastroenterostomy with BMI < 35 kg/$m^2$ in type 2 diabetes mellitus. Obes Surg. Oct 2012; 10:1562-1567.

29. Haidich, A.B. Meta analysis in medical research. Hippokratia. 2010 Dec; 14(1): 29-37.

30. Meijer, R.I, Van Wagensveld, B.A, Siegert, C.E, Eringa, E.C, Serne, E.H, Smulders, Y.M. Bariatric surgery as a novel treatment for type 2 diabetes mellitus. Arch Surg. 2011; 146(6):744-750.

31. Adegbola, S, Tayeh, S, Agrawal, S. Systematic review of laparoscopic adjustable gastric banding in patients with body mass index </= 35 kg/$m^2$. Surgery for obesity and related diseases. 2014; 10:155-160.

32. Rao, W.S, Shan, C.X, Zhang, W, Jiang, D.Z, Qiu, M. A meta analysis of short term outcomes of patients with type 2 diabetes mellitus and BMI </=35 kg/$m^2$ undergoing roux en y gastric bypass. World J Surg. 2015; 39(1):223-230.

33. Muller-Stich, B.P, Senft, J.D, Warschkow, R, Kenngott, H.G, Billeter, A.T, Vit, V, Helfert, S, Diener, M.K, Fischer, L, Buchler, M.W, Nawroth, P.P. Surgical versus medical treatment of type 2 diabetes mellitus in non severely obese patients: a systematic review and meta analysis. Annals of Surgery. 2015; 261(3):421-429.

34. Mantel, N, Haenszel, W. Statistical aspects of the analysis of data from retrospective studies of disease. JNCI. Apr 1959; 22 (4):719-748.

35. Liberati, A, Altman, D.G, Tetzlaff, J, Mulrow, C, Gotzsche, P.C, Ioannidis, J.P, Clarke, M, Devereaux, P.J, Kleijnen, J, Moher, D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 2009; 6(7):1-23.

36. Cochrane consumers and communication group resources for authors accessed from: http://cccrg.cochrane.org/author-resources.

37. Cochrane handbook. Accessed from: http://handbook-5.1.cochrane.org/chapter_10/10_4_1_funnel_plots.htm.

38. Viera, A.J. Odds ratios and risk ratios: what's the difference and why does it matter? Southern Medical Journal. 2008; 101 (7):730-734.

39. Higgins, P.T, Thompson, S.G, Deeks, J.J, Altman, D.G. Measuring inconsistency in meta analyses. BMJ. 2003; 327(6):557-560.

40. Kakoulidis, T.P, Karringer, A, Gloaguen, B.S, Arvidsson, D. Initial results with sleeve gastrectomy for patients with class I obesity (BMI 30-35 kg/$m^2$). Surgery for obesity and related diseases. 2009; 5:425-428.

41. Encinosa, W.E, Bernard, D.M, Du, D, Steiner, C. Recent improvements in bariatric surgery outcomes. Medical Care. 2009; 47(5):531-535.

42. Yang, J, Wang, C, Cao, G, Yang, W, Yu, S, Zhai, H, Pan, Y. Long term effects of laparoscopic sleeve gastrectomy versus roux en y gastric bypass for the treatment of Chinese type 2 diabetes mellitus patients with body mass index 28-35 kg/$m^2$. BMC Surgery. 2015; 15:88-95.

43. Parikh, M, Duncombe, M.B, Fielding, G.A. Laparoscopic adjustable gastric banding for patients with body mass index of </= 35 kg/$m^2$. Surgery for obesity and related diseases. 2006; 2:518-522.

44. Abbatini, F, Cappoccia, D, Casella, G, Coccia, F, Leonetti, F, Basso, N. Type 2 diabetes in obese patients with body mass index of 30-35 kg/$m^2$: sleeve gastrectomy versus medical treatment. Surgery for obesity and related diseases. 2012; 8:20-24.

45. Cohen, R, Pinheiro, J.S, Correa, J.L, Schiavon, C.A. Laparoscopic roux en y gastric bypass for BMI < 35 kg/$m^2$: a tailored approach. Surgery for obesity and related diseases. 2006; 2: 401-404.

46. Huang, C.K, Shabbit, A, Lo, C.H, Tai, C.M, Chen, Y.S, Houng, J.Y. Laparoscopic roux en y gastric bypass for the treatment of type 2 diabetes mellitus in Chinese patients with body mass index of 25-35. Obes Surg. 2011; 21:1344-1349.

47. Lanzarini, E, Csendes, A, Gutierres, L, Cuevas, P, Lembach, H, Molina, J, Musleh, M, Braghetto, I, Papapietro, K, Araya, V, Lahsen, R, Olguin, P, Riffo, A, Diaz, E, Montt, D, Marin, L. Type 2 diabetes mellitus in patients with mild obesity: preliminary results of surgical treatment. Obes Surg. 2013; 23(2): 234-240.

48. Lee, W.J, Ser, K.H, Chong, K, Lee, Y.C, Chen, S.C, Tsou, J.J, Chen, J.C, Chen, C.M. Laparoscopic sleeve gastrectomy for diabetes treatment in non morbidly obese patients: efficacy and change of insulin secretion. Surgery. 2010; 147(5): 664-669

49. O'Brien, P.E, Dixon, J.B, Laurie, C, Skinner, S, Proietto, J, McNeil, J, Strauss, B, Marks, S, Schachter, L, Chapman, L, Anderson, M. Treatment of mild to moderate

obesity with laparoscopic adjustable gastric banding or an intensive medical program: a randomized trial. Ann Intern Med. 2006; 144:625-633.

50. Scopinaro, N, Adami, G.F, Papadia, F.S, Camerini, G, Carlini, F, Briatore, L, Andraghetti, G, Catalano, M, Cordera R. Effects of gastric bypass on type 2 diabetes in patients with BMI 30 to 35. Obes Surg (2014) 24: 1036-1043.

51. Sultan, S, Parikh, M, Youn, H, Kurian, M, Fielding, G, Ren, C. Early U.S. outcomes after laparoscopic adjustable gastric banding in patients with body mass index less than 35 kg/$m^2$. Surg Endosc. 2009; 23:1569-1573

## Tables

Table 1 – Summary data Abbatini et. al. 2012[44]

| Abbatini et. al. 2012 | |
|---|---|
| **Risk of bias** | **Some concerns**<br>Details regarding blinding and randomization were lacking |
| **Study design** | **Randomized controlled trial** |
| **Population** | n = 18<br>single center<br>Italy<br>Included patients with BMI < 35 and T2DM |
| **Intervention** | Laparoscopic sleeve gastrectomy |
| **Control** | Conventional medical therapy |
| **Outcome** | T2DM resolution, dyslipidemia, obstructive sleep apnea<br>Hypertension |
| **Results** | T2DM and hypertension resolution occurred<br>In 88.9% of patients Dyslipidemia was corrected in all surgical patients. Sleep apnea was resolved in 1 patient.<br>In medical group all patients continued to have T2DM and Required antihypertensive and hypolipemic therapies |

Table 2 – Summary data Cohen et. al. 2006[45]

| Cohen et. al. 2006 | |
|---|---|
| **Risk of bias** | **Moderate**<br>Non randomized, non blinded |
| **Study design** | **Prospective Cohort** |
| **Population** | n = 37<br>single center<br>Brazil<br>Included patients with BMI < 35 and T2DM |
| **Intervention** | Laparoscopic roux en y gastric bypass |
| **Control** | Pre surgical group |
| **Outcome** | T2DM, hypertension, lipid disorder, GERD and sleep apnea. Also excess weight loss |
| **Results** | At 2 years of follow up 36/37 patients had complete resolution of all obesity related comorbidities studied. Mean excess weight loss was 81% |

Table 3 – Summary data Cohen et. al. 2012[9]

| Cohen et. al. 2012 | |
|---|---|
| **Risk of bias** | **Moderate**<br>Non randomized, non blinded |
| **Study design** | **Prospective Cohort** |
| **Population** | n = 66<br>single center<br>Brazil<br>Included patients with BMI < 35 and T2DM |
| **Intervention** | Laparoscopic roux en y gastric bypass |
| **Control** | Pre surgical group |
| **Outcome** | T2DM resolution or improvement, waist circumference, total body weight loss, hypertension, hypercholesterolemia. |
| **Results** | After 6 years of follow up T2DM resolution was maintained in 88% and glycemic improvement in 11%. HbA1c was also significantly reduced. Weight loss did not correlate with glucose homeostasis. Hypertension and dyslipidemia were also improved. Neither mortality nor major surgical morbidity were observed. |

Table 4 – Summary data DePaula et. al. 2008[26]

DePaula et. al. 2008

| | |
|---|---|
| **Study design** | Prospective cohort |
| **Population** | n = 39<br>BMI between 23.4-34.9 kg/m$^2$<br>Diabetics with HbA1c > 7.5<br>Multi center<br>Brazil |
| **Intervention** | ileal interposition with sleeve gastrectomy (ILSG) or ileal interposition with diverted sleeve gastrectomy (ILDSG). |
| **Control** | Preoperative group |
| **Outcome** | BMI, excess weight loss, resolution of diabetes, perioperative complications. |
| **Results** | Average follow up was 7 months. Major complications noted in 10.3% and minor complications in 15.4% of patients. Mean BMI decreased from 30.1kg.m$^2$ to 24.9kg/m$^2$. Diabetes was resolved in 86.9% of patients. Remaining patients were on single hypoglycemic agent and no patients were on insulin postoperatively. |

Table 5 – Summary data Dixon et. al. 2002[22]

Dixon et. al. 2002

| | |
|---|---|
| **Study design** | Retrospective cohort |
| **Population** | n = 50<br>BMI > 35 with diabetes<br>Single center<br>Australia |
| **Intervention** | Laparoscopic adjustable gastric banding |
| **Control** | Cohort of patients receiving medical therapy |
| **Outcome** | Weight loss, resolution of comorbidities and complications. |
| **Results** | Total follow up of one year. Mean excess weight loss was 38% which was statistically significantly lower compared to the control group. Significant improvements were also seen in the surgical group with respect to FPG, HbA1c, lipid improvement and improvement in liver enzymes. |

Table 6 – Summary data Huang et. al. 2011[46]

| Huang et. al. 2011 | |
|---|---|
| **Risk of bias** | **Moderate**<br>Non randomized, non blinded |
| **Study design** | **Prospective Cohort** |
| **Population** | n = 22<br>single center<br>Taiwan<br>Included patients with BMI < 35 and T2DM |
| **Intervention** | Laparoscopic roux en y gastric bypass |
| **Control** | Pre surgical group |
| **Outcome** | T2DM resolution or improvement, BMI, cholesterol and triglycerides |
| **Results** | After 12 months of follow up 14 patients (63.6%) showed T2DM remission, 6 (27.3%) showed glycemic control and 2 (9.1%) showed improvement. Comparing patients who had T2DM remission versus control or improvement there was a significant difference in BMI and duration of DM with patients who were in remission having slightly higher BMI and shorter duration of DM. no difference was seen with respect to cholesterol and triglycerides. |

Table 7 – Summary data Lanazarini et. al. 2013[47]

| Lanzarini et. al. 2013 | |
|---|---|
| **Risk of bias** | **Moderate**<br>Non randomized, non blinded |
| **Study design** | **Prospective Cohort** |
| **Population** | n = 31<br>single center<br>Chile<br>Included patients with BMI < 35 and T2DM<br>Age between 18 and 65<br>Adequate nutritional, psychological and endocrinological assessments |
| **Intervention** | Laparoscopic roux en y gastric bypass |
| **Control** | Pre surgical group |
| **Outcome** | T2DM, use of oral hypoglycemic agents, BMI |
| **Results** | At 36 months follow up BMI significantly decreased to mean of 24.7kg/m$^2$. All patients showed improvement in glycemic control and 93.6% met criteria for remission. Only one patient had a postoperative complication (hemoperitoneum). |

Table 8 – Summary data Lee et. al. 2010[48]

| Lee et. al. 2010 | |
|---|---|
| **Risk of bias** | **Moderate**<br>Non randomized, non blinded |
| **Study design** | **Prospective Cohort** |
| **Population** | n = 20<br>single center<br>Taiwan<br>Included patients with BMI between 25 and 35.<br>Patients with T2DM |
| **Intervention** | Laparoscopic sleeve gastrectomy |
| **Control** | Pre surgical group |
| **Outcome** | T2DM, use of oral hypoglycemic agents, BMI, waist circumference, excess weight loss. |
| **Results** | At one year follow up there were significant improvements in BMI, waist circumference, EWL, fasting glucose, and HbA1c |

Table 9 – Summary data Lee et. al. 2011[27]

| Lee. et. al. 2011 | |
|---|---|
| **Risk of bias** | **Low** |
| **Study design** | **Randomized controlled trial** |
| **Population** | n = 60<br>single center<br>Taiwan<br>Included patients with BMI between 25 and 35<br>Age < 60<br>Poorly controlled T2DM |
| **Intervention** | **Laparoscopic roux en y gastric bypass versus Laparoscopic sleeve gastrectomy** |
| **Control** | Sleeve gastrectomy group |
| **Outcome** | Remission of T2DM, BMI, waist circumference, hypertension, blood lipid levels. |
| **Results** | At 12 months after surgery there were significant differences favoring the gastric bypass group including successful treatment of T2DM which occurred in 17 patients (57%) as compared to 0 in the sleeve gastrectomy group. There were also significant differences noted with respect to HDL and persistence of metabolic syndrome. HbA1c and FPG were significantly different. No major surgical complications in either group. |

Table 10 – Summary data Mason et. al. 1967[10]

Mason et. al. 1967

| | |
|---|---|
| **Study design** | Retrospective |
| **Population** | Not well defined, perceived morbidly obese patients |
| **Intervention** | Open gastric bypass |
| **Control** | Cohort study, control group is based on preoperative assessment |
| **Outcome** | Weight loss, perioperative complications, resolution of peptic ulcer related symptoms. |
| **Results** | Patients were operated over a 3 year period. Average weight loss was 86 pounds. 2 deaths reported and 2 major complications in the follow up period. One patient underwent revisional surgery because of jejunal ulcer however details regarding this were lacking in the report. |

Table 11 – Summary data Nocca et. al. 2007[16]

Nocca et. al. 2007

| | |
|---|---|
| **Study design** | Prospective cohort |
| **Population** | n = 167<br>multicenter<br>BMI > 40 or > 35 with comorbidities<br>France, USA |
| **Intervention** | Laparoscopic sleeve gastrectomy |
| **Control** | Pre surgical group |
| **Outcome** | Excess weight loss, perioperative complications. |
| **Results** | Total follow up duration of 2 years. Mean excess weight loss was 61.5%. Complication rate was 7.4% with reoperative rate of 4.9%. Mean BMI went from 45.9kg/$m^2$ to 31.6kg/$m^2$. |

Table 12 – Summary data Noya et. al. 1998[24]

Noya et. al. 1998

| | |
|---|---|
| **Study design** | Retrospective |
| **Population** | n = 10<br>BMI < 35<br>Diabetes, hypercholesterolemia and hypertension<br>Single center<br>Italy |
| **Intervention** | Biliopancreatic diversion |
| **Control** | Preoperative group |
| **Outcome** | Glycemic indices, resolution of hypertension and hypercholesterolemia |
| **Results** | Maximum follow up was 12 months. All patients had resolution of hypertriglyceridemia and hypertension. Mean BMI decreased from 33.2kg/$m^2$ to 30.5kg/$m^2$. 9 of 10 patients had normoglycemic values postoperatively. |

Table 13 – Summary data O'Brien et. al. 2006[49]

| O'Brien et. al. 2006 | |
|---|---|
| **Risk of bias** | **Some Concerns**<br>Non blinded |
| **Study design** | **Randomized controlled trial** |
| **Population** | n = 72<br>single center<br>Australia<br>Included patients with BMI between 30 and 35 |
| **Intervention** | **Laparoscopic Adjustable gastric banding versus intense medical therapy** |
| **Control** | Medical group |
| **Outcome** | Weight change, presence of metabolic syndrome and quality of life |
| **Results** | Patients were followed for 2 years. Surgical group was found to have greater weight loss and excess weight loss. Better resolution of metabolic syndrome was seen in the surgical group with 14/15 patients being resolved versus 8/15 in the medical group. Quality of life was improved significantly more in the surgical group than in the medical group (100% versus 38%). 18% of surgical patients had complications related to surgery with 10% requiring revisional surgery. |

Table 14 – Summary data Parikh et. al. 2006[43]

Parikh et. al. 2006

| | |
|---|---|
| **Risk of bias** | **Serious**<br>Data given on BMI and EWL<br>however no statistical analysis was performed to compare treatment and control groups. |
| **Study design** | **Prospective Cohort** |
| **Population** | **n = 93**<br>Single center<br>Australia<br>Patients with BMI 30-35 |
| **Intervention** | **Laparoscopic Adjustable gastric band** |
| **Control** | **Pre surgical group** |
| **Outcome** | **Weight (kg), BMI, Resolution of comorbidities** |
| **Results** | Weight reduced to 72kg at 3 years<br>EWL was 53.8%<br>BMI was reduced to 18-24 in 34% of patients<br>All diabetic patients were off medications after surgery<br>6 of 8 patients with hypertension were off medication<br>4 of 5 patients with asthma were off medication<br>6 of 7 patients with sleep apnea did not need CPAP<br>8 of 9 patients with arthritis improved<br>4 of 11 patients with depression did not need medication |

Table 15 – Summary data Pories et. al. 1982[21]

Pories et. al. 1982

| | |
|---|---|
| **Study design** | Randomized controlled trial |
| **Population** | n = 87<br>Patients identified double normal weight<br>Average weight 308.5 pounds<br>Range 219-491 pounds<br>Single center<br>USA |
| **Intervention** | Randomized to open gastric bypass versus gastric partition |
| **Control** | Gastric partition group |
| **Outcome** | Weight loss, complications. |
| **Results** | Total follow up was 15 months. Gastric bypass had superior weight loss and at that point randomization was discontinued and patients exclusively underwent gastric bypass. At 18 months follow up gastric bypass had surpassed gastric partition in terms of weight loss by 68 pounds which was statistically significant. Complications were comparable between groups. |

Table 16 – Summary data Scopinaro et. al. 2014[50]

| Scopinaro et. al. 2014 | |
|---|---|
| **Risk of bias** | **Moderate**<br>Non randomized, non blinded |
| **Study design** | **Prospective Cohort/ Matched Case Control** |
| **Population** | n = 20 (27 matched controls who underwent medical therapy)<br>single center<br>Italy<br>Included patients with BMI between 30 and 34.9.<br>Age 35-70 years<br>Patients with T2DM<br>Patients compared to post surgical versus baseline and also to matched controls. |
| **Intervention** | Laparoscopic roux-en-Y gastric bypass |
| **Control** | Pre surgical group |
| **Outcome** | T2DM, body weight, BMI, waist circumference, hypertension, triglycerides and cholesterol values. |
| **Results** | Significant improvements in body weight, BMI, waist circumference, fasting plasma glucose, HbA1c as well as resolution of hypertension, hypercholesterolemia and hypertriglyceridemia. Body weight and BMI were significantly improved compared to control patients however effects on comorbidities were not translated to this group. No mortalities, 1 case of perioperative bleeding was noted. |

Table 17 – Summary data Serrot et. al. 2011[4]

Serrot et. al. 2011

| | |
|---|---|
| **Study design** | Retrospective matched case control |
| **Population** | n = 34<br>BMI < 35<br>Diabetes<br>Single center<br>USA |
| **Intervention** | Laparoscopic roux en y gastric bypass |
| **Control** | Medical therapy |
| **Outcome** | BMI, HbA1c, medication use, hypertension and LDL cholesterol. |
| **Results** | Maximum follow up was 34 months. Surgical group had change in BMI from 34.6+/-0.8 $kg/m^2$ to 25.8+/-2.5 $kg/m^2$ compared to essentially no change in the medical group from 34.1+/-1.0 $kg/m^2$ to 34.3+/-2.1 $kg/m^2$ ( $p < 0.001$). HbA1c decreased in the REYGB group from 8.2 +/-2.0 to 6.1+/-2.7 but did not change in the medical group (7.0+/-0.7 to 7.1+/-1.8 P < 0.001). Blood pressure and LDL cholesterol did not significantly change in either group. Fewer patients in the surgical group were taking medications for glycemic control at the end of the study with 71% of patients taking less medications at one year compared to 6% in the non surgical group ($p < 0.001$). |

Table 18 – Summary data Sultan et. al. 2009[51]

Sultan et. al. 2009

| | |
|---|---|
| **Risk of bias** | **Serious**<br>Statistical analysis was substandard<br>Only descriptives given with no analysis to<br>Determine level of significance |
| **Study design** | **Prospective Cohort** |
| **Population** | n = 53<br>single center<br>USA<br>Included patients with BMI < 35 |
| **Intervention** | Laparoscopic adjustable gastric banding |
| **Control** | Pre surgical group |
| **Outcome** | BMI, weight loss, excess weight loss and<br>Resolution of comorbidities.<br>Surgical complications |
| **Results** | BMI decreased to 25.8 at 2 years of follow up.<br>EWL 69.7%<br>13.2% rate of complications. |

Table 19 – Summary data Wittgrove et. al. 1996[15]

Wittgrove et. al. 1996

| | |
|---|---|
| **Study design** | Prospective Cohort |
| **Population** | n = 27<br>BMI > 40 or BMI > 35 with comorbidities<br>Single center<br>USA |
| **Intervention** | Laparoscopic roux en y gastric bypass |
| **Control** | Pre operative group |
| **Outcome** | Weight loss, resolution of comorbidities, perioperative complications. |
| **Results** | Follow up for total of 18 months. Mean excess weight loss of 80%. Obesity related comorbidities resolved in 98%. 9 patients suffered complications with one anastomotic leak. Remainder of patients were managed conservatively. |

Table 20 – Summary data Yang et. al. 2015[42]

| Yang et. al. 2015 | |
|---|---|
| **Risk of bias** | **Some Concerns**<br>Details regarding blinding were lacking |
| **Study design** | **Randomized controlled trial** |
| **Population** | n = 55<br>single center<br>China<br>Included patients with BMI between 28 and 35<br>HbA1c >/= 7.0% |
| **Intervention** | **Laparoscopic roux en y gastric bypass versus laparoscopic sleeve gastrectomy** |
| **Control** | Sleeve gastrectomy group |
| **Outcome** | Percentage excess weight loss, BMI, waist circumference, HbA1c, FPG and serum lipid levels. |
| **Results** | After 3 years of follow up there were no significant differences between surgical groups with respect to HbA1c, serum lipids and FPG. Gastric bypass resulted in significantly greater excess weight loss and change in BMI than sleeve gastrectomy. |

Table 21 – Summary data Zhu et. al. 2012[28]

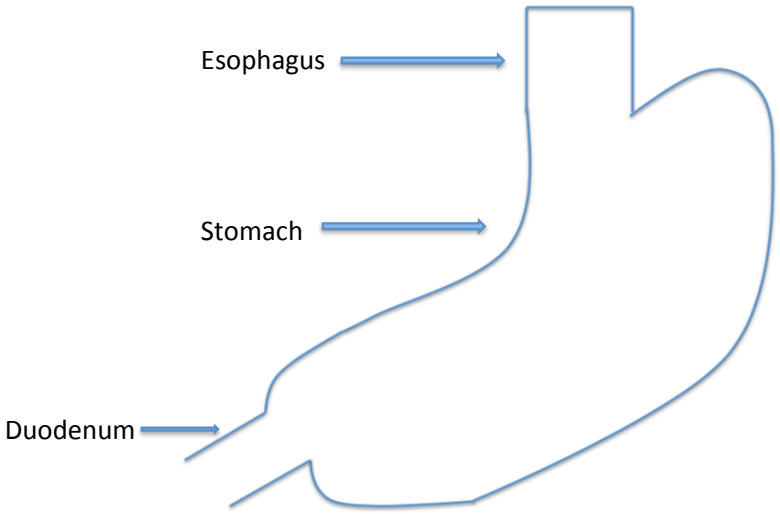
Zhu et. al. 2012

| | |
|---|---|
| **Risk of bias** | **Moderate**<br>Non randomized, non blinded |
| **Study design** | **Prospective Cohort** |
| **Population** | n = 30<br>single center<br>China<br>Included patients with BMI < 35 |
| **Intervention** | **Laparoscopic roux en y gastric bypass** |
| **Control** | Pre surgical group |
| **Outcome** | T2DM, BMI, waist circumference |
| **Results** | Significant reduction in HbA1c was noted at one year of follow up. T2DM was completely resolved in 9 cases. Significant improvements were also noted in BMI, waist circumference, fasting plasma glucose and BMI. No significant perioperative complications were reported. |

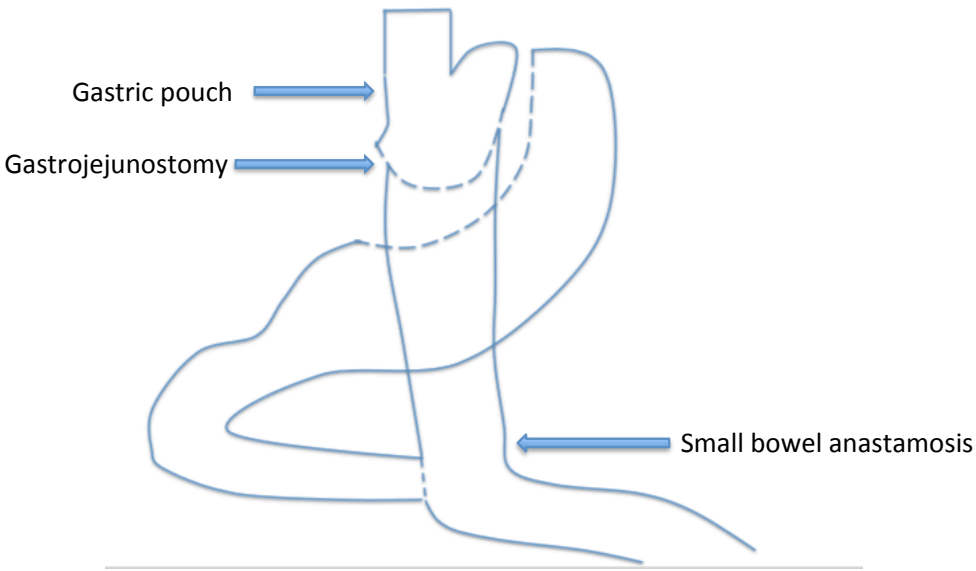
## Table 22 – Summary of study characteristics

| | Risk of Bias | Study Design | No. of patients | Country | Inclusion criteria | Intervention | Control | Outcome | Results |
|---|---|---|---|---|---|---|---|---|---|
| Abbatini et. al. 2012 | Some concerns | Randomized controlled trial | 18 | Italy (single center) | BMI <35 with T2DM | LSG | Medical therapy | T2DM resolution, dyslipidemia,OSA and HTN | 88.9% of patients had resolved T2DM and hypertension . Dyslipidemia resolved in all surgical patients |
| Cohen et. al. 2006 | Moderate | Prospective cohort | 37 | Brazil (single center) | BMI <35 with T2DM | LREYGB | Pre-surgical group | T2DM, HTN, lipid disorder, GERD, OSA and excess weight loss | 36/37 patients had complete resolution of comorbidities 81% EWL |
| Cohen et. al. 2012 | Moderate | Prospective cohort | 66 | Brazil (single center) | BMI <35 with T2DM | LREYGB | Pre-surgical group | T2DM, waist circumference, total weight loss, HTN, hypercholesterolemia | 88% resolution T2DM at 6 years, glycemic improvement in 11%, |
| Huang et. al. 2011 | Moderate | Prospective cohort | 22 | Taiwan (single center) | BMI <35 with T2DM | LREYGB | Pre-surgical group | T2DM resolution or improvement, BMI, cholesterol, triglycerides | Follow up 14 patients, at 12 months. 63.6% T2DM remission, 27.3% glycemic control |
| Lanzarini et. al. 2013 | Moderate | Prospective cohort | 31 | Chile (single center) | BMI <35 with T2DM and age 18-65 | LREYGB | Pre-surgical group | T2DM, use of oral hypoglycemic agents, BMI | BMI decreased by mean 24.7kg/m$^2$, all had improved glycemic control and 93.6% in remission. |
| Lee et. al. 2010 | Moderate | Prospective cohort | 20 | Taiwan (single center) | BMI 25-35 with T2DM | LSG | Pre-surgical group | T2DM, use of oral hypoglycemic agents, BMI, waist circumference, excess weight loss. | Significant improvements in BMI, waist circumference, excess weight loss, fasting glucose and HbA1C at one year |
| Lee. et. al. 2011 | Low | Randomized controlled trial | 60 | Taiwan (single center) | BMI 25-35, age 30 - 60, poorly controlled T2DM | LREYGB | LSG | Remission of T2DM, BMI, waist circumference, hypertension, blood lipid levels. | 93% remission T2DM in LREYGB compared to 47% in LSG at 12 months follow up. Significant differences in HbA1C, FPG and hyperlipidemia favoring LREYGB |
| O'Brien et. al. 2006 | Some concerns | Randomized controlled trial | 72 | Australia (single center) | BMI 30-35 | Laparoscopic adjustable gastric band | Medical therapy | Weight change, presence of metabolic syndrome, change in quality of life | Greater weight loss in surgical group. Resolution of metabolic syndrome in 14/15 patients in surgical group |
| Parikh et. al. 2006 | Serious | Prospective cohort | 93 | Australia (single center) | BMI 30-35 | Laparoscopic adjustable gastric band | Pre-surgical group | Weight, BMI, resolution of comorbidities (asthma, diabetes, hypertension, sleep apnea) | Mean weight reduced from 98 to 72 kg at 2 years, BMI reduced to 18-24 in 34%, all diabetic patients off medications at 2 years. |
| Scopinaro et. al. 2014 | Moderate | Prospective cohort/matched case control | 20 | Italy (single center) | BMI 30-35, aged 35-70 with T2DM | LREYGB | Pre-surgical and matched controls | T2DM, body weight, BMI, waist circumference, hypertension, triglycerides, cholesterol | Improvements in body weight, BMI, waist circumference, FPG, HbA1C, resolution of HTN, hypercholesterolemia |
| Sultan et. al. 2009 | Serious | Prospective cohort | 53 | USA (single center) | BMI <35 | Laparoscopic adjustable gastric band | Pre-surgical group | BMI, weight loss, excess weight loss, resolution of comorbidities and surgical complications | BMI decreased to mean of 25.8 at 2 year follow up. Excess weight loss of 69.7%. 13.2% rate of complications |
| Yang et. al. 2015 | Some | Randomized controlled trial | 55 | China (single center) | BMI 28-35 with HbA1C >/= 7.0% | LREYGB | LSG | Excess weight loss, BMI, waist circumference, HbA1C, FPG, serum lipid levels | LREYGB had significantly improved excess weight loss and decreased BMI |
| Zhu et. al. 2012 | Moderate | Prospective cohort | 30 | China (single center) | BMI <35 | LREYGB | Pre-surgical group | T2DM, BMI, waist circumference | Significant reduction in HbA1C and improvements in BMI, WC, and FPG in surgical group |

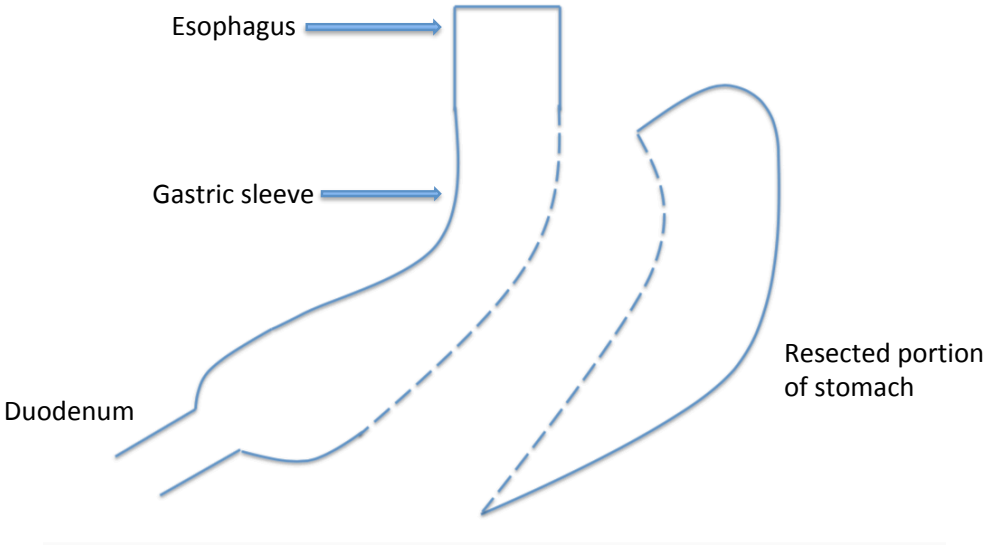
Appendix 1 – Normal Gastric anatomy



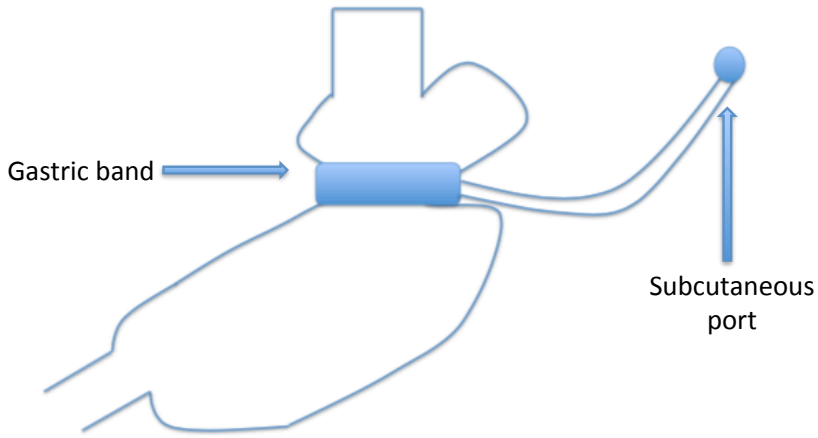
Appendix 2 – Roux-en-Y gastric bypass diagram



Appendix 3 – sleeve gastrectomy diagram



Appendix 4 – adjustable gastric band diagram



Appendix 5 – List of search terms/ search strategy

1 exp Bariatrics/ (11647)
2 Anastomosis, Roux-en-Y/ (2502)
3 Biliopancreatic Diversion/ (675)
4 (bariatric$ or gastric bypass or Gastroplast$ or Lipectom$ or jejunoileal bypass or gastric band$ or biliopancreatic diversion$).tw. (11379)
5 ((vertical or sleeve) adj2 gastrectom$).tw. (652)
6 (weight loss adj2 (surg$ or operat$)).tw. (959)
7 (obes$ adj2 (surg$ or operat$)).tw. (1671)
8 LAGB.tw. (549)
9 lap band$.tw. (214)
10 Roux-en-Y.tw. (4719)
11 duodenal switch$.tw. (350)
12 or/1-11 (19656)
13 ((bmi or body mass index) adj3 ("25$" or "26$" or "27$" or "28$" or "29$" or "30$" or "31$" or "32$" or "33$" or "34$")).tw. (20377)
14 bmi < 35$.tw. (644)
15 bmi<35$.tw. (56)
16 body mass index < 35$.tw. (199)
17 bmi less than 35$.tw. (5)
18 body mass index less than 35$.tw. (10)
19 bmi lower than 35$.tw. (6)
20 bmi below 35$.tw. (3)
21 body mass index below 35$.tw. (1)
22 (non adj2 obes$).tw. (6525)
23 (moderat$ adj2 obes$).tw. (863)
24 (class 1 adj1 obes$).tw. (28)
25 normal weight.tw. (9030)
26 or/13-25 (34299)
27 12 and 26 (1226)
28 animals/ not humans/ (3611731)
29 27 not 28 (1210)
30 "Diabetes Remission and insulin secretion after gastric bypass in patients with body mass index".ti. (1)
31 "Effect of Laparoscopic mini gastric bypass for type 2 diabetes mellitus".ti. (1)
32 "Early postoperative outcomes of metabolic surgery to treat diabetes from sites participating in the ASMBS bariatric surgery center".ti. (1)
33 "Outcomes of bariatric surgery in patients with body mass index".ti. (1)
34 "laparoscopic sleeve gastrectomy for diabetes treatment in nonmorbidly obese patients".ti.
35 or/30-34 (5)
36 29 or 35 (1210)

Appendix 6 – Risk of Bias Assessment (ROBINS-I) tool – Cohen et. al. 2006

**Risk of bias assessment**

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

| Signalling questions | Description | Response options |
|---|---|---|
| **Bias due to confounding** | | |
| 1.1 Is there potential for confounding of the effect of intervention in this study?<br>**If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered | Yes – study design and differences amongst study participants will contribute to confounding effect. Study authors also excluded patients who had appropriate BMI criteria but not other metabolic associated criteria. | Y / PY / PN / N |
| **If Y/PY to 1.1**: determine whether there is a need to assess time-varying confounding: | | |
| 1.2. Was the analysis based on splitting participants' follow up time according to intervention received?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, go to question 1.3. | No | NA / Y / PY / PN / N / NI |
| 1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8) | | NA / Y / PY / PN / N / NI |
| **Questions relating to baseline confounding only** | | |
| 1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains? | Yes – appropriate statistical analysis with 2 sided t test was performed for continuous variables. | NA / Y / PY / PN / N / NI |
| 1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | No | NA / Y / PY / PN / N / NI |
| 1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention? | No | NA / Y / PY / PN / N / NI |
| **Questions relating to baseline and time-varying confounding** | | |
| 1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding? | Yes | NA / Y / PY / PN / N / NI |
| 1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | Yes | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to confounding? | Unpredictable | Favours experimental / Favours comparator / Unpredictable |

| Bias in selection of participants into the study | | |
|---|---|---|
| 2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention?<br>**If N/PN to 2.1:** go to 2.4 | No | Y / PY / PN / N / NI |
| 2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention? | | NA / Y / PY / PN / N / NI |
| 2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome? | | NA / Y / PY / PN / N / NI |
| 2.4. Do start of follow-up and start of intervention coincide for most participants? | Yes | Y / PY / PN / N / NI |
| 2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to selection of participants into the study? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in classification of interventions | | |
|---|---|---|
| 3.1 Were intervention groups clearly defined? | Yes | Y / PY / PN / N / NI |
| 3.2 Was the information used to define intervention groups recorded at the start of the intervention? | Yes | Y / PY / PN / N / NI |
| 3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to classification of interventions? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias due to deviations from intended interventions | | |
|---|---|---|
| **If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2** | | |
| 4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice? | No | Y / PY / PN / N / NI |
| 4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome? | | NA / Y / PY / PN / N / NI |
| **If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6** | | |
| 4.3. Were important co-interventions balanced across intervention groups? | | Y / PY / PN / N / NI |
| 4.4. Was the intervention implemented successfully for most participants? | | Y / PY / PN / N / NI |
| 4.5. Did study participants adhere to the assigned intervention regimen? | | Y / PY / PN / N / NI |
| 4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | |
| Optional: What is the predicted direction of bias due to deviations from the intended interventions? | | |

| Bias due to missing data | | |
|---|---|---|
| 5.1 Were outcome data available for all, or nearly all, participants? | Yes | Y / PY / PN / N / NI |
| 5.2 Were participants excluded due to missing data on intervention status? | No | Y / PY / PN / N / NI |
| 5.3 Were participants excluded due to missing data on other variables needed for the analysis? | No | Y / PY / PN / N / NI |
| 5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions? | | NA / Y / PY / PN / N / NI |
| 5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to missing data? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in measurement of outcomes | | |
|---|---|---|
| 6.1 Could the outcome measure have been influenced by knowledge of the intervention received? | No | Y / PY / PN / N / NI |
| 6.2 Were outcome assessors aware of the intervention received by study participants? | Yes – non randomized study, non blinded. | Y / PY / PN / N / NI |
| 6.3 Were the methods of outcome assessment comparable across intervention groups? | Yes | Y / PY / PN / N / NI |
| 6.4 Were any systematic errors in measurement of the outcome related to intervention received? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to measurement of outcomes? | Favours experimental | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in selection of the reported result | | |
|---|---|---|
| Is the reported effect estimate likely to be selected, on the basis of the results, from... | No | |
| 7.1. ... multiple outcome *measurements* within the outcome domain? | | Y / PY / PN / N / NI |
| 7.2 ... multiple *analyses* of the intervention-outcome relationship? | No | Y / PY / PN / N / NI |
| 7.3 ... different *subgroups*? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to selection of the reported result? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Overall bias | | |
|---|---|---|
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the overall predicted direction of bias for this outcome? | Unpredictable | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

Appendix 7 – Risk of Bias Assessment (ROBINS-I) tool – Cohen et. al. 2012

**Risk of bias assessment**

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

| Signalling questions | Description | Response options |
|---|---|---|
| **Bias due to confounding** | | |
| 1.1 Is there potential for confounding of the effect of intervention in this study?<br>**If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered | Yes – some bias with respect to number of study participants as this was defined as number of patients who could be followed given the grant money that was funded. Also some characteristics were dissimilar at baseline. | Y / PY / PN / N |
| **If Y/PY to 1.1**: determine whether there is a need to assess time-varying confounding: | | |
| 1.2. Was the analysis based on splitting participants' follow up time according to intervention received?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, go to question 1.3. | No | NA / Y / PY / PN / N / NI |
| 1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8) | | NA / Y / PY / PN / N / NI |
| **Questions relating to baseline confounding only** | | |
| 1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains? | Yes – bonferroni adjusted repeated measures ANOVA was used to measure changes in baseline metabolic parameters. All tests were two sided with p values < 0.05. | NA / Y / PY / PN / N / NI |
| 1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | Yes | NA / Y / PY / PN / N / NI |
| 1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention? | No | NA / Y / PY / PN / N / NI |
| **Questions relating to baseline and time-varying confounding** | | |
| 1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding? | Yes | NA / Y / PY / PN / N / NI |
| 1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | Yes | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to confounding? | Unpredictable | Favours experimental / Favours comparator / Unpredictable |

| Bias in selection of participants into the study | | |
|---|---|---|
| 2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? | No | Y / PY / PN / N / NI |
| **If N/PN to 2.1:** go to 2.4 | | |
| 2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention? | | NA / Y / PY / PN / N / NI |
| 2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome? | | NA / Y / PY / PN / N / NI |
| 2.4. Do start of follow-up and start of intervention coincide for most participants? | Yes | Y / PY / PN / N / NI |
| 2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to selection of participants into the study? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in classification of interventions | | |
|---|---|---|
| 3.1 Were intervention groups clearly defined? | Yes | Y / PY / PN / N / NI |
| 3.2 Was the information used to define intervention groups recorded at the start of the intervention? | Yes | Y / PY / PN / N / NI |
| 3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to classification of interventions? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias due to deviations from intended interventions | | |
|---|---|---|
| **If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2** | | |
| 4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice? | No | Y / PY / PN / N / NI |
| 4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome? | | NA / Y / PY / PN / N / NI |
| **If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6** | | |
| 4.3. Were important co-interventions balanced across intervention groups? | | Y / PY / PN / N / NI |
| 4.4. Was the intervention implemented successfully for most participants? | | Y / PY / PN / N / NI |
| 4.5. Did study participants adhere to the assigned intervention regimen? | | Y / PY / PN / N / NI |
| 4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | |
| Optional: What is the predicted direction of bias due to deviations from the intended interventions? | | |

| Bias due to missing data | | |
|---|---|---|
| 5.1 Were outcome data available for all, or nearly all, participants? | Yes | Y / PY / PN / N / NI |
| 5.2 Were participants excluded due to missing data on intervention status? | No | Y / PY / PN / N / NI |
| 5.3 Were participants excluded due to missing data on other variables needed for the analysis? | No | Y / PY / PN / N / NI |
| 5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions? | | NA / Y / PY / PN / N / NI |
| 5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to missing data? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in measurement of outcomes | | |
|---|---|---|
| 6.1 Could the outcome measure have been influenced by knowledge of the intervention received? | No | Y / PY / PN / N / NI |
| 6.2 Were outcome assessors aware of the intervention received by study participants? | Yes – prospective study, non randomized, non blinded | Y / PY / PN / N / NI |
| 6.3 Were the methods of outcome assessment comparable across intervention groups? | Yes | Y / PY / PN / N / NI |
| 6.4 Were any systematic errors in measurement of the outcome related to intervention received? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to measurement of outcomes? | Favours experimental, patients may be more likely to have aggressive therapy post operatively given only one intervention group. | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in selection of the reported result | | |
|---|---|---|
| Is the reported effect estimate likely to be selected, on the basis of the results, from...<br>7.1. ... multiple outcome *measurements* within the outcome domain? | No | Y / PY / PN / N / NI |
| 7.2 ... multiple *analyses* of the intervention-outcome relationship? | No | Y / PY / PN / N / NI |
| 7.3 ... different *subgroups*? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to selection of the reported result? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Overall bias | | |
|---|---|---|
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the overall predicted direction of bias for this outcome? | Unpredictable | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

Appendix 8 – Risk of Bias Assessment (ROBINS-I) tool – Huang et. al. 2011

**Risk of bias assessment**

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

| Signalling questions | Description | Response options |
|---|---|---|
| **Bias due to confounding** | | |
| 1.1 Is there potential for confounding of the effect of intervention in this study?<br>**If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered | Yes – study design and differences amongst study participants will contribute to confounding effect. | Y / PY / PN / N |
| **If Y/PY to 1.1**: determine whether there is a need to assess time-varying confounding: | | |
| 1.2. Was the analysis based on splitting participants' follow up time according to intervention received?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, go to question 1.3. | No | NA / Y / PY / PN / N / NI |
| 1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8) | | NA / Y / PY / PN / N / NI |
| **Questions relating to baseline confounding only** | | |
| 1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains? | Yes – paired t test and Fisher's exact test were used when appropriate. All tests were two tailed. | NA / Y / PY / PN / N / NI |
| 1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | No | NA / Y / PY / PN / N / NI |
| 1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention? | No | NA / Y / PY / PN / N / NI |
| **Questions relating to baseline and time-varying confounding** | | |
| 1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding? | Yes | NA / Y / PY / PN / N / NI |
| 1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | Yes | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to confounding? | | Favours experimental / Favours comparator / Unpredictable |

| Bias in selection of participants into the study | | |
|---|---|---|
| 2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? | No | Y / PY / PN / N / NI |
| **If N/PN to 2.1:** go to 2.4 | | |
| 2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention? | | NA / Y / PY / PN / N / NI |
| 2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome? | | NA / Y / PY / PN / N / NI |
| 2.4. Do start of follow-up and start of intervention coincide for most participants? | Yes | Y / PY / PN / N / NI |
| 2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to selection of participants into the study? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in classification of interventions | | |
|---|---|---|
| 3.1 Were intervention groups clearly defined? | Yes | Y / PY / PN / N / NI |
| 3.2 Was the information used to define intervention groups recorded at the start of the intervention? | Yes | Y / PY / PN / N / NI |
| 3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to classification of interventions? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias due to deviations from intended interventions | | |
|---|---|---|
| **If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2** | | |
| 4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice? | No | Y / PY / PN / N / NI |
| 4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome? | | NA / Y / PY / PN / N / NI |
| **If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6** | | |
| 4.3. Were important co-interventions balanced across intervention groups? | | Y / PY / PN / N / NI |
| 4.4. Was the intervention implemented successfully for most participants? | | Y / PY / PN / N / NI |
| 4.5. Did study participants adhere to the assigned intervention regimen? | | Y / PY / PN / N / NI |
| 4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | |
| Optional: What is the predicted direction of bias due to deviations from the intended interventions? | | |

| Bias due to missing data | | |
|---|---|---|
| 5.1 Were outcome data available for all, or nearly all, participants? | Yes | Y / PY / PN / N / NI |
| 5.2 Were participants excluded due to missing data on intervention status? | No | Y / PY / PN / N / NI |
| 5.3 Were participants excluded due to missing data on other variables needed for the analysis? | No | Y / PY / PN / N / NI |
| 5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions? | | NA / Y / PY / PN / N / NI |
| 5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to missing data? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in measurement of outcomes | | |
|---|---|---|
| 6.1 Could the outcome measure have been influenced by knowledge of the intervention received? | No | Y / PY / PN / N / NI |
| 6.2 Were outcome assessors aware of the intervention received by study participants? | Yes – non randomized study, non blinded. | Y / PY / PN / N / NI |
| 6.3 Were the methods of outcome assessment comparable across intervention groups? | Yes | Y / PY / PN / N / NI |
| 6.4 Were any systematic errors in measurement of the outcome related to intervention received? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to measurement of outcomes? | Favours experimental | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in selection of the reported result | | |
|---|---|---|
| Is the reported effect estimate likely to be selected, on the basis of the results, from... | No | |
| 7.1. ... multiple outcome *measurements* within the outcome domain? | | Y / PY / PN / N / NI |
| 7.2 ... multiple *analyses* of the intervention-outcome relationship? | No | Y / PY / PN / N / NI |
| 7.3 ... different *subgroups*? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to selection of the reported result? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Overall bias | | |
|---|---|---|
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the overall predicted direction of bias for this outcome? | Unpredictable | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

Appendix 9: Risk of Bias Assessment (ROBINS-I) tool – Lanzarini et. al. 2013

**Risk of bias assessment**

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

| Signalling questions | Description | Response options |
|---|---|---|
| **Bias due to confounding** | | |
| 1.1 Is there potential for confounding of the effect of intervention in this study?<br>**If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered | Yes – study design and differences amongst study participants will contribute to confounding effect. | Y / PY / PN / N |
| **If Y/PY to 1.1**: determine whether there is a need to assess time-varying confounding: | | |
| 1.2. Was the analysis based on splitting participants' follow up time according to intervention received?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, go to question 1.3. | No | NA / Y / PY / PN / N / NI |
| 1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8) | | NA / Y / PY / PN / N / NI |
| **Questions relating to baseline confounding only** | | |
| 1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains? | Yes – 2 sided t tests were used for continuous variables where appropriate. | NA / Y / PY / PN / N / NI |
| 1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | No | NA / Y / PY / PN / N / NI |
| 1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention? | No | NA / Y / PY / PN / N / NI |
| **Questions relating to baseline and time-varying confounding** | | |
| 1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding? | Yes | NA / Y / PY / PN / N / NI |
| 1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | Yes | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to confounding? | | Favours experimental / Favours comparator / Unpredictable |

| Bias in selection of participants into the study | | |
|---|---|---|
| 2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? | No | Y / PY / PN / N / NI |
| **If N/PN to 2.1:** go to 2.4 | | |
| 2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention? | | NA / Y / PY / PN / N / NI |
| 2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome? | | NA / Y / PY / PN / N / NI |
| 2.4. Do start of follow-up and start of intervention coincide for most participants? | Yes | Y / PY / PN / N / NI |
| 2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to selection of participants into the study? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in classification of interventions | | |
|---|---|---|
| 3.1 Were intervention groups clearly defined? | Yes | Y / PY / PN / N / NI |
| 3.2 Was the information used to define intervention groups recorded at the start of the intervention? | Yes | Y / PY / PN / N / NI |
| 3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to classification of interventions? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias due to deviations from intended interventions | | |
|---|---|---|
| **If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2** | | |
| 4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice? | No | Y / PY / PN / N / NI |
| 4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome? | | NA / Y / PY / PN / N / NI |
| **If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6** | | |
| 4.3. Were important co-interventions balanced across intervention groups? | | Y / PY / PN / N / NI |
| 4.4. Was the intervention implemented successfully for most participants? | | Y / PY / PN / N / NI |
| 4.5. Did study participants adhere to the assigned intervention regimen? | | Y / PY / PN / N / NI |
| 4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | |
| Optional: What is the predicted direction of bias due to deviations from the intended interventions? | | |

| Bias due to missing data | | |
|---|---|---|
| 5.1 Were outcome data available for all, or nearly all, participants? | Yes | Y / PY / PN / N / NI |
| 5.2 Were participants excluded due to missing data on intervention status? | No | Y / PY / PN / N / NI |
| 5.3 Were participants excluded due to missing data on other variables needed for the analysis? | No | Y / PY / PN / N / NI |
| 5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions? | | NA / Y / PY / PN / N / NI |
| 5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to missing data? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in measurement of outcomes | | |
|---|---|---|
| 6.1 Could the outcome measure have been influenced by knowledge of the intervention received? | No | Y / PY / PN / N / NI |
| 6.2 Were outcome assessors aware of the intervention received by study participants? | Yes – non randomized study, non blinded. | Y / PY / PN / N / NI |
| 6.3 Were the methods of outcome assessment comparable across intervention groups? | Yes | Y / PY / PN / N / NI |
| 6.4 Were any systematic errors in measurement of the outcome related to intervention received? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to measurement of outcomes? | Favours experimental | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in selection of the reported result | | |
|---|---|---|
| Is the reported effect estimate likely to be selected, on the basis of the results, from... | No | |
| 7.1. ... multiple outcome *measurements* within the outcome domain? | | Y / PY / PN / N / NI |
| 7.2 ... multiple *analyses* of the intervention-outcome relationship? | No | Y / PY / PN / N / NI |
| 7.3 ... different *subgroups*? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to selection of the reported result? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Overall bias | | |
|---|---|---|
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the overall predicted direction of bias for this outcome? | Unpredictable | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

Appendix 10 – Risk of Bias Assessment (ROBINS-I) tool – Lee et. al. 2010

**Risk of bias assessment**

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

| Signalling questions | Description | Response options |
|---|---|---|
| **Bias due to confounding** | | |
| 1.1 Is there potential for confounding of the effect of intervention in this study?<br>**If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered | Yes – Patients were dissimilar at baseline with respect to insulin use along with other baseline characteristics. | Y / PY / PN / N |
| **If Y/PY to 1.1**: determine whether there is a need to assess time-varying confounding: | | |
| 1.2. Was the analysis based on splitting participants' follow up time according to intervention received?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, go to question 1.3. | No | NA / Y / PY / PN / N / NI |
| 1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8) | | NA / Y / PY / PN / N / NI |
| **Questions relating to baseline confounding only** | | |
| 1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains? | Yes – Appropriate statistical analysis with chi square, pearson correlation and 1-way repeated analysis of variance. Continuous variables expressed as mean and standard deviation and compared using bonferroni test. | NA / Y / PY / PN / N / NI |
| 1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | Yes | NA / Y / PY / PN / N / NI |
| 1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention? | No | NA / Y / PY / PN / N / NI |
| **Questions relating to baseline and time-varying confounding** | | |
| 1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding? | Yes – Patients were followed up at multiple points in time | NA / Y / PY / PN / N / NI |
| 1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | Yes | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to confounding? | | Favours experimental / Favours comparator / Unpredictable |

| Bias in selection of participants into the study | | |
|---|---|---|
| 2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? | No | Y / PY / PN / N / NI |
| **If N/PN to 2.1:** go to 2.4 | | |
| 2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention? | | NA / Y / PY / PN / N / NI |
| 2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome? | | NA / Y / PY / PN / N / NI |
| 2.4. Do start of follow-up and start of intervention coincide for most participants? | Yes | Y / PY / PN / N / NI |
| 2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to selection of participants into the study? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in classification of interventions | | |
|---|---|---|
| 3.1 Were intervention groups clearly defined? | Yes | Y / PY / PN / N / NI |
| 3.2 Was the information used to define intervention groups recorded at the start of the intervention? | Yes | Y / PY / PN / N / NI |
| 3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to classification of interventions? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias due to deviations from intended interventions | | |
|---|---|---|
| **If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2** | | |
| 4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice? | No | Y / PY / PN / N / NI |
| 4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome? | | NA / Y / PY / PN / N / NI |
| **If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6** | | |
| 4.3. Were important co-interventions balanced across intervention groups? | | Y / PY / PN / N / NI |
| 4.4. Was the intervention implemented successfully for most participants? | | Y / PY / PN / N / NI |
| 4.5. Did study participants adhere to the assigned intervention regimen? | | Y / PY / PN / N / NI |
| 4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | |
| Optional: What is the predicted direction of bias due to deviations from the intended interventions? | | |

| Bias due to missing data | | |
|---|---|---|
| 5.1 Were outcome data available for all, or nearly all, participants? | Yes | Y / PY / PN / N / NI |
| 5.2 Were participants excluded due to missing data on intervention status? | No | Y / PY / PN / N / NI |
| 5.3 Were participants excluded due to missing data on other variables needed for the analysis? | No | Y / PY / PN / N / NI |
| 5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions? | | NA / Y / PY / PN / N / NI |
| 5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to missing data? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in measurement of outcomes | | |
|---|---|---|
| 6.1 Could the outcome measure have been influenced by knowledge of the intervention received? | No – Prospective cohort design | Y / PY / PN / N / NI |
| 6.2 Were outcome assessors aware of the intervention received by study participants? | Yes – assesors were not blinded given study type | Y / PY / PN / N / NI |
| 6.3 Were the methods of outcome assessment comparable across intervention groups? | Yes | Y / PY / PN / N / NI |
| 6.4 Were any systematic errors in measurement of the outcome related to intervention received? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to measurement of outcomes? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in selection of the reported result | | |
|---|---|---|
| Is the reported effect estimate likely to be selected, on the basis of the results, from... | No | |
| 7.1. ... multiple outcome *measurements* within the outcome domain? | | Y / PY / PN / N / NI |
| 7.2 ... multiple *analyses* of the intervention-outcome relationship? | No | Y / PY / PN / N / NI |
| 7.3 ... different *subgroups*? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to selection of the reported result? | Low | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Overall bias | | |
|---|---|---|
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the overall predicted direction of bias for this outcome? | Favours experimental group | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

Appendix 11 – Risk of Bias Assessment (ROBINS-I) tool – Parikh et. al. 2006

**Risk of bias assessment**

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

| Signalling questions | Description | Response options |
|---|---|---|
| **Bias due to confounding** | | |
| 1.1 Is there potential for confounding of the effect of intervention in this study?<br>**If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered | Yes – study design and differences amongst study participants will contribute to confounding effect. | Y / PY / PN / N |
| **If Y/PY to 1.1**: determine whether there is a need to assess time-varying confounding: | | |
| 1.2. Was the analysis based on splitting participants' follow up time according to intervention received?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, go to question 1.3. | No | NA / Y / PY / PN / N / NI |
| 1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8) | | NA / Y / PY / PN / N / NI |
| **Questions relating to baseline confounding only** | | |
| 1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains? | No - data was given on BMI and EWL however no statistical analysis was performed to compare treatment and control groups. | NA / Y / PY / PN / N / NI |
| 1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | | NA / Y / PY / PN / N / NI |
| 1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention? | No | NA / Y / PY / PN / N / NI |
| **Questions relating to baseline and time-varying confounding** | | |
| 1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding? | No | NA / Y / PY / PN / N / NI |
| 1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Serious | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to confounding? | Likely favour experimental | Favours experimental / Favours comparator / Unpredictable |

| Bias in selection of participants into the study | | |
|---|---|---|
| 2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? | No | Y / PY / PN / N / NI |
| **If N/PN to 2.1:** go to 2.4 | | |
| 2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention? | | NA / Y / PY / PN / N / NI |
| 2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome? | | NA / Y / PY / PN / N / NI |
| 2.4. Do start of follow-up and start of intervention coincide for most participants? | Yes | Y / PY / PN / N / NI |
| 2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to selection of participants into the study? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in classification of interventions | | |
|---|---|---|
| 3.1 Were intervention groups clearly defined? | Yes | Y / PY / PN / N / NI |
| 3.2 Was the information used to define intervention groups recorded at the start of the intervention? | Yes | Y / PY / PN / N / NI |
| 3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to classification of interventions? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias due to deviations from intended interventions | | |
|---|---|---|
| **If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2** | | |
| 4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice? | No | Y / PY / PN / N / NI |
| 4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome? | | NA / Y / PY / PN / N / NI |
| **If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6** | | |
| 4.3. Were important co-interventions balanced across intervention groups? | | Y / PY / PN / N / NI |
| 4.4. Was the intervention implemented successfully for most participants? | | Y / PY / PN / N / NI |
| 4.5. Did study participants adhere to the assigned intervention regimen? | | Y / PY / PN / N / NI |
| 4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | |
| Optional: What is the predicted direction of bias due to deviations from the intended interventions? | | |

| Bias due to missing data | | |
|---|---|---|
| 5.1 Were outcome data available for all, or nearly all, participants? | Yes | Y / PY / PN / N / NI |
| 5.2 Were participants excluded due to missing data on intervention status? | No | Y / PY / PN / N / NI |
| 5.3 Were participants excluded due to missing data on other variables needed for the analysis? | No | Y / PY / PN / N / NI |
| 5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions? | | NA / Y / PY / PN / N / NI |
| 5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to missing data? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in measurement of outcomes | | |
|---|---|---|
| 6.1 Could the outcome measure have been influenced by knowledge of the intervention received? | No | Y / PY / PN / N / NI |
| 6.2 Were outcome assessors aware of the intervention received by study participants? | Yes – non randomized study, non blinded. | Y / PY / PN / N / NI |
| 6.3 Were the methods of outcome assessment comparable across intervention groups? | Yes | Y / PY / PN / N / NI |
| 6.4 Were any systematic errors in measurement of the outcome related to intervention received? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to measurement of outcomes? | Favours experimental | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in selection of the reported result | | |
|---|---|---|
| Is the reported effect estimate likely to be selected, on the basis of the results, from... | No | |
| 7.1. ... multiple outcome *measurements* within the outcome domain? | | Y / PY / PN / N / NI |
| 7.2 ... multiple *analyses* of the intervention-outcome relationship? | No | Y / PY / PN / N / NI |
| 7.3 ... different *subgroups*? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to selection of the reported result? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Overall bias | | |
|---|---|---|
| **Risk of bias judgement** | Serious | Low / Moderate / Serious / Critical / NI |
| Optional: What is the overall predicted direction of bias for this outcome? | Likely experimental | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

Appendix 12 – Risk of Bias Assessment (ROBINS-I) tool – Scopinaro et. al. 2014

**Risk of bias assessment**

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

| Signalling questions | Description | Response options |
|---|---|---|
| **Bias due to confounding** | | |
| 1.1 Is there potential for confounding of the effect of intervention in this study?<br>**If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered | Yes – due to nature of study design and some dissimilarities between study participants. | Y / PY / PN / N |
| **If Y/PY to 1.1**: determine whether there is a need to assess time-varying confounding: | | |
| 1.2. Was the analysis based on splitting participants' follow up time according to intervention received?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, go to question 1.3. | No | NA / Y / PY / PN / N / NI |
| 1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8) | | NA / Y / PY / PN / N / NI |
| **Questions relating to baseline confounding only** | | |
| 1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains? | Yes – Wilcoxon signed rank test for paired group comparisons. Percentages compared with $X^2$ test and simple regression used for analysis of correlations between data. | NA / Y / PY / PN / N / NI |
| 1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | Yes | NA / Y / PY / PN / N / NI |
| 1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention? | No | NA / Y / PY / PN / N / NI |
| **Questions relating to baseline and time-varying confounding** | | |
| 1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding? | Yes | NA / Y / PY / PN / N / NI |
| 1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | Yes | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to confounding? | Unpredictable | Favours experimental / Favours comparator / Unpredictable |

| Bias in selection of participants into the study | | |
|---|---|---|
| 2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? | No | Y / PY / PN / N / NI |
| **If N/PN to 2.1:** go to 2.4 | | |
| 2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention? | | NA / Y / PY / PN / N / NI |
| 2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome? | | NA / Y / PY / PN / N / NI |
| 2.4. Do start of follow-up and start of intervention coincide for most participants? | Yes | Y / PY / PN / N / NI |
| 2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to selection of participants into the study? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in classification of interventions | | |
|---|---|---|
| 3.1 Were intervention groups clearly defined? | Yes | Y / PY / PN / N / NI |
| 3.2 Was the information used to define intervention groups recorded at the start of the intervention? | Yes | Y / PY / PN / N / NI |
| 3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to classification of interventions? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias due to deviations from intended interventions | | |
|---|---|---|
| **If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2** | | |
| 4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice? | No | Y / PY / PN / N / NI |
| 4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome? | | NA / Y / PY / PN / N / NI |
| **If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6** | | |
| 4.3. Were important co-interventions balanced across intervention groups? | | Y / PY / PN / N / NI |
| 4.4. Was the intervention implemented successfully for most participants? | | Y / PY / PN / N / NI |
| 4.5. Did study participants adhere to the assigned intervention regimen? | | Y / PY / PN / N / NI |
| 4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | |
| Optional: What is the predicted direction of bias due to deviations from the intended interventions? | | |

| Bias due to missing data | | |
|---|---|---|
| 5.1 Were outcome data available for all, or nearly all, participants? | Yes | Y / PY / PN / N / NI |
| 5.2 Were participants excluded due to missing data on intervention status? | No | Y / PY / PN / N / NI |
| 5.3 Were participants excluded due to missing data on other variables needed for the analysis? | No | Y / PY / PN / N / NI |
| 5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions? | | NA / Y / PY / PN / N / NI |
| 5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to missing data? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in measurement of outcomes | | |
|---|---|---|
| 6.1 Could the outcome measure have been influenced by knowledge of the intervention received? | No | Y / PY / PN / N / NI |
| 6.2 Were outcome assessors aware of the intervention received by study participants? | Yes – non blinded, non randomized study | Y / PY / PN / N / NI |
| 6.3 Were the methods of outcome assessment comparable across intervention groups? | Yes | Y / PY / PN / N / NI |
| 6.4 Were any systematic errors in measurement of the outcome related to intervention received? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to measurement of outcomes? | Unpredictable | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in selection of the reported result | | |
|---|---|---|
| Is the reported effect estimate likely to be selected, on the basis of the results, from... | | |
| 7.1. ... multiple outcome *measurements* within the outcome domain? | No | Y / PY / PN / N / NI |
| 7.2 ... multiple *analyses* of the intervention-outcome relationship? | No | Y / PY / PN / N / NI |
| 7.3 ... different *subgroups*? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to selection of the reported result? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Overall bias | | |
|---|---|---|
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the overall predicted direction of bias for this outcome? | Unpredictable | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

Appendix 13 – Risk of Bias Assessment (ROBINS-I) tool – Sultan et. al. 2008

**Risk of bias assessment**

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

| Signalling questions | Description | Response options |
|---|---|---|
| **Bias due to confounding** | | |
| 1.1 Is there potential for confounding of the effect of intervention in this study?<br>**If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered | Yes – study design and differences amongst study participants will contribute to confounding effect. | Y / PY / PN / N |
| **If Y/PY to 1.1**: determine whether there is a need to assess time-varying confounding: | | |
| 1.2. Was the analysis based on splitting participants' follow up time according to intervention received?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, go to question 1.3. | No | NA / Y / PY / PN / N / NI |
| 1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8) | | NA / Y / PY / PN / N / NI |
| **Questions relating to baseline confounding only** | | |
| 1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains? | No - Descriptive statistics were given only with means and standard deviation. No statistical analysis to determine if any statistical significance was performed. | NA / Y / PY / PN / N / NI |
| 1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | | NA / Y / PY / PN / N / NI |
| 1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention? | No | NA / Y / PY / PN / N / NI |
| **Questions relating to baseline and time-varying confounding** | | |
| 1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding? | No | NA / Y / PY / PN / N / NI |
| 1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Serious | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to confounding? | Likely Favours experimental | Favours experimental / Favours comparator / Unpredictable |

| Bias in selection of participants into the study | | |
|---|---|---|
| 2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? | No | Y / PY / PN / N / NI |
| **If N/PN to 2.1:** go to 2.4 | | |
| 2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention? | | NA / Y / PY / PN / N / NI |
| 2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome? | | NA / Y / PY / PN / N / NI |
| 2.4. Do start of follow-up and start of intervention coincide for most participants? | Yes | Y / PY / PN / N / NI |
| 2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to selection of participants into the study? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in classification of interventions | | |
|---|---|---|
| 3.1 Were intervention groups clearly defined? | Yes | Y / PY / PN / N / NI |
| 3.2 Was the information used to define intervention groups recorded at the start of the intervention? | Yes | Y / PY / PN / N / NI |
| 3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to classification of interventions? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias due to deviations from intended interventions | | |
|---|---|---|
| **If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2** | | |
| 4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice? | No | Y / PY / PN / N / NI |
| 4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome? | | NA / Y / PY / PN / N / NI |
| **If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6** | | |
| 4.3. Were important co-interventions balanced across intervention groups? | | Y / PY / PN / N / NI |
| 4.4. Was the intervention implemented successfully for most participants? | | Y / PY / PN / N / NI |
| 4.5. Did study participants adhere to the assigned intervention regimen? | | Y / PY / PN / N / NI |
| 4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | |
| Optional: What is the predicted direction of bias due to deviations from the intended interventions? | | |

| Bias due to missing data | | |
|---|---|---|
| 5.1 Were outcome data available for all, or nearly all, participants? | Yes | Y / PY / PN / N / NI |
| 5.2 Were participants excluded due to missing data on intervention status? | No | Y / PY / PN / N / NI |
| 5.3 Were participants excluded due to missing data on other variables needed for the analysis? | No | Y / PY / PN / N / NI |
| 5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions? | | NA / Y / PY / PN / N / NI |
| 5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to missing data? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in measurement of outcomes | | |
|---|---|---|
| 6.1 Could the outcome measure have been influenced by knowledge of the intervention received? | No | Y / PY / PN / N / NI |
| 6.2 Were outcome assessors aware of the intervention received by study participants? | Yes – non randomized study, non blinded. | Y / PY / PN / N / NI |
| 6.3 Were the methods of outcome assessment comparable across intervention groups? | Yes | Y / PY / PN / N / NI |
| 6.4 Were any systematic errors in measurement of the outcome related to intervention received? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to measurement of outcomes? | Favours experimental | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in selection of the reported result | | |
|---|---|---|
| Is the reported effect estimate likely to be selected, on the basis of the results, from... | No | |
| 7.1. ... multiple outcome *measurements* within the outcome domain? | | Y / PY / PN / N / NI |
| 7.2 ... multiple *analyses* of the intervention-outcome relationship? | No | Y / PY / PN / N / NI |
| 7.3 ... different *subgroups*? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to selection of the reported result? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Overall bias | | |
|---|---|---|
| **Risk of bias judgement** | Serious | Low / Moderate / Serious / Critical / NI |
| Optional: What is the overall predicted direction of bias for this outcome? | Likely favours experimental | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

Appendix 14 – Risk of Bias Assessment (ROBINS-I) tool – Zhu et. al. 2012

**Risk of bias assessment**

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sig
posts to other questions, no formatting is used.

| Signalling questions | Description | Response options |
|---|---|---|
| **Bias due to confounding** | | |
| 1.1 Is there potential for confounding of the effect of intervention in this study?<br>**If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered | Yes – study design and differences amongst study participants will contribute to confounding effect. | Y / PY / PN / N |
| **If Y/PY to 1.1**: determine whether there is a need to assess time-varying confounding: | | |
| 1.2. Was the analysis based on splitting participants' follow up time according to intervention received?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, go to question 1.3. | No | NA / Y / PY / PN / N / NI |
| 1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8) | | NA / Y / PY / PN / N / NI |
| **Questions relating to baseline confounding only** | | |
| 1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains? | Yes – Student's t test and $X^2$ test were utilized | NA / Y / PY / PN / N / NI |
| 1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | No | NA / Y / PY / PN / N / NI |
| 1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention? | No | NA / Y / PY / PN / N / NI |
| **Questions relating to baseline and time-varying confounding** | | |
| 1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding? | Yes | NA / Y / PY / PN / N / NI |
| 1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | Yes | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to confounding? | | Favours experimental / Favours comparator / Unpredictable |

| Bias in selection of participants into the study | | |
|---|---|---|
| 2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? | No | Y / PY / PN / N / NI |
| **If N/PN to 2.1:** go to 2.4 | | |
| 2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention? | | NA / Y / PY / PN / N / NI |
| 2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome? | | NA / Y / PY / PN / N / NI |
| 2.4. Do start of follow-up and start of intervention coincide for most participants? | Yes | Y / PY / PN / N / NI |
| 2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to selection of participants into the study? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in classification of interventions | | |
|---|---|---|
| 3.1 Were intervention groups clearly defined? | Yes | Y / PY / PN / N / NI |
| 3.2 Was the information used to define intervention groups recorded at the start of the intervention? | Yes | Y / PY / PN / N / NI |
| 3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to classification of interventions? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias due to deviations from intended interventions | | |
|---|---|---|
| **If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2** | | |
| 4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice? | Yes – Based on the author's description the gastric bypass did not utilize small gastric pouch but rather the intestine was anastamosed to the body of essentially native stomach. | Y / PY / PN / N / NI |
| 4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome? | No | NA / Y / PY / PN / N / NI |
| **If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6** | | |
| 4.3. Were important co-interventions balanced across intervention groups? | | Y / PY / PN / N / NI |
| 4.4. Was the intervention implemented successfully for most participants? | | Y / PY / PN / N / NI |
| 4.5. Did study participants adhere to the assigned intervention regimen? | | Y / PY / PN / N / NI |
| 4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Moderate | |
| Optional: What is the predicted direction of bias due to deviations from the intended interventions? | | |

| Bias due to missing data | | |
|---|---|---|
| 5.1 Were outcome data available for all, or nearly all, participants? | Yes | Y / PY / PN / N / NI |
| 5.2 Were participants excluded due to missing data on intervention status? | No | Y / PY / PN / N / NI |
| 5.3 Were participants excluded due to missing data on other variables needed for the analysis? | No | Y / PY / PN / N / NI |
| 5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions? | | NA / Y / PY / PN / N / NI |
| 5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data? | | NA / Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to missing data? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in measurement of outcomes | | |
|---|---|---|
| 6.1 Could the outcome measure have been influenced by knowledge of the intervention received? | No | Y / PY / PN / N / NI |
| 6.2 Were outcome assessors aware of the intervention received by study participants? | Yes – non randomized study, non blinded. | Y / PY / PN / N / NI |
| 6.3 Were the methods of outcome assessment comparable across intervention groups? | Yes | Y / PY / PN / N / NI |
| 6.4 Were any systematic errors in measurement of the outcome related to intervention received? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to measurement of outcomes? | Favours experimental | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Bias in selection of the reported result | | |
|---|---|---|
| Is the reported effect estimate likely to be selected, on the basis of the results, from... | No | |
| 7.1. ... multiple outcome *measurements* within the outcome domain? | | Y / PY / PN / N / NI |
| 7.2 ... multiple *analyses* of the intervention-outcome relationship? | No | Y / PY / PN / N / NI |
| 7.3 ... different *subgroups*? | No | Y / PY / PN / N / NI |
| **Risk of bias judgement** | Low | Low / Moderate / Serious / Critical / NI |
| Optional: What is the predicted direction of bias due to selection of the reported result? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| Overall bias | | |
|---|---|---|
| **Risk of bias judgement** | Moderate | Low / Moderate / Serious / Critical / NI |
| Optional: What is the overall predicted direction of bias for this outcome? | Unpredictable | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

Appendix 15 – Risk of Bias Assessment (ROB 2.0) tool – Abbatini et. al. 2012

| Domain | Signalling questions | Response options | Description/Support for judgement |
|---|---|---|---|
| **Bias arising from the randomization process** | 1.1 Was the allocation sequence random? | Y / PY / PN / N / NI | Yes |
| | 1.2 Was the allocation sequence concealed until participants were recruited and assigned to interventions? | Y / PY / PN / N / NI | NI |
| | 1.3 Were there baseline imbalances that suggest a problem with the randomization process? | Y / PY / PN / N / NI | No - Stratified randomization given low numbers of patients, matched for BMI, HbA1c, cpeptide levels. |
| | **Risk of bias judgement** | Low / High / Some concerns | Low |
| | Optional: What is the predicted direction of bias arising from the randomization process? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | |
| **Bias due to deviations from intended interventions** | 2.1. Were participants aware of their assigned intervention during the trial? | Y / PY / PN / N / NI | NI |
| | 2.2. Were carers and trial personnel aware of participants' assigned intervention during the trial? | Y / PY / PN / N / NI | NI |
| | 2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention beyond what would be expected in usual practice? | NA / Y / PY / PN / N / NI | No |
| | 2.4. If Y/PY to 2.3: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome? | NA / Y / PY / PN / N / NI | |
| | 2.5 Were any participants analysed in a group different from the one to which they were assigned? | Y / PY / PN / N / NI | No |
| | 2.6 If Y/PY/NI to 2.5: Was there potential for a substantial impact (on the estimated effect of intervention) of analysing participants in the wrong group? | NA / Y / PY / PN / N / NI | |

| | | | |
|---|---|---|---|
| | **Risk of bias judgement** | Low / High / Some concerns | Some concerns |
| | Optional: What is the predicted direction of bias due to deviations from intended interventions? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | |
| **Bias due to missing outcome data** | 3.1 Were outcome data available for all, or nearly all, participants randomized? | Y / PY / PN / N / NI | Yes |
| | 3.2 If N/PN/NI to 3.1: Are the proportions of missing outcome data and reasons for missing outcome data similar across intervention groups? | NA / Y / PY / PN / N / NI | |
| | 3.3 If N/PN/NI to 3.1: Is there evidence that results were robust to the presence of missing outcome data? | NA / Y / PY / PN / N / NI | |
| | **Risk of bias judgement** | Low / High / Some concerns | Low |
| | Optional: What is the predicted direction of bias due to missing outcome data? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | |
| **Bias in measurement of the outcome** | 4.1 Were outcome assessors aware of the intervention received by study participants? | Y / PY / PN / N / NI | NI |
| | 4.2 If Y/PY/NI to 4.1: Was the assessment of the outcome likely to be influenced by knowledge of intervention received? | NA / Y / PY / PN / N / NI | No |
| | **Risk of bias judgement** | Low / High / Some concerns | Low |
| | Optional: What is the predicted direction of bias due to measurement of the outcome? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | Unpredictable |
| **Bias in selection of** | Are the reported outcome data likely to have been selected, on the basis of the results, from... | | |

| | | | |
|---|---|---|---|
| **the reported result** | 5.1. ... multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain? | Y / PY / PN / N / NI | No |
| | 5.2 ... multiple analyses of the data? | Y / PY / PN / N / NI | No |
| | **Risk of bias judgement** | Low / High / Some concerns | Low |
| | Optional: What is the predicted direction of bias due to selection of the reported result? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | Unpredictable |
| **Overall bias** | **Risk of bias judgement** | Low / High / Some concerns | Some concerns |
| | Optional:<br>What is the overall predicted direction of bias for this outcome? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | Unpredictable. Details regarding randomization and blinding of participants were lacking in this study although given one of the treatment groups was surgical blinding may have been difficult. |

Appendix 16 – Risk of Bias Assessment (ROB 2.0) tool - Lee et. al. 2011

| Domain | Signalling questions | Response options | Description/Support for judgement |
|---|---|---|---|
| **Bias arising from the randomization process** | 1.1 Was the allocation sequence random? | Y / PY / PN / N / NI | Y |
| | 1.2 Was the allocation sequence concealed until participants were recruited and assigned to interventions? | Y / PY / PN / N / NI | Y |
| | 1.3 Were there baseline imbalances that suggest a problem with the randomization process? | Y / PY / PN / N / NI | N |
| | **Risk of bias judgement** | Low / High / Some concerns | Low |
| | Optional: What is the predicted direction of bias arising from the randomization process? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | |
| **Bias due to deviations from intended interventions** | 2.1. Were participants aware of their assigned intervention during the trial? | Y / PY / PN / N / NI | N |
| | 2.2. Were carers and trial personnel aware of participants' assigned intervention during the trial? | Y / PY / PN / N / NI | N |
| | 2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention beyond what would be expected in usual practice? | NA / Y / PY / PN / N / NI | N |
| | 2.4. If Y/PY to 2.3: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome? | NA / Y / PY / PN / N / NI | |
| | 2.5 Were any participants analysed in a group different from the one to which they were assigned? | Y / PY / PN / N / NI | N |
| | 2.6 If Y/PY/NI to 2.5: Was there potential for a substantial impact (on the estimated effect of intervention) of analysing participants in the wrong group? | NA / Y / PY / PN / N / NI | |

| | | | |
|---|---|---|---|
| | **Risk of bias judgement** | Low / High / Some concerns | Low |
| | Optional: What is the predicted direction of bias due to deviations from intended interventions? | Favours experimental / Favours comparator / Towards null / Away from null / Unpredictable | |
| **Bias due to missing outcome data** | 3.1 Were outcome data available for all, or nearly all, participants randomized? | Y / PY / PN / N / NI | Y |
| | 3.2 If N/PN/NI to 3.1: Are the proportions of missing outcome data and reasons for missing outcome data similar across intervention groups? | NA / Y / PY / PN / N / NI | |
| | 3.3 If N/PN/NI to 3.1: Is there evidence that results were robust to the presence of missing outcome data? | NA / Y / PY / PN / N / NI | |
| | **Risk of bias judgement** | Low / High / Some concerns | Low |
| | Optional: What is the predicted direction of bias due to missing outcome data? | Favours experimental / Favours comparator / Towards null / Away from null / Unpredictable | |
| **Bias in measurement of the outcome** | 4.1 Were outcome assessors aware of the intervention received by study participants? | Y / PY / PN / N / NI | N |
| | 4.2 If Y/PY/NI to 4.1: Was the assessment of the outcome likely to be influenced by knowledge of intervention received? | NA / Y / PY / PN / N / NI | N |
| | **Risk of bias judgement** | Low / High / Some concerns | Low |
| | Optional: What is the predicted direction of bias due to measurement of the outcome? | Favours experimental / Favours comparator / Towards null / Away from null / Unpredictable | |
| **Bias in selection of** | Are the reported outcome data likely to have been selected, on the basis of the results, from... | | |

| | | | |
|---|---|---|---|
| **the reported result** | 5.1. ... multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain? | Y / PY / PN / N / NI | N |
| | 5.2 ... multiple analyses of the data? | Y / PY / PN / N / NI | N |
| | **Risk of bias judgement** | Low / High / Some concerns | Low |
| | Optional: What is the predicted direction of bias due to selection of the reported result? | Favours experimental / Favours comparator / Towards null / Away from null / Unpredictable | |
| **Overall bias** | **Risk of bias judgement** | Low / High / Some concerns | Low |
| | Optional:<br>What is the overall predicted direction of bias for this outcome? | Favours experimental / Favours comparator / Towards null / Away from null / Unpredictable | |

Appendix 17 – Risk of Bias Assessment (ROB 2.0) tool – Obrien et. al. 2006

| Domain | Signalling questions | Response options | Description/Support for judgement |
|---|---|---|---|
| **Bias arising from the randomization process** | 1.1 Was the allocation sequence random? | Y / PY / PN / N / NI | Y |
| | 1.2 Was the allocation sequence concealed until participants were recruited and assigned to interventions? | Y / PY / PN / N / NI | Y |
| | 1.3 Were there baseline imbalances that suggest a problem with the randomization process? | Y / PY / PN / N / NI | N |
| | **Risk of bias judgement** | Low / High / Some concerns | Low |
| | Optional: What is the predicted direction of bias arising from the randomization process? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | |
| **Bias due to deviations from intended interventions** | 2.1. Were participants aware of their assigned intervention during the trial? | Y / PY / PN / N / NI | NI |
| | 2.2. Were carers and trial personnel aware of participants' assigned intervention during the trial? | Y / PY / PN / N / NI | NI |
| | 2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention beyond what would be expected in usual practice? | NA / Y / PY / PN / N / NI | N |
| | 2.4. If Y/PY to 2.3: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome? | NA / Y / PY / PN / N / NI | |
| | 2.5 Were any participants analysed in a group different from the one to which they were assigned? | Y / PY / PN / N / NI | N |
| | 2.6 If Y/PY/NI to 2.5: Was there potential for a substantial impact (on the estimated effect of intervention) of analysing participants in the wrong group? | NA / Y / PY / PN / N / NI | |

| | | | |
|---|---|---|---|
| | **Risk of bias judgement** | Low / High / Some concerns | Some concerns |
| | Optional: What is the predicted direction of bias due to deviations from intended interventions? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | |
| **Bias due to missing outcome data** | 3.1 Were outcome data available for all, or nearly all, participants randomized? | Y / PY / PN / N / NI | Y |
| | 3.2 If N/PN/NI to 3.1: Are the proportions of missing outcome data and reasons for missing outcome data similar across intervention groups? | NA / Y / PY / PN / N / NI | |
| | 3.3 If N/PN/NI to 3.1: Is there evidence that results were robust to the presence of missing outcome data? | NA / Y / PY / PN / N / NI | |
| | **Risk of bias judgement** | Low / High / Some concerns | Low |
| | Optional: What is the predicted direction of bias due to missing outcome data? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | |
| **Bias in measurement of the outcome** | 4.1 Were outcome assessors aware of the intervention received by study participants? | Y / PY / PN / N / NI | NI |
| | 4.2 If Y/PY/NI to 4.1: Was the assessment of the outcome likely to be influenced by knowledge of intervention received? | NA / Y / PY / PN / N / NI | PN |
| | **Risk of bias judgement** | Low / High / Some concerns | Some Concerns |
| | Optional: What is the predicted direction of bias due to measurement of the outcome? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | Unpredictable |

| | | | |
|---|---|---|---|
| **Bias in selection of the reported result** | Are the reported outcome data likely to have been selected, on the basis of the results, from... | | |
| | 5.1. ... multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain? | Y / PY / PN / N / NI | N |
| | 5.2 ... multiple analyses of the data? | Y / PY / PN / N / NI | N |
| | **Risk of bias judgement** | Low / High / Some concerns | Low |
| | Optional: What is the predicted direction of bias due to selection of the reported result? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | |
| **Overall bias** | **Risk of bias judgement** | Low / High / Some concerns | Some concerns |
| | Optional:<br>What is the overall predicted direction of bias for this outcome? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | Unpredictable. Details were lacking however blinding would be difficult in this trial as one group was surgical and the other did not receive any surgical intervention. |

Appendix 18 – Risk of Bias Assessment (ROB 2.0) tool – Yang et. al. 2015

| Domain | Signalling questions | Response options | Description/Support for judgement |
|---|---|---|---|
| **Bias arising from the randomization process** | 1.1 Was the allocation sequence random? | Y / PY / PN / N / NI | Y |
| | 1.2 Was the allocation sequence concealed until participants were recruited and assigned to interventions? | Y / PY / PN / N / NI | Y |
| | 1.3 Were there baseline imbalances that suggest a problem with the randomization process? | Y / PY / PN / N / NI | N |
| | **Risk of bias judgement** | Low / High / Some concerns | Low |
| | Optional: What is the predicted direction of bias arising from the randomization process? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | |
| **Bias due to deviations from intended interventions** | 2.1. Were participants aware of their assigned intervention during the trial? | Y / PY / PN / N / NI | NI |
| | 2.2. Were carers and trial personnel aware of participants' assigned intervention during the trial? | Y / PY / PN / N / NI | NI |
| | 2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention beyond what would be expected in usual practice? | NA / Y / PY / PN / N / NI | N |
| | 2.4. If Y/PY to 2.3: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome? | NA / Y / PY / PN / N / NI | |
| | 2.5 Were any participants analysed in a group different from the one to which they were assigned? | Y / PY / PN / N / NI | N |
| | 2.6 If Y/PY/NI to 2.5: Was there potential for a substantial impact (on the estimated effect of intervention) of analysing participants in the wrong group? | NA / Y / PY / PN / N / NI | |

| | | | |
|---|---|---|---|
| | **Risk of bias judgement** | Low / High / Some concerns | Some concerns |
| | Optional: What is the predicted direction of bias due to deviations from intended interventions? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | |
| **Bias due to missing outcome data** | 3.1 Were outcome data available for all, or nearly all, participants randomized? | Y / PY / PN / N / NI | Y |
| | 3.2 If N/PN/NI to 3.1: Are the proportions of missing outcome data and reasons for missing outcome data similar across intervention groups? | NA / Y / PY / PN / N / NI | |
| | 3.3 If N/PN/NI to 3.1: Is there evidence that results were robust to the presence of missing outcome data? | NA / Y / PY / PN / N / NI | |
| | **Risk of bias judgement** | Low / High / Some concerns | Low |
| | Optional: What is the predicted direction of bias due to missing outcome data? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | |
| **Bias in measurement of the outcome** | 4.1 Were outcome assessors aware of the intervention received by study participants? | Y / PY / PN / N / NI | NI |
| | 4.2 If Y/PY/NI to 4.1: Was the assessment of the outcome likely to be influenced by knowledge of intervention received? | NA / Y / PY / PN / N / NI | PN |
| | **Risk of bias judgement** | Low / High / Some concerns | Some Concerns |
| | Optional: What is the predicted direction of bias due to measurement of the outcome? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | Unpredictable |
| **Bias in selection of** | Are the reported outcome data likely to have been selected, on the basis of the results, from... | | |

| | | | |
|---|---|---|---|
| **the reported result** | 5.1. ... multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain? | Y / PY / PN / N / NI | N |
| | 5.2 ... multiple analyses of the data? | Y / PY / PN / N / NI | N |
| | **Risk of bias judgement** | Low / High / Some concerns | Low |
| | Optional: What is the predicted direction of bias due to selection of the reported result? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | |
| **Overall bias** | **Risk of bias judgement** | Low / High / Some concerns | Some concerns |
| | Optional:<br>What is the overall predicted direction of bias for this outcome? | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable | Unpredictable. Details were lacking howeve blinding would be possible given both group were surgical. |