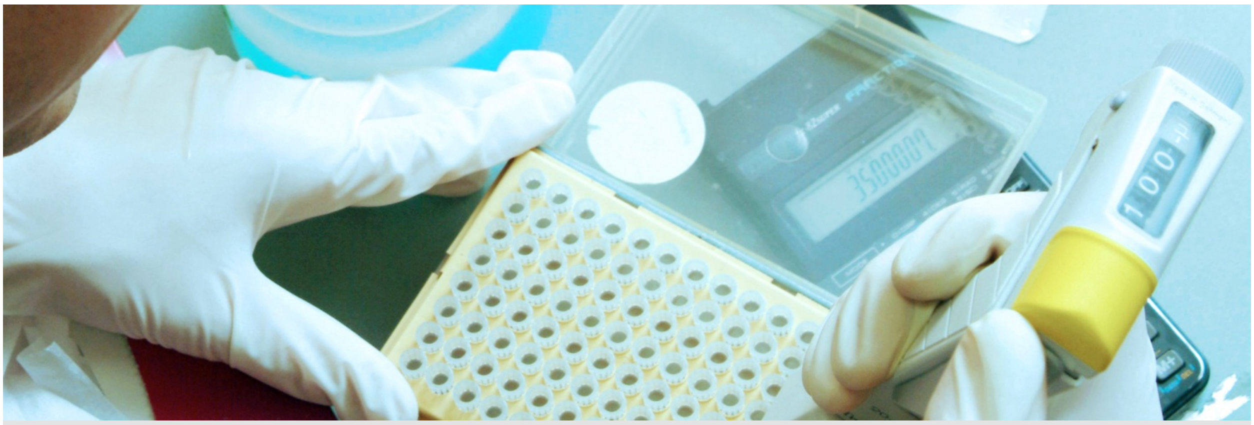
swiss medical board

Bariatric surgery vs. conservative treatment for obesity and overweight



Assessment

Final report 16 june 2016

Impressum

Swiss Medical Board
Stampfenbachstrasse 30
Postfach, 8090 Zürich

Geschäftsstelle
Susanna Marti C.

Telefon +41 43 259 5211
info@medical-board.ch
www.medical-board.ch

Assessment

The assessment team included the following persons (in alphabetical order)
Zanfina Ademi[1], Heiner C. Bucher[2], Dominik Glinz[2], Viktoria Gloy[2], Samia Hurst[3], Heike Raatz[2], Urs Saxer[4], Matthias Schwenkglenks[1,5], Joris Van Stiphout[1], Yuki Tomonaga[5]

Clinical Experts:

PD Dr. med. Ralph Peterli, Leitender Arzt Viszeralchirurgie, Claraspital, Basel, Switzerland
PD Dr. med. Igor Langer, V.E.T. Viszeral Experten Team AG, Praxis am Lindenhofspital, Bern, Switzerland
Dr. med. Fabian Meienberg, Leiter Endokrinologie-Sprechstunde, University Hospital Basel, Switzerland

The assessment team thanks the clinical experts for their support. The assessment team is responsible for the contents of this report.

Contributions:
Coordination: HR and MS; **Scoping:** HR, DG, MS, ZA and JVS; **Systematic review:** VG and HR wrote the protocol; VG, HR and DG developed and VG conducted the literature search; DG and VG screened the literature; DG, VG and HR were responsible for data extraction; DG, VG and HR graded the risk of bias and the quality of evidence; DG, VG and HR conducted data analysis; DG, VG and HR are responsible for data interpretation and wrote this systematic review; **Health economic analyses:** YT, MS and ZA wrote the protocol; JVS and YT conducted the literature search; JVS, YT screened the identified literature; YT, ZA and JVS were responsible for the data extraction and for the quality assessment of the literature; YT adapted the costs for Switzerland; ZA, YT, MS and JVS are responsible for data interpretation and wrote the systematic review; **Ethical domain:** SH elaborated the ethical domain and wrote the section about; **Legal domain:** US elaborated the legal domain and wrote the section about; **final report:** ZA, HCB, DG, VG, HR, MS, JVS and YT agreed with the final content of the final report.

[1]Universität Basel, Institut für Pharmazeutische Medizin
[2]Basel Institute for Clinical Epidemiology and Biostatistics (CEB), Universitätsspital Basel
[3]Universität Genf, Institut Ethique Histoire Humanités (iEH2)
[4] University of Zurich and Steinbrüchel Hüssy Rechtsanwälte, Zürich
[5]Universität Zürich, Institut für Epidemiologie, Biostatistik und Prävention (EBPI)

Contents

Abbreviations

AE	Adverse events
BAG	Bundesamt für Gesundheit
BMI	Body mass index
BPD	Biliopancreatic diversion
CHEERS	Consolidated Health Economic Evaluation Reporting Standards
CI	Confidence Interval
CRD	Centre for Review and Dissemination
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HTA	Health Technology Assessment
ICER	Incremental Cost-Effectiveness Ratio
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
LAGB	Laparoscopic adjustable gastric banding
MD	Mean difference
OKP	Obligatorische Krankenpflegeversicherung
OR	Odds Ratio
MD	Mean difference
N.A.	Not applicable
NHS EED	UK National Health Service's Economic Evaluation Database
n.r.	Not reported
RCT	Randomized controlled trial
RoB	Risk of bias
RR	Relative risk ratio
RYGB	Roux-en-Y gastric bypass
SAE	Serious adverse events
SD	Standard deviation
SG	Sleeve gastrectomy
SoF	Summary of Findings (GRADE output)
T2DM	Type 2 diabetes mellitus
VBG	Vertical banded gastroplasty

vs.	versus
WHO	World Health Organization

Tables

Figures

1 Executive summaries

1.1 Summary

1.1.1 Background

Obesity is a serious health problem associated with an increased risk for comorbidities, such as type 2 diabetes, hypertension, obstructive sleep apnoea and musculoskeletal disorders, and increased mortality. In Switzerland, the estimated health care costs due to obesity were 5,755 Mio. CHF in 2006 and 7,990 Mio. CHF in 2011. Of the Swiss population 30.8% are overweight (defined as body mass index, BMI 25.0-29.9 kg/m^2) and 10.3% are obese (BMI ≥30 kg/m^2). In Switzerland an adequate conservative treatment of obesity includes any combination of nutritional counselling by a nutritionist or physician, behavioural therapy including psychotherapy, diets to reduce caloric intake, physical therapy including physiotherapy or medication.

If adequate conservative treatment fails, bariatric (weight-loss) surgery can be considered. Conservative treatment is regarded as failed if in two years of conservative treatment or thereafter a BMI of <35 kg/m^2 cannot be reached and maintained. Swiss statutory health insurance ("obligatorische Krankenpflegeversicherung") covers bariatric surgery in obese individuals with a BMI ≥35 kg/m^2 independent of existing comorbidities. In Switzerland 4,167 bariatric operations were performed in 2014 showing a marked increase from 2001 when only 703 operations were performed. About 82.5% of all bariatric operations are gastric bypasses (Roux-en Y gastric bypass, RYGB), 15.7% are sleeve gastrectomies (SG), 0.8% are adjustable gastric bandings (AGB) and 0.7% are biliopancreatic diversions (BPD). Vertical banded gastroplasty (VBG) is covered but de facto no longer performed in Switzerland. Most of the operations are performed laparoscopically instead of open surgery. The RYGB involves surgical rearrangements of the stomach (the stomach becomes a smaller pouch) and small intestine limiting calorie intake and absorption of calories. In the case of AGB, an adjustable ring placed around the stomach allows only the ingestion of small food portions. The ring can be adjusted via an access port. In the case of SG the size of the stomach is reduced to a sleeve-like small pouch, while the pouch in the case of VBG is created using a (non-adjustable) band and staples.

While there is a relatively extensive international literature, there are only few studies for Switzerland which examine the overall cost of obesity. There is clearly no Swiss modelling of the cost-effectiveness of bariatric surgery or gastric bypass for overweight (BMI 25-29.9 kg/m^2) and obese (BMI ≥30 kg/m^2) adults.

1.1.2 Aim

The aim of this HTA report is to assess

- the effectiveness and safety
- the cost-effectiveness and budget impact analyses,
- legal as well as ethical implications

of bariatric surgery compared to conservative treatment, both in the population currently covered by the OKP (i.e. obese individuals with a BMI ≥35 kg/m^2) and in patients currently not covered by the OKP (i.e. overweight or obese individuals with a BMI of 25 - 35 kg/m^2). While all the surgical interventions currently used were being considered, the main focus was gastric bypass.

1.1.3 Clinical effectiveness and safety

For this HTA report the clinical effectiveness and safety of bariatric surgery compared to conservative treatment was assessed. The search was conducted in July 2015 and filters for randomised controlled studies (RCT) were used. The search was not limited for the time period, but was restricted to the languages English, German, French and Italian. Study characteristics and results of the included studies were presented per study in tables and summarized descriptively. The main focus of the analysis were the combined results after 2 or 3 years follow-up, but the combined results of the 6, 9 and 12 months follow-up were analysed and presented as well. Risk of bias and quality of evidence were assessed for included studies with 2 or 3 years follow-up as the primary focus was on long-term outcome. For additional information results after a follow-up of 6 months to 1 year were pooled as well but not evaluated according to GRADE. Risk of bias was assessed according to the Cochrane Handbook and the quality of evidence was assessed according to GRADE. Where possible, outcome results were summarized quantitatively in a meta-analysis by using inverse variance models assuming random effects. Effect estimates (summary and single for each trial) with corresponding 95% confidence interval were presented as forest plots. Relative risks were calculated for binary outcomes. Continuous outcomes were presented as mean differences. In case of considerable heterogeneity, methodological and clinical factors that might explain the heterogeneity were explored in subgroup and sensitivity analyses where possible. Some of the pre-specified subgroup analyses were the surgical method (i.e. gastric bypass vs. adjustable gastric banding vs. other surgical interventions) and patients with a BMI ≥35 kg/m^2 vs. patients with a BMI 25-34.9 kg/m^2.

Sixteen RCTs fulfilled the inclusion criteria with 10 RCTs reporting 2 or 3 year data. Most of the RCTs with 2 or 3 year data included patients with specific comorbidities (type 2 diabetes: 6 RCTs, sleep apnoea: 2 RCTs, mix of possible comorbidities: 1 RCT), only one RCT did not specify a comorbidity as inclusion criterion. Of the 13 outcomes assessed those classified as critical were body weight, quality of life, $HBA1_c$ (diabetes control), stroke and myocardial infarction.

The effects observed in these meta-analyses for the 2 to 3 years results were very consistent with a statistically significant benefit in the surgical arm compared to the conservative treatment for the following critical outcomes: percent body weight change, physical health (quality of life summary component score), HbA1c and diabetes remission; as well as for important outcomes such as hypertension and dyslipidaemia.

No statistically significant effect was shown for mental health (quality of life summary component score), stroke, all-cause mortality, serious adverse events (SAE) including reoperations, sleep apnoea, and cancer.

Data was extremely sparse (max. 1 event) or not existing for the critical outcomes stroke and the important outcomes mortality, myocardial infarction, and revision rates, therefore no conclusions can be drawn for these outcomes.

In the stratified presentation of results for the different surgical techniques or on BMI (BMI 25-34.9 kg/m^2 vs. BMI ≥35 kg/m^2), the direction of the observed effects – when they were present – consistently showed a benefit for bariatric surgery compared to conservative treatment. Results were stratified by surgical technique but effect estimates were not compared statistically by subgroup analyses because of the small number of studies within strata.

The quality of evidence was moderate for the critical outcome percentage body weight change and low for change in HbA1c. The overall quality of the evidence is judged to be very low because of the very low quality of evidence for the critical outcomes mental health and stroke and the absence of data from RCTs on myocardial infarction.

Significant uncertainties remain regarding outcomes in the long-term (>3 years) and the overweight population, where little or no data were available.

1.1.4 Cost-effectivenss and budget impact analyses

Relevant databases including Medline, Embase, the Cochrane Library, the Centre for Review and Dissemination (CRD) database, and the UK National Health Service's Economic Evaluation Database (NHS EED) were systematically searched for relevant health economic articles. Quality of reporting was assessed against the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) 24-item checklist, recommended by the ISPOR Health Economic Evaluations Publication Guidelines Task Force. For international studies, qualitative transferability to Switzerland was assessed as described in section 4.2.4. In the case of studies classified as qualitatively transferable to Switzerland, direct medical costs data were adapted to improve comparability in three distinct steps: correction for different levels of resource utilisation, for different prices of healthcare services, and for change in level of resource utilisation and prices over time. Subsequently, adapted incremental cost-effectiveness ratios (ICERs) were calculated. Budget impact has been addressed by combining Swiss epidemiological data with Swiss and adapted international cost data. Considering the limited data available, two approaches were pursued. Firstly, the potential surgery costs for treating all eligible Swiss patients in a single year were estimated. Secondly, the potential costs of bariatric surgery over a time period of ten years, from 2011 until 2020, were assessed from a Swiss healthcare system perspective.

Twenty-one cost-effectiveness analyses and one systematic review were finally included in this report and were assessed using the CHEERS checklist. About half of the studies were performed in Europe whereas the rest was conducted in the US or Australia; and vast majority assessed cost-effectiveness as cost per quality-adjusted life year (QALY) gained, over time horizons ranging from 5 years to lifelong. A total of 15 articles fulfilled criteria for qualitative transferability and underwent numerical adaptation of ICER results to Switzerland. Studies adopted healthcare, payer or societal perspectives. Their modelling of clinical effectiveness was based on short-term clinical trials and longer-term observational data (such as the SOS study).

All studies of patients with BMI values >35 kg/m^2 indicated bariatric surgery to be cost-saving or cost-effective, based on the criteria set by the authors. Adapted ICERs for Switzerland were below CHF 50,000 per QALY gained.

When studies like Borg et al. adopted a societal perspective, adapted ICERs for females aged 45-54 with BMI 40-44 kg/m^2 were below CHF 3,000 when gastric bypass was compared to conventional treatment. For males aged 45-54 with BMI 40-44 kg/m^2, gastric bypass was cost-saving and thus dominant. Michaud et al. also reported cost-effectiveness from a societal perspective, for US patients with BMI >40 kg/m^2 or with BMI >35 kg/m^2 and a high risk of co-morbidities. The adapted ICER for gastric bypass compared to conventional treatment was CHF 8,158 per life year gained.

For patients in the BMI class <35 kg/m^2, the adapted ICER from societal perspective was CHF 10,458 per QALY gained for females, and the adapted ICER for males was CHF 12,365 per QALY gained in the Borg et al, study. For BMI >35 kg/m^2 or BMI >30 kg/m^2 with qualifying co-morbidities, the adapted ICER for gastric bypass compared to conventional treatment reported by Michaud et al. was CHF 10,502 per life year gained.

Procedure-specific differences in terms of benefit and efficiency appear to exist. When gastric bypass and gastric banding were compared to conventional treatment, gastric bypass appeared to be better than gastric banding in terms of benefits (QALYs gained and life years saved), but also more expensive.

In the budget impact part, the first approach was to estimate the potential surgery costs of treating all eligible Swiss patients. In 2012 there were almost 25,000 persons in Switzerland with a BMI >40 kg/m^2. The costs for surgical treatment of all these individuals in a single year would reach CHF 353 Mio. In addition, if 10-30% of people in Switzerland with BMI 35-40 kg/m^2 were surgically treated due to the presence of relevant co-morbidities, the total costs of surgery would reach CHF 496-781 Mio. In 2014, the estimated total surgery costs for the 4,153 surgeries actually performed in Switzerland were CHF 61.1 Mio.

In the second approach, the budget impact estimation was based on a single, European cost-effectiveness study by Ackroyd et al., which primarily based the effectiveness of the intervention on the remission of diabetes. This selection was driven by the availability of sufficient numerical detail. Results over ten years strongly depended on the assumptions made. For example, if the diabetes prevalence in obese patients was 10%, the incremental costs of bariatric surgery in 2016 might reach CHF 55.1 Mio. In contrast, a diabetes prevalence in obese patients of 40% might make bariatric surgery cost-saving (CHF -1.2 Mio.).

Virtually all studies found bariatric surgery to be cost-saving or cost-effective, compared to conventional treatment. Adapted ICER results for patients with a BMI >35 kg/m^2 indicated a cost-saving (dominant) situation or showed ICERs of CHF 8,000 to CHF 44,000 per QALY gained. Adapted ICER results for patients with a BMI <35 kg/m^2 suggested a cost-effective situation with cost of CHF 3,000 to CHF 50,000 per QALY gained. Differences were due to approaches to the modelling of effectiveness (in terms of duration of BMI changes and dependent effects on morbidity, mortality and costs), time horizon, population studied, exact type of intervention studied, and possibly other reasons.

The first part of the budget impact analysis can only be used to describe the potential magnitude of immediate surgical intervention costs at the national level. The second part of analysis over ten years was based on single study for Germany, selected for availability of sufficient numerical detail. It was performed under numerous assumptions and is subject to substantial limitations. The results can't be considered accurate and definitive, but may give a rough idea of the potential magnitude of bariatric surgery-related budget impact.

1.1.5 Assessment of the legal domain

Starting point of the legal analysis are the pertinent provisions of the Federal Swiss Health Insurance Act (HIA). The HIA establishes a system of compulsory social health insurance for all Swiss residents.

Obesity and overweight are not generally considered as illness according to the legal definition in Art. 1a (2) (a) HIA. Only extremer forms qualify as an illness. The details are regulated by an ordinance of the Swiss Government and by the Medical Guidelines of the SMOB. Overweight has to pass a certain threshold to qualify as an illness. To cure that illness, the applicable legal rules provide for both conservative treatments as well as bariatric surgery. The methods and treatments not covered by the mandatory health insurance may be applied nonetheless if they are medically sound and based on an agreement between doctor and patient. Coverage of the costs, however, depends on whether the status of a person may be qualified as an illness.

According to the findings of the report as well as of the SMOB in the Medical Guidelines there is some evidence for a certain superiority of bariatric surgery over conservative treatments. The applicable legal norms, however, require as a general rule first a two year conservative treatment. The costs of earlier bariatric surgery are therefore usually not covered by the public health insurance. Health law thus displays a bias for conservative treatment and against bariatric surgery. This may not be justified as shown by the present report, because for some outcomes bariatric surgery was shown to be more effective e.g. on body weight loss at lower costs than conservative treatment.

1.1.6 Assessment of ethical issues

The evidence for the ethical assessment was reviewed based on:

- issues which became apparent during scoping, and in subsequent discussions during assessment;
- a systematic analysis of possible ethical issues based on three grids;
- a literature search in Pubmed and EBSCO for ethical issues associated with bariatric surgery using keywords associated with both of these terms, followed by screening of resulting titles, abstracts and papers. As recommended by EUnetHTA, this literature search was complemented by a reflective process of literature consultations on ethical issues associated with other, more studied, situations or technologies that pose similar issues.

The purpose of the ethics component of the assessment phase is to yield a series of questions, issues and comments to be integrated during ethical evaluation in the appraisal phase. The interventions involved in weight loss surgery are, for the most part, not new and have been used for other indications. The main ethical issues are thus not associated with the use of new interventions per se but with their increasing use in the treatment of obesity. The main questions raised here are associated with the status of obesity as a disease, the risk of stigmatizing patients or of medicalising social problems through the surgical approach for obesity. Although questions related to personal responsibility for health have been raised in the literature, counter-arguments to integrating these aspects into decisions regarding treatment and coverage have also been brought forward. Other issues are raised by the increased use of bariatric surgery itself, such as how to respect patient autonomy, what requirements can be justified before undergoing surgery, operating on children and adolescents, unequal access, and risks associated with a commercialized innovation.

1.2 Zusammenfassung

1.2.1 Hintergrund

Adipositas (Fettleibigkeit) ist ein ernst zu nehmendes Gesundheitsproblem, das mit einem erhöhten Risiko für Begleiterkrankungen wie Diabetes mellitus Typ 2, Bluthochdruck, obstruktive Schlafapnoe und Erkrankungen des Bewegungsapparates sowie mit einer erhöhten Sterblichkeit einhergeht. In der Schweiz betrugen, die auf Adipositas zurückzuführenden Kosten im Gesundheitswesen im Jahr 2006 geschätzte 5'755 Mio. CHF und im Jahr 2011 7'990 Mio. CHF. 30,8 % der schweizerischen Bevölkerung sind übergewichtig (definiert als ein Body-Mass-Index, BMI, von 25,0-29,9 kg/m^2), und 10,3 % sind adipös (BMI ≥30 kg/m^2). In der Schweiz gehört zu einer angemessenen konservativen Behandlung der Fettleibigkeit eine Kombination aus Massnahmen wie Ernährungsberatung durch einen Ernährungsberater oder Arzt, Verhaltenstherapie einschliesslich Psychotherapie, kalorienreduzierte Diäten, und eine körperliche Therapie einschliesslich Physiotherapie oder Medikamente.

Bleibt eine angemessene konservative Behandlung erfolglos, kann die bariatrische (Gewichtsreduktions-)Chirurgie in Betracht gezogen werden. Eine konservative Therapie gilt dann als erfolglos, wenn in zwei Jahren konservativer Behandlung oder danach kein BMI von <35 kg/m^2 erreicht oder beibehalten werden kann. Die schweizerische Krankenversicherung („obligatorische Krankenpflegeversicherung") deckt bei adipösen Personen mit einem BMI ≥35 kg/m^2 bariatrische Chirurgie ab, und zwar unabhängig von bereits vorliegenden Begleiterkrankungen. Im Jahr 2014 wurden in der Schweiz 4'167 bariatrische Operationen durchgeführt. Dies ist ein deutlicher Anstieg gegenüber dem Jahr 2001, in dem nur 703 solcher Operationen durchgeführt wurden. Etwa 82,5 % aller bariatrischen Operationen sind Magenbypass-Operationen (Roux-en-Y-Magenbypass, RYGB), 15,7 % sind Schlauchgastrektomien (SG), 0,8 % sind Behandlungen mit verstellbarem Magenband, und bei 0,7 % wird die Methode der biliopankreatischen Diversion (BPD) angewendet. Die vertikale Gastroplastik ist ebenfalls abgedeckt, wird aber in der Schweiz nicht mehr praktiziert. Die meisten Operationen werden laparoskopisch durchgeführt und nicht als offene Eingriffe. Die RYGB erfordert eine chirurgische Anpassung des Magens (der Magen wird zu einem verkleinerten Beutel (Magenpouch)) und des Dünndarms, wodurch Kalorienaufnahme und -resorption eingeschränkt werden. Bei einer Magenband-Operation wird ein verstellbarer Ring um den Magen gelegt, der nur den Verzehr kleiner Nahrungsmengen zulässt. Der Ring lässt sich über einen Zugang verstellen. Bei einer Schlauchgastrektomie wird die Grösse des Magens auf einen schlauchähnlichen kleinen Beutel reduziert, während bei der vertikalen Gastroplastik durch ein (nicht verstellbares) Band und Klammern ein Beutel gebildet wird.

Obgleich es international relativ umfangreiche Literatur gibt, liegen für die Schweiz nur wenige Untersuchungen über die Gesamtkosten von Adipositas vor. Es fehlt definitiv ein schweizerisches Modell für die Rentabilität von bariatrischen Operationen oder Magenbypässen bei übergewichtigen (BMI 25-29,9 kg/m^2) und adipösen (BMI ≥30 kg/m^2) Erwachsenen.

1.2.2 Ziel

Das Ziel dieses HTA-Berichts (HTA – Health Technology Assessment) ist die Bestimmung

- der Wirksamkeit und Sicherheit,
- der Kosteneffektivität und Budget-Impact,
- der rechtlichen und ethischen Auswirkungen

der bariatrischen Chirurgie im Vergleich zu konservativen Behandlungsmethoden, sowohl bei der derzeit durch die obligatorische Krankenpflegeversicherung (OKP) versorgten Bevölkerung (d.h. adipöse Personen mit einem BMI ≥35 kg/m^2) als auch bei Patienten, die derzeit nicht durch die OKP versorgt sind (d.h. übergewichtige oder adipöse Personen mit einem BMI von 25 – 35 kg/m^2). Der Schwerpunkt liegt auf dem Magenbypass, auch wenn alle augenblicklich angewendeten chirurgischen Eingriffe berücksichtigt werden.

1.2.3 Klinische Wirksamkeit und Sicherheit

Für diesen HTA-Bericht wurden die klinische Wirksamkeit und Sicherheit der bariatrischen Chirurgie im Vergleich zu konservativen Behandlungsmethoden untersucht. Die Recherche wurde im Juli 2015 durchgeführt, und es wurden Filter für randomisiert kontrollierte Studien (RCT) verwendet. Die Recherche war nicht auf einen Zeitraum begrenzt, wurde aber auf die Sprachen Englisch, Deutsch, Französisch und Italienisch beschränkt. Die Merkmale der Studien und die Ergebnisse der einbezogenen Studien wurden für jede Studie einzeln und nach deskriptiven Methoden zusammengefasst in Tabellen dargestellt. Das Hauptaugenmerk der Analyse lag auf den zusammengefassten Ergebnissen nach 2 bzw. 3 Jahren Nachbeobachtung, aber auch die zusammengefassten Ergebnisse von Nachbeobachtungen über 6, 9 und 12 Monate wurden analysiert und dargestellt. Das Risiko einer Verzerrung und die Qualität der Evidenz wurden für die einbezogenen Studien mit 2 bzw. 3 Jahren Nachbeobachtung bewertet, da der Hauptschwerpunkt auf den langfristigen Resultaten lag. Um weitere Informationen zu gewinnen, wurden ausserdem die Ergebnisse nach einer Nachbeobachtung von 6 Monaten bis zu 1 Jahr zusammengefasst, aber nicht nach dem GRADE System ausgewertet. Das Risiko einer Verzerrung wurde nach dem Cochrane-Handbuch bewertet und die Qualität der Evidenz nach GRADE. Wenn möglich, wurden die Behandlungsergebnisse in einer Metaanalyse quantitativ zusammengefasst, wobei inverse Varianzmodelle unter Annahme von Zufallseffekten verwendet wurden. Effektschätzer (gesamt und einzeln für jede Studie) mit den entsprechenden 95 %-Konfidenzintervallen wurden in Form von Forest Plots dargestellt. Für binäre Endpunkte wurden relative Risiken berechnet. Kontinuierliche Endpunkte wurden als Mittelwertdifferenzen dargestellt. Bei erheblicher Heterogenität wurden nach Möglichkeit methodische und klinische Faktoren, die als Erklärung für die Heterogenität dienen könnten, in Subgruppen- und Sensitivitätsanalysen untersucht. Einige der im Voraus festgelegten Subgruppenanalysen betrafen die chirurgische Methode (d.h. Magenbypass vs. verstellbares

Magenband vs. sonstige chirurgische Eingriffe) und die Patienten mit einem BMI ≥35 kg/m² vs. Patienten mit einem BMI von 25-34,9 kg/m².

Sechzehn RCTs erfüllten die Einschlusskriterien, wobei für 10 RCTs Daten von 2 bzw. 3 Jahren vorlagen. Die meisten RCTs mit Daten über 2 bzw. 3 Jahre schlossen Patienten mit spezifischen Begleiterkrankungen ein (Diabetes mellitus Typ 2: 6 RCTs, Schlafapnoe: 2 RCTs, Mix aus möglichen Begleiterkrankungen: 1 RCT); bei lediglich einer RCT waren Begleiterkrankungen nicht als Einschlusskriterium spezifiziert. Von den 13 bewerteten Endpunkten waren die als kritisch klassifizierten Endpunkte Körpergewicht, Lebensqualität, $HBA1_c$ (Blutzuckerkontrolle), Schlaganfall und Herzinfarkt.

Die beobachteten Ergebnisse in den Metaanalysen für die Ergebnisse nach 2 bzw. 3 Jahren waren sehr konsistent mit statistisch signifikanten besseren Effekten im chirurgischen verglichen mit dem konservativen Arm für die folgenden kritischen Endpunkte: prozentuale Veränderung des Körpergewichts, körperliche Gesundheit (Lebensqualitätskomponente Summenscore), HbA1c und Diabetesremission sowie in Bezug auf die wichtigen Endpunkte: Hypertonie und Dyslipidämie.

Für die Endpunkte psychische Gesundheit (Lebensqualitätskomponente Summenscore), Schlaganfall, Gesamtmortalität, schwerwiegende unerwünschte Ereignisse, einschliesslich Reoperationen, Schlafapnoe und Krebs wurden keine statistisch signifikanten Effekte nachgewiesen.

Es lagen sehr wenige oder keine Daten für den kritischen Endpunkt Schlaganfall und für die wichtigen Endpunkte Sterblichkeit, Herzinfarkt und Revisionsrate vor, so dass für diese Endpunkte keine Schlussfolgerungen gezogen werden können.

In der stratifizierten Darstellung der Ergebnisse für die verschiedenen chirurgischen Methoden oder den BMI (BMI 25-34,9 kg/m² vs. BMI ≥35 kg/m²) deutete die Richtung der beobachteten Effekte – sofern sie vorhanden waren – durchwegs auf einen Vorteil der bariatrischen Chirurgie gegenüber der konservativen Behandlung hin. Die Ergebnisse wurden nach dem chirurgischen Verfahren stratifiziert, aber wegen der geringen Zahl von Studien innerhalb der Strata wurde kein statistischer Vergleich von Effektschätzungen mittels Subgruppenanalysen durchgeführt.

Die Qualität der Evidenz für den kritischen Endpunkt prozentuale Veränderung des Körpergewichts war moderate und die für die Veränderung des HbA1c niedrig. Auf Grund der sehr niedrigen Qualität der Evidenz für die kritischen Endpunkte psychische Gesundheit und Schlaganfall und wegen des Fehlens von Daten zu Herzinfarkten aus RCTs wird die Qualität der Evidenz insgesamt als sehr niedrig bewertet.

Signifikante Unsicherheiten bleiben bezüglich langfristigen (>3 Jahre) Ergebnissen und der übergewichtigen Population bestehen, da zu diesen Gruppen keine oder nur wenige Daten vorlagen.

1.2.4 Kosten-Effektivität und Budget-Impact Analyse

Relevante Datenbanken, namentlich *Medline*, *Embase*, die *Cochrane Library*, die *Centre for Review and Dissemination* (CRD)-Datenbank und die *UK National Health Service's Economic Evaluation Database* (NHS EED), wurden systematisch nach relevanten gesundheitsökonomischen Artikeln durchsucht. Die Qualität der Artikel wurde mit der *Consolidated Health Economic Evaluation*

Reporting Standards (CHEERS)-Checkliste mit 24 Punkten überprüft, wie es von der *ISPOR Health Economic Evaluations Publication Guidelines Task Force* empfohlen wird. Bei internationalen Studien wurde die qualitative Übertragbarkeit auf die Schweiz bewertet, wie in Teil 4.2.4. beschrieben. Bei denjenigen Studien, die als qualitativ übertragbar auf die Schweiz eingestuft wurden, wurden die direkten medizinischen Kostendaten adaptiert, um die Vergleichbarkeit zu verbessern. Diese Adaptation wurde in drei Schritten durchgeführt: erstens erfolgte eine Mengenkorrektur für internationale Unterschiede in der Inanspruchnahme von Gesundheitsleistungen, zweitens eine Korrektur für Preisunterschiede der Gesundheitsleistungen, drittens eine Korrektur für Änderungen im Umfang der Ressourcennutzung und der Preisentwicklung im Laufe der Zeit. Auf dieser Basis wurden die inkrementellen Kosten-Effektivitäts-Verhältnisse (*incremental cost-effectivess ratios*, ICERs) neu berechnet. Der Budget-Impact geschätzt, indem epidemiologischen Daten aus der Schweiz mit schweizerischen sowie adaptierten, internationalen Kostendaten kombiniert wurden. In Anbetracht der geringen zur Verfügung stehenden Daten wurden zwei Ansätze verfolgt. Erstens wurden die potentiellen Operationskosten für die Behandlung aller geeigneten schweizerischen Patienten in einem einzigen Jahr geschätzt. Zweitens wurden die potentiellen Kosten von bariatrischen Operationen, unter Berücksichtigung von resultierenden Folgekosten und Einsparungen, über einen Zeitraum von zehn Jahren, von 2011 bis 2020, aus der Perspektive des schweizerischen Gesundheitssystems berechnet.

Einundzwanzig Kosten-Effektivitätsanalysen und eine systematische Übersichtsarbeit wurden letztendlich mit der CHEERS-Checkliste evaluiert und in diesem Bericht berücksichtigt. Ungefähr die Hälfte der Studien wurde in Europa durchgeführt, die restlichen in den USA oder Australien. Die Mehrheit der Studien schätzte die Kosten-Effektivität als Kosten pro gewonnenem qualitätsadjustiertem Lebensjahr (QALY), wobei die gewählten Zeithorizonte zwischen 5 Jahren und lebenslang lagen. Insgesamt 15 Artikel erfüllten die Kriterien für qualitative Übertragbarkeit und wurden numerisch für die Schweiz adaptiert. Diese Studien wurden aus einer Gesundheitswesen-, Zahler- oder gesellschaftlichen Perspektive durchgeführt. Die Modellierung der Kosten-Effektivität basierte auf kurzfristigen klinischen Studien sowie langfristigen Beobachtungsstudien (z.B. der Schwedischen SOS Studie).

Alle Studien, die Patienten mit einem BMI >35 kg/m^2 einschlossen, befanden die bariatrische Chirurgie für kostensparend oder kosteneffektiv, basiert auf den von den jeweiligen Autoren gewählten Kriterien. Die für die Schweiz adaptierten ICERs betrugen unter 50'000 CHF pro gewonnenem QALY.

Wenn Studien wie die von Borg et al. eine gesellschaftliche Perspektive einnahmen, lagen die ICERs für Magenbypass im Vergleich zur konventionellen Behandlung für Frauen im Alter zwischen 45-54 Jahren mit einem BMI von 40-44 kg/m^2 unter 3’000 CHF pro gewonnen QALY. Bei Männern im Alter zwischen 45-54 Jahren mit einem BMI von 40-44 kg/m^2 erwies sich der Magenbypass als kostensparend und daher dominant. Michaud et al. befanden den Magenbypass in Vergleich zur konventionellen Behandlung bei US-Patienten mit einem BMI > 40 kg/m^2 oder Patienten mit einem BMI > 35 kg/m^2 und einem hohen Komorbiditätsrisiko für kosteneffektiv. Das adaptierte ICER, aus gesellschaftlicher Perspektive, betrug 8’158 CHF pro gewonnenem Lebensjahr.

Für Patienten mit einem BMI <35 kg/m^2 betrugen die adaptierten ICERs aus gesellschaftlicher Perspektive laut der Studie von Borg et al. 10’458 CHF pro gewonnen QALY bei Frauen und

12’365 CHF pro gewonnenem QALY bei Männern. Das ICER für Patienten mit BMI >35 kg/m^2 oder BMI >30 kg/m^2 mit relevanten Komorbiditäten lag in der Studie von Michaud et al. ähnlich, bei 10'502 CHF pro gewonnenem QALY.

Die verschiedenen Verfahren der bariatrischen Chirurgie scheinen sich in ihrem klinischen Nutzen und ihrer Effizienz zu unterscheiden. Jeweils im Vergleich zur konventionellen Behandlung schien der Magenbypass im Vergleich zum Magenband von höherem Nutzen (d.h. gewonnenen QALYs oder Lebensjahren) zu sein, aber auch teurer.

Im Budget-Impact-Teil bestand der erste Ansatz darin, die potentiellen Operationskosten aller qualifizierenden Patienten in der Schweiz zu schätzen. Im Jahr 2012 hatten fast 25'000 Personen einen BMI >40 kg/m^2. Die unmittelbaren Kosten einer chirurgischen Behandlung all dieser Menschen in einem einzigen Jahr betrügen 353 Mio. CHF. Würden zusätzlich auch 10-30% der Menschen in der Schweiz mit einem BMI von 35-40 kg/m^2 und relevanten Komorbiditäten operativ behandelt, würden die unmittelbaren Kosten 496-781 Mio. CHF erreichen. Die geschätzten unmittelbaren Kosten für die 4'153 Operationen, die im Jahr 2014 tatsächlich in der Schweiz durchgeführt wurden, waren 61.1 Mio. CHF.

Im zweiten Teil der Budget-Impact-Analyse basierten die Berechnungen auf einer einzelnen, europäischen Kosten-Effektivitäts-Studie von Ackroyd et al., die sich hinsichtlich der Wirksamkeit der Intervention auf die Rückbildung diabetischer Symptome konzentrierte. Die Auswahl dieser Studie erfolgte primär, weil sie ausreichende numerische Details zur Verfügung stellte. Die Kostenschätzungen über 10 Jahren waren stark von verschiedenen Annahmen abhängig. Wenn beispielsweise 10 % der adipösen Patienten unter Diabetes litten, betrügen die inkrementellen Kosten der bariatrischen Chirurgie im Jahr 2016 55.1 Mio. CHF. Bei einer Diabetesprävalenz von 40% wären bariatrische Interventionen hingegen kostensparend (-1.2 Mio. CHF).

Praktisch alle Studien klassifizierten die bariatrische Chirurgie als kostensparende oder kosteneffektive Intervention, verglichen mit konventioneller Behandlung. Die adaptierten Berechnungen für Patienten mit einem BMI >35 kg/m^2 zeigten eine kostensparende (dominante) Situation oder ICERs zwischen 8'000 CHF und 44'000 CHF pro gewonnenem QALY. Die adaptierten Berechnungen für Patienten mit einem BMI <35 kg/m^2 zeigten eine kosteneffektive Situation mit ICERs zwischen 3'000 CHF und 50'000 pro gewonnen QALY. Unterschiede zwischen den Studienresultaten resultierten aus verschiedenen Vorgehensweisen bei der Modellierung der klinischen Effektivität (z.B. betreffend BMI-Reduktion und davon abhängiger Effekte auf Morbidität, Mortalität und Kosten), verschiedenen Zeithorizonten, verschiedenen untersuchten Populationen, im Detail unterschiedlichen Interventionen und möglicherweise anderem.

Der erste Teil der Budget-Impact-Analyse ist nur geeignet, die potentielle Grössenordnung der unmittelbaren Kosten der bariatrischen Chirurgie auf nationaler Ebene zu schätzen. Der zweite Teil der Analyse mit einer zehnjährigen Perspektive beruhte auf einer einzigen Studie für Deutschland, die wegen der Verfügbarkeit von ausreichenden numerischen Daten ausgewählt wurde. Sie wurde unter zahlreichen Hypothesen durchgeführt und unterliegt erheblichen Einschränkungen. Die Ergebnisse können nicht als genau und endgültig betrachtet werden, geben aber eine grobe Vorstellung von der potentiellen Grössenordnung des Budget-Impact der bariatrischen Chirurgie.

1.2.5 Bestimmung der rechtlichen Aspekte

Ausgangspunkt für die rechtliche Analyse sind die einschlägigen Bestimmungen des schweizerischen Bundesgesetzes über die Krankenversicherung (KVG). Das KVG begründet ein System der obligatorischen Krankenversicherung für alle Bewohner der Schweiz.

Adipositas und Übergewicht werden laut Gesetzesdefinition in Art. 1a (2) (a) des KVG grundsätzlich nicht als Krankheit eingestuft. Lediglich die extremeren Ausprägungen gelten als Krankheit. Die Einzelheiten sind durch eine Verordnung des Bundesrates und durch die Medizinischen Richtlinien der Swiss Society for the Study of Morbid Obesity and Metabolic Disorders (SMOB) geregelt. Übergewicht muss, um als Krankheit angesehen zu werden, eine gewisse Schwelle überschreiten. Zur Heilung dieser Krankheit sehen die geltenden gesetzlichen Regelungen sowohl konservative Behandlungen als auch die bariatrische Chirurgie vor. Die Methoden und Behandlungen, die nicht durch die obligatorische Krankenversicherung abgedeckt sind, können trotzdem durchgeführt werden, wenn sie medizinisch fundiert und zwischen Arzt und Patient vereinbart sind. Die Kostenübernahme hängt allerdings davon ab, ob der Zustand des Patienten als Krankheit anerkannt wird.

Nach den Erkenntnissen des Berichts und nach den medizinischen Richtlinien des SMOB gibt es Hinweise auf eine gewisse Überlegenheit der bariatrischen Chirurgie gegenüber konservativen Behandlungen. Die geltende Rechtslage erfordert allerdings generell zuerst eine zweijährige konservative Behandlung. Die Kosten einer früher durchgeführten bariatrischen Operation werden daher normalerweise nicht von der obligatorischen Krankenversicherung übernommen. Daher führen die geltenden gesundheitsrechtlichen Normen zu einem Trend zur konservativen Behandlung und gegen die bariatrische Chirurgie. Dies mag, wie aus dem vorliegenden Bericht ersichtlich ist, nicht gerechtfertigt sein, weil gezeigt wurde, dass die bariatrische Chirurgie zur Erreichung bestimmter Ergebnisse effizienter war und z.B. bei Gewichtsabnahme niedrigere Kosten verursachte, als eine konservative Behandlung.

1.2.6 Bestimmung der ethischen Aspekte

Das Assessment ethischer Bewertungskriterien erfolgte anhand von:

- Aspekten, die sich während des Scopings sowie in den nachfolgenden Diskussionen während des Assessments zeigten;
- einer systematischen Analyse möglicher ethischer Aspekte, basierend auf drei Rastern, und ;
- einer Literaturrecherche in PubMed, EBSCO und PsycINFO zu ethischen Aspekten in Verbindung mit der bariatrischen Chirurgie mithilfe von Schlüsselbegriffen, die mit diesen beiden Ausdrücken in Zusammenhang stehen, gefolgt von einer Durchsicht der sich daraus ergebenden Titel, Abstracts und wissenschaftlichen Arbeiten. Wie von EUnetHTA empfohlen, wurde diese Sichtung der Literatur ergänzt durch einen reflektiven Vorgang von Literaturkonsultationen zu ethischen Gesichtspunkten in Zusammenhang mit anderen, besser untersuchten Situationen oder Technologien, die ähnliche Gesichtspunkte aufwerfen.

Der Zweck der ethischen Komponente der Bewertungsphase ist es, eine Reihe von Fragen, Themen und Kommentaren zu liefern, die während der ethischen Evaluation in die Einschätzung einbezogen werden sollten. Die Eingriffe, die bei der Chirurgie zur Gewichtsabnahme in Frage kommen, sind überwiegend nicht neu und wurden bereits bei anderen Indikationen angewendet. Die wichtigsten ethischen Gesichtspunkte stehen daher nicht mit der Anwendung neuer Eingriffe per se in Zusammenhang, sondern vielmehr mit der zunehmenden Anwendung chirurgischer Eingriffe in der Adipositas-Behandlung. Die wichtigsten Fragen, die hier gestellt werden, beruhen auf dem Status der Adipositas als Krankheit, dem Risiko einer Stigmatisierung der Patienten und der Verschiebung gesellschaftlicher Probleme durch die chirurgische Behandlung der Adipositas auf die medizinische Ebene. Obgleich in der Literatur Fragen zur persönlichen Verantwortung für die Gesundheit gestellt werden, werden auch Gegenargumente für die Einbeziehung dieses Aspekts in Entscheidungen über Behandlung und Kostenübernahme vorgebracht. Andere Aspekte, wie die Respektierung der Autonomie des Patienten, die Rechtfertigung bestimmter Anforderungen vor der Durchführung einer Operation, Operationen an Kindern und Jugendlichen, der ungleiche Zugang, sowie Risiken im Zusammenhang mit kommerzialisierten Innovationen, ergeben sich aus der zunehmenden Anwendung der bariatrischen Chirurgie an sich.

1.3 Résumé

1.3.1 Contexte

L'obésité est un grave problème de santé, associé à un risque accru de comorbidités, telles que le diabète de type 2, l'hypertension, l'apnée obstructive du sommeil, les troubles musculo-squelettiques et une mortalité accrue. En Suisse, les coûts de soins de santé liés à l'obésité étaient estimés à 5 755 millions CHF en 2006 et à 7 990 millions de CHF en 2011. 30,8 % de la population suisse sont en surcharge pondérale (caractérisée par un indice de masse corporelle (IMC) compris entre 25,0 et 29,9 kg/m^2) et 10,3 % sont obèses (IMC ≥ 30 kg/m^2). En Suisse, un traitement conservateur adéquat de l'obésité comprend toute combinaison de conseils nutritionnels dispensés par un nutritionniste ou un médecin, de thérapie comportementale, y compris la psychothérapie, de régimes alimentaires visant à réduire l'apport calorique, de thérapie physique, y compris la physiothérapie, ou de médication.

En cas d'échec d'un traitement conservateur adéquat, la chirurgie bariatrique peut être envisagée. On considère que le traitement conservateur a échoué s'il n'est pas possible d'atteindre et de maintenir un IMC < 35 kg/m^2 après deux années de traitement conservateur ou par la suite. L'assurance obligatoire des soins (AOS) suisse prend en charge chirurgie bariatrique chez des individus obèses présentant un IMC ≥ 35 kg/m^2, indépendamment des comorbidités existantes. 4167 interventions bariatriques ont été effectuées en Suisse en 2014, soit une hausse marquée par rapport à 2001 où seulement 703 opérations ont été effectuées. Les bypass gastriques (bypass gastrique Roux-en Y, RYGB) représentent environ 82,5 % de toutes les interventions bariatriques, contre 15,7 % pour les gastrectomies longitudinales (GL), 0,8 % pour les anneaux gastriques ajustables (AGA) et 0,7 % pour la dérivation biliopancréatique (DBP). La gastroplastie verticale calibrée (GVC) est également prise en charge, mais n'est de fait plus pratiquée en Suisse. La plupart des interventions sont pratiquées par laparoscopie au lieu d'une intervention à ciel ouvert. Le RYGB consiste en un réarrangement chirurgical de l'estomac (l'estomac est rétréci) et de l'intestin grêle de manière à limiter l'apport et l'absorption de calories. Dans le cas de l'AGA, un anneau ajustable placé autour de l'estomac ne permet d'ingérer que de petites portions d'aliments. L'anneau peut être ajusté par le biais d'un port d'accès. Dans le cas de la GL, la taille de l'estomac est réduite à une petite poche en forme de tube ; dans le cas de la GVC, la poche est créée à l'aide d'un anneau (non ajustable) et d'agrafes.

Bien que la littérature internationale soit relativement abondante, seules quelques études ont examiné le coût global de l'obésité en Suisse. Il n'existe clairement aucune modélisation suisse du rapport coût-efficacité de la chirurgie bariatrique ou du bypass gastrique chez les adultes en surcharge pondérale (IMC = 25-29,9 kg/m^2) et obèses (IMC ≥ 30 kg/m^2).

1.3.2 But

Le but du présent rapport HTA est d'évaluer :

- l'efficacité et la sécurité cliniques,
- les analyses du rapport coût-efficacité et de l'impact budgétaire,
- les implications juridiques et éthiques

de la chirurgie bariatrique comparée au traitement conservateur, tant pour la population actuellement couverte par l'AOS (c.-à-d. les personnes obèses présentant un IMC ≥ 35 kg/m^2) que pour les patients qui ne sont actuellement pas pris en charge par l'AOS (c.-à-d. les personnes en surcharge pondérale ou obèses présentant un IMC de 25-35 kg/m^2). Bien que toutes les interventions chirurgicales actuellement pratiquées aient été examinées, l'accent porte essentiellement sur le bypass gastrique.

1.3.3 Efficacité et sécurité cliniques

Dans le cadre du présent rapport HTA, l'efficacité et la sécurité cliniques de la chirurgie bariatrique comparée à un traitement conservateur ont été évaluées. L'étude a été menée en juillet 2015 et a appliqué des filtres pour trouver les essais contrôlés randomisés (ECR). L'étude n'a pas été limitée dans le temps, mais a été limitée à l'anglais, l'allemand, le français et l'italien. Les caractéristiques et les conclusions des études retenues ont été présentées par étude sous forme de tableaux, accompagnés d'un résumé descriptif. L'analyse a essentiellement porté sur les résultats combinés après un suivi de 2 ou 3 ans, mais les résultats combinés des suivis après 6, 9 et 12 mois ont été également analysés et présentés. Le risque de biais et la qualité des données scientifiques ont été évalués pour les études retenues avec un suivi de 2 ou 3 ans, dans la mesure où l'accent portait essentiellement sur les résultats à long terme. À titre d'information complémentaire, les résultats des suivis après 6 mois à 1 an ont également été regroupés, mais n'ont pas été évalués selon les critères GRADE. Le risque de biais a été évalué conformément au manuel Cochrane et la qualité des données scientifiques a été évaluée selon les critères GRADE. Dans la mesure du possible, les résultats ont été résumés quantitativement dans le cadre d'une méta-analyse grâce une pondération par l'inverse de la variance avec effets aléatoires. Les estimations des effets (résumés et simples pour chaque essai) ont été présentées sous forme de graphiques en forêt, avec un intervalle de confiance correspondant de 95 %. Les risques relatifs ont été calculés pour les résultats binaires. Les résultats continus ont été présentés sous forme de différences moyennes. En cas d'hétérogénéité considérable, les facteurs méthodologiques et cliniques susceptibles d'expliquer cette hétérogénéité ont été examinés dans le cadre d'analyses de sous-groupes et de sensibilité, dans la mesure du possible. Certaines des analyses de sous-groupes pré-spécifiées portaient sur la méthode chirurgicale (c.-à-d. bypass gastrique comparé aux anneaux gastriques ajustables et aux autres interventions chirurgicales) ainsi que sur les patients présentant un IMC ≥ 35 kg/m^2 par rapport aux patients ayant un IMC de 25 à 34,9 kg/m^2.

Seize ECR ont satisfait aux critères d'inclusion, dont 10 ECR portant sur des données de 2 ou 3 ans. La plupart des ECR relatives à des données de 2 ou 3 ans comprenaient des patients présentant des comorbidités spécifiques (diabète de type 2 : 6 ECR, apnée du sommeil : 2 ECR, mélange de comorbidités potentielles : 1 ECR) ; seule une ECR ne spécifiait pas de comorbidité en tant que critère d'inclusion. Sur les 13 résultats évalués, le poids corporel, la qualité de vie, l'HBA1c (contrôle du diabète), les AVC et l'infarctus du myocarde ont été considérés comme critiques.

Les effets observés dans les méta-analyses pour les résultats de 2 à 3 ans indiquaient de manière très consistante un bénéfice statistiquement important de l'approche chirurgicale par rapport au traitement conservateur pour les résultats critiques suivants : pourcentage de variation du poids corporel, santé physique (score composite de qualité de vie), HbA1c et rémission du diabète, et les résultats importants suivants : hypertension et dyslipidémie.

Aucun effet statistiquement important n'a été identifié pour la santé mentale (score composite de qualité de vie), les AVC, la mortalité toutes causes confondues, les événements indésirables graves (EIG), y compris les réopérations, l'apnée du sommeil, et le cancer.

Les données étaient extrêmement rares (1 événement maximum) ou inexistantes pour le résultat critique AVC et les résultats importants suivants : mortalité, infarctus du myocarde et taux de révision ; il n'a donc pas été possible de tirer des conclusions pour ces résultats.

Dans la présentation stratifiée des résultats pour les différentes techniques chirurgicales ou l'IMC (IMC 25-34,9 kg/m^2 contre IMC ≥ 35 kg/m^2), la direction des effets observés, quand ils étaient présents, a constamment montré l'avantage de la chirurgie bariatrique par rapport au traitement conservateur. Les résultats ont été stratifiés par technique chirurgicale, mais les estimations des effets n'ont pas été comparées statistiquement via des analyses de sous-groupes en raison du petit nombre d'études par strate.

La qualité des données scientifiques était modérée pour le résultat critique pourcentage de variation du poids corporel et faible pour l'évolution de l'HbA1c. La qualité générale des données est considérée comme très faible en raison de la très faible qualité des données pour les résultats critiques santé mentale et AVC et de l'absence de données provenant d'ECR sur l'infarctus du myocarde.

Des incertitudes importantes subsistent concernant les résultats à long terme (> 3 ans) et la population en surcharge pondérale, pour lesquels il n'y a guère ou pas de données disponibles.

1.3.4 Analyses du rapport coût-efficacité et de l'impact budgétaire

Les bases de données pertinentes, y compris Medline, Embase, la Cochrane Library, la base de données du Centre for Review and Dissemination (CRD) et la base de données d'évaluation économique du National Health Service (NHS EED) britannique, ont été systématiquement examinées en vue d'identifier des articles pertinents consacrés à l'économie de la santé. La qualité des rapports a été évaluée par rapport aux 24 points de la liste de contrôle CHEERS (Consolidated Health Economic Evaluation Reporting Standards), recommandée par le groupe de travail chargé des directives de publication des évaluations d'économie de la santé de l'ISPOR. Pour les études internationales, la transférabilité qualitative vers la Suisse a été évaluée de la façon décrite à la section 4.2.4. Dans le cas des études classées comme qualitativement transférables vers la Suisse, les données des coûts médicaux directs ont été adaptées pour améliorer la comparabilité, en trois étapes distinctes : correction pour tenir compte de différents niveaux d'utilisation des ressources, de différents prix des services de santé et de l'évolution du niveau d'utilisation des ressources et des prix dans le temps. Des rapports coût-efficacité incrémentiels (ICER) adaptés ont ensuite été calculés. L'impact budgétaire a été abordé en combinant des données épidémiologiques suisses avec des données de coûts suisses et internationales adaptées. Compte tenu des données limitées disponibles, deux approches ont été adoptées. Les coûts chirurgicaux potentiels pour traiter tous les patients suisses éligibles en une seule année ont tout d'abord été estimés. Les coûts potentiels de la chirurgie bariatrique sur une période de 10 ans, de 2011 à 2020, ont ensuite été évalués du point de vue du système de santé suisse.

21 analyses du rapport coût-efficacité et une revue systématique ont finalement été incluses dans ce rapport et ont été évaluées à l'aide de la liste de contrôle CHEERS. La moitié des études environ ont été conduites en Europe, tandis que les autres ont été menées aux États-Unis ou en Australie ; la vaste majorité ont évalué le rapport coût-efficacité en tant que coût par année de vie pondérée par la qualité (QALY) gagnée, pour des horizons temporels allant de 5 ans à la durée de vie complète. 15 articles au total ont satisfait aux critères de transférabilité qualitative et ont été soumis à une adaptation numérique des résultats en termes d'ICER pour la Suisse. Les études adoptées portaient sur le point de vue sociétal, du payeur ou des soins de santé. Leur modélisation de l'efficacité clinique reposait sur des essais cliniques à court terme et des données d'observation à plus long terme (telles que l'étude SOS).

Toutes les études de patients présentant des valeurs IMC > 35 kg/m^2 ont indiqué que la chirurgie bariatrique était économique ou rentable, selon les critères définis par les auteurs. Les ICER adaptés pour la Suisse étaient en deçà de 50 000 CHF par QALY gagnée.

Lorsque des études comme Borg et al. ont adopté un point de vue sociétal, les ICER adaptés pour les femmes âgées de 45 à 54 ans présentant un IMC de 40-44 kg/m^2 étaient inférieurs à 3 000 CHF quand le bypass gastrique était comparé à un traitement conventionnel. Pour les hommes âgés de 45 à 54 ans avec un IMC de 40-44 kg/m^2, le bypass gastrique était économique et par-là même dominant. Michaud et al. indiquent également un rapport coût-efficacité du point de vue sociétal chez des patients américains présentant un IMC > 40 kg/m^2 ou un IMC > 35 kg/m^2 avec un risque élevé de comorbidités. L'ICER adapté pour le bypass gastrique par rapport au traitement conventionnel était de 8 158 CHF par année de vie gagnée.

Pour les patients avec un IMC < 35 kg/m^2, l'ICER adopté d'un point de vue sociétal était de 10 458 CHF par QALY gagnée pour les femmes, tandis que l'ICER adopté pour les hommes était de 12 365 CHF par QALY gagnée dans l'étude Borg et al. Pour les patients avec un IMC > 35 kg/m^2 ou IMC > 30 kg/m^2 avec comorbidités qualifiantes, l'ICER adapté pour le bypass gastrique par rapport au traitement conventionnel était de 10 502 CHF par année de vie gagnée, dans l'analyse de Michauld et al.

Des différences spécifiques aux procédures semblent exister en termes de bénéfices et d'efficacité. Quand le bypass gastrique et l'anneau gastrique ont été comparés au traitement conventionnel, le bypass gastrique s'est avéré supérieur à l'anneau gastrique en termes de bénéfices (nombre de QALY gagnées et années de vie sauvées), mais également plus onéreux.

Dans la partie de l'impact budgétaire, la première approche était d'estimer les coûts potentiels de la chirurgie de traiter tous les patients Suisses éligibles. La Suisse comptait près de 25 000 personnes présentant un IMC > 40 kg/m² en 2012. Les coûts du traitement chirurgical de toutes ces personnes en une seule année s'élèveraient à 353 millions de CHF. En outre, si 10 à 30 % des personnes vivant en Suisse et présentant un IMC de 35-40 kg/m² subissaient un traitement chirurgical en raison de la présence de comorbidités pertinentes, le total des coûts chirurgicaux atteindrait de 496 à 781 millions de CHF. En 2014, le total des coûts chirurgicaux pour les 4 153 interventions chirurgicales effectivement réalisées en Suisse était estimé à 61,1 millions de CHF.

Dans la seconde approche, l'estimation de l'impact budgétaire reposait sur une seule étude européenne du rapport coût-efficacité, menée par Ackroyd et al. Cette étude repose principalement l'efficacité de l'intervention sur la rémission du diabète. Cette décision a été faite par la disponibilité de détail numérique suffisante. Les résultats sur 10 ans dépendent fortement des hypothèses. Par exemple, si la prévalence du diabète chez les patients obèses était de 10 %, les coûts incrémentiels de la chirurgie bariatrique en 2016 pourraient atteindre 55,1 millions de CHF. À l'opposé, une prévalence de 40 % du diabète chez les patients obèses pourrait rendre la chirurgie bariatrique plus avantageuse (-1,2 million de CHF).

La quasi-totalité des études ont conclu que la chirurgie bariatrique était économique ou rentable par rapport au traitement conventionnel. Les résultats ICER adaptés pour les patients présentant un IMC > 35 kg/m² ont indiqué des économies (dominantes) ou affiché des ICER allant de 8 000 à 44 000 CHF par QALY gagnée. Les résultats ICER adaptés pour les patients présentant un IMC < 35 kg/m² suggéraient une situation rentable, avec un coût de 3 000 à 50 000 CHF par QALY gagnée. Les différences étaient imputables aux approches de modélisation de l'efficacité (en termes de durée des changements d'IMC et des effets dépendants sur la morbidité, la mortalité et les coûts), de l'horizon temporel, de la population étudiée, du type exact d'intervention étudiée et d'autres raisons potentielles.

La première partie de l'analyse d'impact budgétaire ne peut être utilisée que pour décrire la magnitude potentielle des coûts d'une intervention chirurgicale immédiate au niveau national. La seconde partie de l'analyse sur 10 ans reposait sur une étude unique pour l'Allemagne, retenue en raison de la disponibilité de détails numériques suffisants. Elle a été effectuée à partir de nombreuses hypothèses et est sujette à des limites importantes. Les résultats ne peuvent être considérés comme précis et définitifs, mais peuvent donner une idée générale de la magnitude potentielle de l'impact budgétaire lié à la chirurgie bariatrique.

1.3.5 Évaluation du domaine juridique

Les dispositions pertinentes de la Loi fédérale sur l'assurance-maladie (LAMal) constituent le point de départ de cette analyse juridique. La LAMal prévoit un système social d'assurance maladie obligatoire pour toutes les personnes résidant en Suisse.

L'obésité et la surcharge pondérale ne sont généralement pas considérées comme des maladies selon la définition juridique de l'Art. 1a (2) (a) de la LAMal. Seules les formes les plus extrêmes sont considérées comme des maladies. Les détails sont réglementés par une ordonnance du gouvernement suisse ainsi que par les directives médicales de la SMOB (Société *suisse* pour l'étude de l'obésité morbide). La surcharge pondérale doit franchir un certain seuil pour être considérée comme une maladie. Pour soigner cette maladie, les règles juridiques applicables prévoient à la fois des traitements conservateurs et la chirurgie bariatrique. Les méthodes et traitements qui ne sont pas couverts par l'assurance maladie obligatoire peuvent néanmoins être appliqués s'ils sont médicalement sérieux et basés sur un accord entre le médecin et le patient. La prise en charge des coûts dépend cependant de l'état du patient, selon qu'il est ou non considéré comme une maladie.

Selon les conclusions du rapport et des directives médicales du SMOB, il semblerait que la chirurgie bariatrique présente une certaine supériorité par rapport aux traitements conservateurs. Les normes juridiques applicables imposent cependant, en règle générale, un traitement conservateur préalable pendant deux ans. Les coûts d'une chirurgie bariatrique précoce ne sont donc pas généralement couverts par l'assurance maladie publique. La loi sur la santé affiche ainsi une préférence pour le traitement conservateur, au détriment de la chirurgie bariatrique. Cela pourrait ne pas être justifié, comme le montre le présent rapport, dans la mesure où la chirurgie bariatrique s'avère plus efficace pour certains résultats et permet par exemple d'obtenir une perte de poids à un coût moindre par rapport à un traitement conservateur.

1.3.6 Considération des questions éthiques

Les données de cette évaluation éthique ont été examinées à partir :

- des problèmes qui sont devenus apparents durant le cadrage et les discussions ultérieures pendant l'évaluation ;
- d'une analyse systématique des questions éthiques potentielles basées sur trois grilles ;
- d'une recherche de la littérature sous Pubmed et EBSCO relative aux questions éthiques associées à la chirurgie bariatrique, à l'aide de mots-clés associés à ces deux termes, suivi d'un filtrage des titres, résumés et articles identifiés. Conformément aux recommandations de l'EUnetHTA, cette recherche de littérature est allée de pair avec un processus réflexif de consultation de la littérature portant sur des questions éthiques associées à d'autres situations ou technologies, plus étudiées, qui soulèvent des questions similaires.

L'objectif de cet élément éthique de la phase d'évaluation est de soulever une série de questions, problématiques et commentaires qui seront incorporés dans l'évaluation éthique lors de la phase d'appréciation. Les interventions concernées par la chirurgie bariatrique n'ont, pour la plupart, rien de nouveau et ont été utilisées pour d'autres maladies. Les principales questions éthiques ne sont donc pas associées à l'utilisation de nouvelles interventions en soi, mais à leur utilisation croissante dans le traitement de l'obésité. Les principales questions soulevées ici sont associées au statut de l'obésité en tant que maladie et au risque de stigmatiser les patients ou de médicaliser les problèmes sociaux par le biais d'une approche chirurgicale de l'obésité. Bien que des questions liées à la responsabilité personnelle en matière de santé aient été évoquées dans la littérature, des contre-arguments à l'intégration de cet aspect dans les décisions relatives au traitement et à la prise en charge ont également été avancés. L'utilisation accrue de la chirurgie bariatrique elle-même soulève d'autres questions, notamment par rapport au respect de l'autonomie des patients, aux conditions à remplir avant de subir une intervention chirurgicale, aux interventions sur les enfants et les adolescents, à l'inégalité de l'accès et aux risques associés à une innovation lorsque celle-ci est commercialisée.

2 Introduction

2.1 Background

Obesity is a serious health problem associated with an increased risk for comorbidities, such as type 2 diabetes, hypertension, obstructive sleep apnoea and musculoskeletal disorders, and increased mortality.[1] In Switzerland, the estimated health care costs due to obesity were 5,755 Mio. CHF in 2006 and 7,990 Mio. CHF in 2011.[2] Of the Swiss population 30.8% are overweight (defined as body mass index, BMI 25.0-29.9 kg/m^2) and 10.3% are obese (BMI ≥30 kg/m^2).[3] In Switzerland an adequate conservative treatment of obesity includes any of the following or their combination[3]:

1. nutritional counselling by a nutritionist or physician with expertise in nutrition
2. behavioural therapy including psychotherapy
3. diets to reduce caloric intake
4. physical therapy including physiotherapy
5. medication

If adequate conservative treatment fails, bariatric (weight-loss) surgery can be considered. Conservative treatment is regarded as failed if in two years of conservative treatment or thereafter a BMI of <35 kg/m^2 cannot be reached and maintained. Swiss statutory health insurance ("obligatorische Krankenpflegeversicherung, OKP") covers certain types of bariatric surgery in obese individuals with a BMI ≥35 kg/m^2 independent of existing comorbidities.[4] The following surgical techniques are performed in Switzerland: gastric bypass (Roux-en Y gastric bypass, RYGB), adjustable gastric banding (AGB), sleeve gastrectomy (SG), biliopancreatic diversion (BPD) with or without duodenal switch. Vertical banded gastroplasty (VBG) is covered but de facto no longer performed in Switzerland.[3] [5] The technical details of how the surgical interventions are performed can vary. Multiple surgical interventions can be unplanned, due to reoperations following complications or due to revisions because of ineffective initial interventions or pre-planned as staged approaches, for example in cases where the patients need to lose some more weight before the definitive surgical intervention can be safely performed. The fitting of a gastric balloon, which is not a bariatric surgery technique in a narrower sense, is not covered by the Swiss statutory health insurance.

The RYGB involves surgical rearrangements of the stomach (the stomach becomes a smaller pouch) and small intestine limiting caloric intake and absorption. In the case of adjustable gastric banding, an adjustable ring placed around the stomach allows only the ingestion of small food portions. The ring can be adjusted via an access port. In the case of sleeve gastrectomy the size of the stomach is reduced to a sleeve-like small pouch, while the pouch in the case of VBG is created using non-adjustable band and staples.

In Switzerland 4,167 surgeries were performed in 2014, many more than in 2001 when only 703 were done.

Gastric bypass is the most frequently used surgical technique. About 82.5% of all bariatric surgeries are gastric bypasses, 15.7% are sleeve gastrectomies, 0.8% are gastric bandings and

0.7% are biliopancreatic diversions.[5] Most of the surgeries are performed laparoscopically instead of open surgery.

An earlier health technology assessment (HTA) report for the „Bundesamt für Gesundheit, BAG" entitled „Effectiveness of preventive and therapeutic interventions in overweight and obesity" evaluated only one systematic review[6], which compared bariatric surgery and conservative treatment in terms of QALY gained.[7]

While there is a relatively broad international literature, there are only few studies for Switzerland which examine the overall cost of obesity. There is clearly no Swiss modelling of the cost-effectiveness of bariatric surgery or gastric bypass for overweight (BMI 25-29 kg/m^2) and obese (BMI ≥30 kg/m^2) adults.

2.2 Scoping process

The research question was chosen among the submitted research questions by the "Trägerverein" based on the criteria pre-defined in the SMB methods paper.[8,9]

Suggestions for relevant PICO-questions were developed with the support of clinical experts. The PICO-questions define the **p**opulation to be assessed, the **i**ntervention, the **c**omparators and the relevant **o**utcomes. In the course of the scoping process the clinical experts had the opportunity to comment on the PICO-questions, current and relevant bariatric surgical techniques, the adequate duration of follow-up and the use of medical resources.

Critical and important outcomes were defined together with the clinical experts according to GRADE (Grading of Recommendations Assessment, Development and Evaluation), a tool for the rating of the quality of evidence and the strength of recommendations.

2.3 Aim

The aim of this HTA report is to assess the effectiveness and safety as well as the economic, legal and ethical implications of bariatric surgery compared to conservative treatment; both in the population currently covered by the OKP (i.e. obese individuals with a BMI ≥35 kg/m^2) and in patients currently not covered by the OKP (i.e. overweight or obese individuals with a BMI of 25 - 35 kg/m^2). While all the surgical interventions currently used were being considered, the main focus was gastric bypass.

3 Clinical effectiveness and safety

3.1 Methods

3.1.1 Literature search

The literature search comprised the following databases:

- Central („Cochrane central register of controlled trials"),
- EMBASE via OvidSP,
- Medline via OvidSP and
- in addition via PubMed (top up search for publications not yet indexed in OvidSP Medline)

The strategy was combined with a search filter for randomized controlled trials (RCTs): i.e. the best optimized RCT filter with regard to sensitivity and specificity, by Wong et al. 2006[10] in EMBASE and the "Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity- and precision-maximizing version (2008 revision)" in Medline, combined with the search terms "random" and "randomised". See Appendix A for details on the search strategies used.

Screening of the literature

Two reviewers independently screened the titles and abstracts of the records for potentially eligible studies. Discrepant screening results were discussed and resolved by consensus. Full text articles of potentially eligible studies were subsequently screened by two reviewers independently to identify eligible RCTs. Discrepant screening results were discussed and resolved by consensus or by third party arbitration.

3.1.2 Inclusion criteria

Table 1 Inclusion criteria clinical effectiveness and safety

Population	Overweight and obese adults as defined in section 3.1.2.1
Intervention	Bariatric surgery as defined in section 3.1.2.2
Comparator	Conservative treatments as defined in section 3.1.2.3
Outcomes	Weight loss, quality of life, mortality, reoperations as well as disease specific outcomes as defined in section 3.1.2.4
Study design	Randomised and quasi-randomised controlled trials (see section 3.1.2.5)
Languages	English, German, French, Italian (see section 3.1.2.6)

3.1.2.1 **Population**

Overweight (BMI 25.0-29.9 kg/m^2) and obese (BMI ≥30 kg/m^2) adults.

3.1.2.2 **Intervention**

Any bariatric surgery technique, **except for**

- jejunoileal bypass/plication,
- gastric imbrication/plication (because it is only done in an experimental context) and
- horizontal gastroplasty,
- vertical banded gastroplasty or
- vertical gastroplasty (de facto no longer performed in Switzerland).[5]

A particular focus of the report is the assessment of the effects of gastric bypass. Studies investigating Roux-en-Y gastric bypass and banded gastric bypass were considered to be relevant for the gastric bypass group. Only laparoscopic gastric bypass was considered in this group either with or without robotic assistance.

Studies investigating the following forms of gastric bypass were **excluded from the gastric bypass group** but were considered in the group “other surgical interventions”**:**

- Open gastric bypass
- Distal gastric bypass
- Mini bypass
- Omega loop bypass
- Very long limb bypass

3.1.2.3 **Comparator**

Any conservative, non-surgical intervention including reduced calorie intake, increased physical activity, behavioural therapy, medication or any combination of these treatments was eligible.

3.1.2.4 **Outcomes**

Critical and important outcomes were defined in collaboration with clinical experts of the FMH.

Critical outcomes:

1. Body weight
2. Quality of life
3. $HBA1_c$ (diabetes control)
4. Stroke
5. Myocardial infarction

Important outcomes:

6. All-cause mortality
7. Serious adverse events/ reoperations*
8. Diabetes remission
9. Hypertension
10. Dyslipidaemia
11. Sleep apnoea
12. Cancer
13. Revision rates**

*Reoperation was defined as operations performed due to problems/complications of the initial intervention like treatment of hernias developed post-operatively, insufficiency of the anastomosis etc.

**Revisions, planned or unplanned were performed when the initial intervention was for example ineffective or not complete: i.e. the procedure was changed into another one (Umwandlungsoperation).

Post-hoc data on adverse events and the apnoea-hypopnea index were extracted and analysed.

3.1.2.5 **Study design**

Both randomised controlled trials (RCT) and quasi-RCTs (inadequate method to randomise patients, e.g. alteration) fulfilling the inclusion criteria.

3.1.2.6 **Languages**

Trials published in English, French, German and Italian were included.

3.1.3 Data extraction

Data on study characteristics and outcomes were extracted into a standardized form by one reviewer and checked by another. Discrepancies were resolved by discussion or third party arbitration.

The data extracted included patient recruitment time period, setting and country, previous weight loss attempts of the patients, number of patients who refused the intervention or number of patients who withdrew, adverse events, adherence rate at two years or at last follow-up and description of the study interventions, scheduled follow-up time points and maximum follow-up. The extraction included the following baseline characteristics: age, sex distribution, mean BMI, mean body weight, mean HbA1c, mean duration of type 2 diabetes, and proportion of patients with BMI <35 kg/m^2, type 2 diabetes, dyslipidaemia, and other comorbidities.

Post-hoc the number of patients refusing conservative or surgical treatment or, if this was not reported, the number of patients withdrawing from the intervention they were randomised to was extracted. In addition the rate of patients returning to the last follow-up (adherence rate) was extracted.

For the meta-analysis body weight measured as a percentage change from baseline was used. Quality of life, $HbA1_c$ and apnoea-hypopnea index (AHI; expressed as events/hour) were expressed as absolute changes from baseline. Serious adverse events and reoperations were treated as combined endpoints. The number of patients with serious adverse events/reoperation, diabetes remission, mortality, cancer, hypertension, dyslipidaemia, metabolic syndrome and sleep apnoea at follow-up was extracted. The definition of diabetes remission used was as described by the study authors. The same holds for hypertension, dyslipidaemia, metabolic syndrome and sleep apnoea if reported.

Relevant time points were 6 months follow-up or longer. The analysis combined the results of studies with 6, 9 or 12 months follow-up and those of studies with 2 or 3 years follow-up. The quality of evidence was assessed for data at 2 or 3 years follow-up. Studies with a longer duration of follow-up would have been eligible as well but no data were available.

3.1.4 Risk of bias assessment and quality of evidence assessment

One reviewer assessed the internal validity (risk of bias) of each trial and per outcome. This was checked by a second reviewer. Discrepancies were resolved by discussion or third party arbitration.

To assess the risk of bias of individual trials the following criteria were used[11-27]:

- adequate random sequence generation
- adequate concealment of treatment allocation
- adequate blinding of patients, health carers, and outcome assessors
- complete outcome data
- reporting bias

Blinding of outcome assessors and complete outcome data were judged at the outcome level.

The quality of the evidence was judged according to GRADE (Grading of Recommendations Assessment, Development and Evaluation) for the critical and important outcomes after 2 years (i.e. at the outcome level by considering all available trials for the respective outcome). The following criteria were considered to judge the quality of the evidence[16 25]:

Criteria for rating down the quality of evidence:

- risk of bias (internal validity)
- inconsistency
- indirectness
- imprecision
- publication bias

Criteria for rating up the quality of evidence:

- large magnitude of effect
- dose-response gradient
- all plausible confounders or biases increase the confidence in the estimated effect

Using the GRADE software (GRADEprofiler Version 3.6.1) results were presented in a summary of findings table. The definition of the four levels of evidence according to GRADE is described in Table 2.[14]

Table 2 Quality levels of evidence

Quality level	Definition
High	We are very confident that the true effect lies close to that of the estimate of the effect
Moderate	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low	Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low	We have very little confidence in the effect estimate: The true effect is likely to be substantially different form the estimate of effect

3.1.5 Evidence synthesis

Study characteristics and results of the eligible trials were presented per study in tables and summarized descriptively.

The main focus of the analysis was the combined results after 2 or 3 years follow-up; the combined results of the 6, 9 and 12 months follow-up were analysed (see Appendix 9).

Where possible, outcome results were summarized quantitatively in a meta-analysis by using inverse variance models assuming random effects.[28] The analyses were performed using Review Manager (Version 5.3.5).

In the case of 3-arm studies, the surgical treatment arms were combined (e.g. in case of two different surgical techniques they were combined).

In case that the relevant effect estimates were missing, study authors were asked via email to report them or they were calculated based on other relevant information in the publication. Missing standard deviations were approximated by calculating the median of the available standard deviations.[11]

Effect estimates (summary and single for each trial) with corresponding 95% confidence interval were presented as forest plots.

Relative risks were calculated for binary outcomes.

Continuous outcomes were presented as mean differences. If a continuous outcome would have been measured on different scales, mean differences of the individual trial results would have been standardized using the following formula:

Standard mean difference (SMD) = $(\text{mean}_{\text{surgery}} - \text{mean}_{\text{conservative}})/\text{SD}_{\text{pooled}}$

An effect size of 0.2 standard deviations (SD) corresponds to a small effect; effect sizes of 0.5 and 0.8 SDs correspond to medium and big effects, respectively.[29 30]

The minimal clinically important difference (MCID) for the outcome of quality of life was assessed. An increase/decrease of 3-5 points on SF-36 scales (scores range from 1 to 100) was considered as clinically relevant.[31]

The presence of heterogeneity in the pooled effect was estimated using I^2.The interpretation of I^2 followed the guidance of the Cochrane Handbook for Interventions, where an I^2 of 0% to 40% indicates: there may be no important heterogeneity; 30% to 60%: may represent moderate heterogeneity; 50% to 90%: may represent substantial heterogeneity; 75% to 100%: considerable heterogeneity. The importance of the observed I^2 value depends on (i) magnitude and direction of effects and (ii) strength of evidence for heterogeneity (e.g. p value from the chi-squared test, or a confidence interval for I^2).[11 32]

In case of considerable heterogeneity methodological and clinical factors that might explain the heterogeneity were explored in subgroup and sensitivity analyses.

3.1.6 Subgroup analysis

In order to answer questions regarding possible variations of the effects depending on the patients treated, the type of intervention, the type of study design, and to investigate possible heterogeneous results; subgroup analyses were planned for the following pre-specified subsets:

1. Gastric bypass vs. adjustable gastric banding vs. other surgical interventions

2. Patients with a BMI ≥35 kg/m^2 vs. patients with a BMI 25-34.9 kg/m^2 (studies with a mixed population were classified depending on the BMI of at least 80% of the population)
3. Perioperative (≤30 days after surgery) vs. longer term (>30 days after surgery) serious adverse events/reoperations
4. Adequate vs. inadequate or unclear random sequence generation (planned but not done)
5. Adequate vs. no adequate or unclear allocation concealment (planned but not done)
6. Adequate vs. inadequate or unclear blinding of patients, carers, and outcome assessors (planned but not done)
7. Low vs. high or unclear risk due to incomplete outcome data (planned but not done)

3.1.7 Sensitivity analysis

In a sensitivity analysis the combined endpoint of serious adverse events/reoperations was treated as separate endpoints.

In another sensitivity analysis the effects on HbA1c by excluding studies not done in obese/overweight populations with type 2 diabetes was assessed.

If enough data had been available, a sensitivity analysis would have been performed using a broader definition for gastric bypass by including open gastric bypass, distal gastric bypass, mini bypass, omega loop bypass, and very long limb bypass.

Further exploratory sensitivity analyses were done; if applicable they are being described in the result section for the outcome it concerned.

3.2 Results

3.2.1 Results of the literature search

Results of the literature search are presented in Figure 1. The literature search identified 4,127 records. In the title and abstracts screening 74 records were identified as potentially relevant. After full text screening 21 publications were included corresponding to 10 RCTs with a follow-up of at least 2 years (Table 3) and six RCTs with a follow-up of less than 2 years (but longer than 6 months) (see Appendices: Appendix 6 to Appendix 9). The longest follow-up was 3 years (three studies). One study reported on outcomes 7 years after study completion, beyond the planned study period. Therefore these results were not considered for the analysis. The number of studies included in meta-analyses varied depending on the outcomes and is presented in Table 4 and in the GRADE summary of findings table (Table 7).

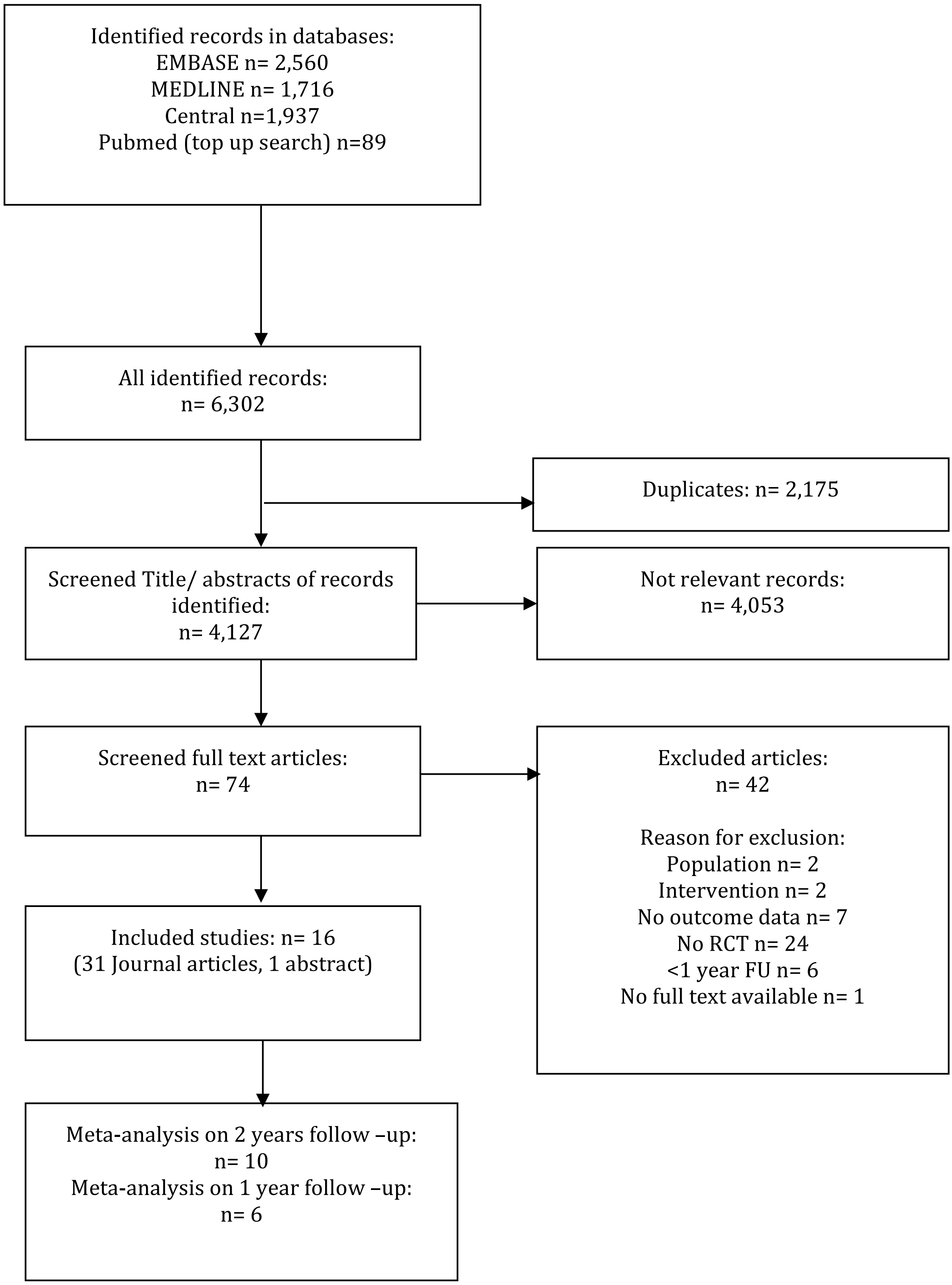


Figure 1 Results on literature search

Table 3 Included RCTs and corresponding publications with a follow-up of at least 2 years

Study identification	Reference
Courcoulas 2014 [33-35]	Courcoulas AP, Goodpaster BH, Eagleton JK, et al. (2014). Surgical vs medical treatments for type 2 diabetes mellitus: a randomized clinical trial. JAMA surgery 149(7):707-15. Courcoulas AP, Belle SH, Neiberg RH, et al. (2015). Three-year outcomes of bariatric surgery vs lifestyle intervention for type 2 diabetes mellitus treatment: a randomized clinical trial. JAMA surgery doi:10.1001/jamasurg.2015.1534.
Dixon 2008 [36]	Dixon, J. B., P. E. O'Brien, et al. (2008). Adjustable gastric banding and conventional therapy for type 2 diabetes: A randomized controlled trial. Obstetrical and Gynecological Survey 63(6): 372-373.
Dixon 2012 [37]	Dixon, J., L. Schachter, et al. (2012). Surgical versus conventional therapy for weight loss treatment of obstructive sleep apnea: A randomized controlled trial. Obesity Research and Clinical Practice 6: 29. Dixon, J. B., L. M. Schachter, et al. (2012). Surgical vs conventional therapy for weight loss treatment of obstructive sleep apnea: A randomized controlled trial. JAMA - Journal of the American Medical Association 308(11): 1142-1149.
Feigel-Cuiller 2015 [38]	Feigel-Guiller B, Drui D, Dimet J, et al. (2015). Laparoscopic Gastric Banding in Obese Patients with Sleep Apnea: A 3-Year Controlled Study and Follow-up After 10 Years. Obes Surg22:22.
Ikramuddin 2015 [39 40]	Ikramuddin, S., J. Korner, et al. (2013). Roux-en-Y gastric bypass vs intensive medical management for the control of type 2 diabetes, hypertension, and hyperlipidemia: the Diabetes Surgery Study randomized clinical trial. JAMA 309(21): 2240-2249. Ikramuddin S, Billington CJ, Lee WJ, et al. (2015). Roux-en-Y gastric bypass for diabetes (the Diabetes Surgery Study): 2-year outcomes of a 5-year, randomised, controlled trial. The Lancet Diabetes & endocrinology 3(6):413-22.
Mingrone 2012 [41 42]	Mingrone, G., S. Panunzi, et al. (2012). Bariatric surgery versus conventional medical therapy for type 2 diabetes. N Engl J Med 366(17): 1577-1585. Pokala, S. (2012). Gastric bypass or biliopancreatic diversion increases remission from type 2 diabetes in obese adults. Annals of Internal Medicine 157(2): 2-12.
O'Brien 2006 [43 44]	O'Brien, P. E., J. B. Dixon, et al. (2006). Treatment of mild to moderate obesity with laparoscopic adjustable gastric banding or an intensive medical program: A randomized trial. Annals of Internal Medicine 144(9): 625-633. Dixon, J. B., B. J. Strauss, et al. (2007). Changes in body composition with weight loss: obese subjects randomized to surgical and medical programs. Obesity (Silver Spring) 15(5): 1187-1198.

Reis 2010 [45 46]	Reis, L. O., W. J. Favaro, et al. (2010). Erectile dysfunction and hormonal imbalance in morbidly obese male is reversed after gastric bypass surgery: A prospective randomized controlled trial. International Journal of Andrology 33(5): 736-744. Oliveira LV, Aguiar IC, Hirata RP, et al. (2011). Sleep study, respiratory mechanics, chemosensitive response and quality of life in morbidly obese patients undergoing bariatric surgery: a prospective, randomized, controlled trial. BMC surgery 11:28.
Schauer 2014 [47-51]	Kashyap SR, Bhatt DL, Schauer PR. (2010). Bariatric surgery vs. advanced practice medical management in the treatment of type 2 diabetes mellitus: rationale and design of the Surgical Therapy And Medications Potentially Eradicate Diabetes Efficiently trial (STAMPEDE). Diabetes, obesity & metabolism 12(5):452-4. Schauer, P. R., S. R. Kashyap, et al. (2012). Bariatric surgery versus intensive medical therapy in obese patients with diabetes. N Engl J Med 366(17): 1567-1576. Kashyap SR, Bhatt DL, Wolski K, et al. (2013). Metabolic effects of bariatric surgery in patients with moderate obesity and type 2 diabetes: analysis of a randomized control trial comparing surgery with intensive medical treatment. Diabetes care 36(8):2175-82. Schauer PR, Bhatt DL, Kirwan JP, et al. (2014). Bariatric surgery versus intensive medical therapy for diabetes - 3-year outcomes. The New England journal of medicine 370(21):2002-13. Malin SK, Samat A, Wolski K, et al. (2014). Improved acylated ghrelin suppression at 2 years in obese patients with type 2 diabetes: Effects of bariatric surgery vs standard medical therapy. International Journal of Obesity38(3):364-70.
Wentworth 2014 [52 53]	Wentworth JM, Playfair J, Laurie C, et al. (2014). Multidisciplinary diabetes care with and without bariatric surgery in overweight people: a randomised controlled trial. The Lancet Diabetes & Endocrinology 2(7):545-52. Wentworth JM, Playfair J, Laurie C, et al. (2015). Gastric Band Surgery Leads to Improved Insulin Secretion in Overweight People with Type 2 Diabetes. Obes Surg 21:21.

Table 4 The number of studies included in meta-analyses for each outcome at 2 years

Study identification	Body weight	Quality of life	$HBA1_c$	Stroke	Myocardial infarction	All-cause mortality	Serious adverse events/ reoperations	Diabetes remission	Hypertension	Dyslipidaemia	Sleep apnoea	Cancer	Revision rates
Courcoulas 2015	x		x			x	x	x					
Dixon 2008	x		x					x					
Dixon 2012	x	x	x			x	x						
Feigel-Guiller 2015	x						x				x		
Ikramuddin 2015	x		x			x		x	x	x		x	
Mingrone 2012	x		x					x					
O'Brien 2006	x												
Reis 2010						x							
Schauer 2014	x		x	x		x		x				x	
Wentworth 2014		x	x					x	x				

3.2.2 Characteristics of the included studies

3.2.2.1 General study characteristics

General study characteristics for studies with 2 or 3 years follow-up are summarized in Table 5 and described in the following section. General study characteristics of studies with maximal 1 year follow-up are presented in Appendix 7.

One RCT included overweight (BMI of 25-30 kg/m²) individuals only **(Wentworth 2014)**, another RCT individuals with a BMI of 27-43 kg/m² **(Schauer 2014)**, one RCT included individuals with a BMI 30-35 kg/m² **(O'Brien 2006)**, three RCTs included individuals with a BMI of 30-40 kg/m² **(Courcoulas 2014; Dixon 2008; Ikramuddin 2015)**, two RCTs included individuals with BMI >35 kg/m² **(Feigel-Guiller 2015; Mingrone 2012)**, one RCT included individuals with BMI 35-55 kg/m² **(Dixon 2012)** and one RCT included individuals with BMI >40 kg/m² **(Reis 2010)**. Six RCTs included solely patients with type 2 diabetes **(Courcoulas 2014; Dixon 2008; Ikramuddin 2015; Mingrone 2012; Schauer 2014; Wentworth 2014),** two RCTs solely patients with obstructive sleep apnoea **(Dixon 2012; Feigel-Guiller 2015)** and one with erectile dysfunction **(Reis 2010)**. The remaining study required at least one obesity-related **(O'Brien 2006)** comorbid condition as inclusion criterion.

In Appendix 3 the inclusion and exclusion criteria of the individual RCTs are described. In most RCTs individuals with severe comorbidities, drug addictions including alcohol or smoking **(Courcoulas 2014; Ding 2015; Dixon 2008; Dixon 2012; Halperin 2014; Ikramuddin 2015; Liang 2013; Mingrone 2002; O'Brien 2006; Reis 2010)** or cognitive or mental impairment **(Courcoulas 2014; Ding 2015; Dixon 2008; Dixon 2012; Halperin 2014; O'Brien 2006; Parikh 2014)** were excluded from study participation. The authors of the RCTs did not always define the term "severe comorbidities".

There were three 3-arm **(Courcoulas 2014; Mingrone 2012; Schauer 2014)** and seven 2-arm RCTs **(Dixon 2008; Dixon 2012; Feigel-Guiller 2015; Ikramuddin 2015; O'Brien 2006; Reis 2010; Wentworth 2014)**. Of the 3-arm studies, all compared gastric bypass with conservative treatment. All gastric bypass operations were done laparoscopically except for one RCT, which did not report on this **(Courcoulas 2014)**. Additionally one RCT included a laparoscopic gastric banding **(Courcoulas 2014)**, one RCT a sleeve gastrectomy **(Schauer 2014)** and one RCT a biliopancreatic diversion treatment arm **(Mingrone 2012)**. Of the 2-arm RCTs, five compared laparoscopic gastric banding **(Dixon 2008; Dixon 2012; Feigel-Guiller 2015; O'Brien 2006; Wentworth 2014)** and two laparoscopic gastric bypass with conservative treatment **(Ikramuddin 2015; Reis 2010)**.

Conservative treatment consisted of lifestyle modification, including reduced caloric intake and increased physical activity, in all studies. Three RCTs reported a medical treatment with orlistat or sibutramine (**Dixon 2008; Ikramuddin 2015; O'Brien 2006**). In six RCTs (**Dixon 2008; Dixon 2012; Feigel-Guiller 2015; Ikramuddin 2015; Schauer 2014; Wentworth 2014**) the surgical intervention was accompanied with the same lifestyle modification intervention as for the conservative treatment group. Details can be found in Appendix 4.

Two RCTs were conducted in the USA **(Courcoulas 2014; Schauer 2014)**. Three RCTs were done in Australia **(Dixon 2008; Dixon 2012; Wentworth 2014)**; one RCT in Italy **(Mingrone 2012)**; one RCT in France (**Feigel-Guiller 2015)** and one RCT in Brazil **(Reis 2010)**. One RCT was done in Taiwan and the USA **(Ikramuddin 2015)**. One RCT did not report the country where it was conducted **(O'Brien 2006)**. Limited information was reported on the setting. Five RCTs were in a single-centre **(Courcoulas 2014; Feigel-Guiller 2015; Mingrone 2012; Reis 2010; Wentworth 2014)**; two were multi-centre RCTs **(Ikramuddin 2015; Dixon 2012)**; the other RCTs did not report whether they were single or multi-centre RCTs. Most RCTs were done in hospitals or university hospitals or specialized centres for the treatment of obesity.

The recruitment periods were all in the time period from 1999 to 2011. The recruit period was not reported for one RCT **(Courcoulas 2014)**. The study periods were lying in the time period from 1999 to 2016 and were not reported for three RCTs **(Ikramuddin 2015; Mingrone 2012;**

Reis 2010). The number of patients randomised in the included studies ranged from 20 (smallest study) (Reis 2010) to 150 (biggest study) **(Schauer 2014)**. Overall, 733 patients were included in these ten studies.

Nine RCTs reported on individuals who refused or withdrew from the intervention **(Courcoulas 2014; Dixon 2008; Dixon 2012; Feigel-Guiller 2015; Ikramuddin 2015; Mingrone 2012, O'Brien 2006; Schauer 2014, Wentworth 2014)**. Of the 346 patients randomised to surgical intervention 10 (2.9%, 95% CI 6.1-13.6%) refused surgery. Thereof 5 of 154 (3.2%), 4 of 122 (3.3%), 0 of 20 (0%) and 1 of 50 (2%) refused RYGB, LAGB, BPD or SG, respectively. For the conservative intervention, the number of patients refusing or who withdrew from the study were 24 of 256 (9.3%, 95% CI 6.1-13.6%).

All ten RCTs reported the number of patients returning for the last visit. The overall adherence rate in the bariatric surgery group was 90.0%, 95% CI 86.7-92.8% (370 of 411) and in the conservative treatment group was 81.4%, 95% CI 76.7-85.5% (262 of 322).

3.2.2.2 Trial populations at baseline

Mean age, mean BMI, mean body weight, mean duration of type 2 diabetes and proportion of female patients, patients with BMI <35 kg/m^2 and comorbidities at baseline are summarized for each study arm in Appendix 2.

Table 5 General study characteristics

Study ID	Population	Comparison	Setting	Recruitment period Study period	n individuals assessed for eligibility and randomized patients who withdrew* from study/refused the intervention**	Follow-up visits Maximal follow-up time	Adherence rate at last follow-up Adherence rate at 2 years
Courcoulas 2014	Individuals with a BMI of 30-40 kg/m² and type 2 diabetes; prior weight loss attempts: n.r. not reported; run in phase: n.r.	Gastric bypass (laparoscopic not reported) vs. laparoscopic gastric banding vs. conservative treatment (for more details see Appendix 4)	Single centre, University of Pittsburgh Medical Center, Pittsburgh, USA	not reported; October 2009 to May 2012	Assessed: n= 667 Randomised: n= 69 RYGB: n= 3 ** LAGB: n= 1** conservative: n= 3 **	RYGB: at 2 weeks, 3, 6, 9, 12, 24 and 36 months LAGB: at 2 weeks, 2, 4, 6, 8, 10, 12, 24 and 36 months; 36 months	Last follow-up: RYGB: 75.0% (18/24) LAGB: 90.9% (20/22) Conservative: 60.9% (14/23) At 2 years: RYGB: 75.0% (18/24) LAGB: 77.3% (17/22) Conservative: 60.9% (14/23)
Dixon 2008	Individuals with a BMI of 30-40 kg/m² and type 2 diabetes; prior weight loss attempts: n.r.; Run in: 3-months, counselling on diet, exercise, glucose self-monitoring and medications	Laparoscopic gastric banding plus conservative treatment vs. conservative treatment	Single centre, "University Obesity Research Center in Australia"	December 2002 to November 2004; December 2002 to December 2006	Assessed: n= 158 Randomised: n= 60 LAGB: n= 1 * conservative: n= 4*	For both study groups: every 6 weeks throughout the 2 years LAGB: every 4-6 weeks assessed by bariatric surgery team; 24 months	Last follow-up: LAGB: 96.7% (29/30) Conservative: 86.7% (26/30) At 2 years: LAGB: 96.7% (29/30) Conservative: 86.7% (26/30)

Dixon 2012	Individuals with a BMI of 35-55 kg/m² and obstructive sleep apnoea; prior weight loss attempts: at least 3 significant attempts; Run-in phase: n.r.	Laparoscopic gastric banding plus conservative treatment vs. conservative treatment	Recruited from 7 sleep clinics in Melbourne, Australia	September 2006 to March 2009; September 2006 to September 2011	Assessed: n= 130 Randomised: n= 60 LAGB: n= 4 ** conservative: n.r.	For both study groups: progress reviewed every 4-6 weeks; 24 months	Last follow-up: LAGB: 93% (28/30) Conservative: 93% (28/30) At 2 years: LAGB: 93% (28/30) Conservative: 93% (28/30)
Feigel-Guiller 2015	Individuals with a BMI>35 kg/m² and obstructive sleep apnoea or obesity-hypoventilation syndrome; prior body weight loss attempts: n.r.; run-in phase: n.r.	Laparoscopic gastric banding plus conservative treatment vs. conservative treatment	Single centre; France, no more details reported	1999 to 2003 (for recruitment and study conduction the same period reported); 1999 to 2003 (for recruitment and study conduction the same period reported)	Assessed: n= 70 Randomised: n= 63 LAGB: n= 1 ** conservative: n= 0 **	For both study groups: monthly for 6 months then every 2 months; 36 months (after completions of RCT, patients were contact at 120 months)	Last follow-up: LAGB: 73.3% (22/30) Conservative: 72.7% (24/33) At 2 years: not reported
Ikramuddin 2015	Individuals with a BMI of 30-39.9 kg/m² and type 2 diabetes; previous body prior body weight loss attempts: n.r.; run-in phase: 2 week (food, exercise and glucose readings)	Laparoscopic gastric bypass plus conservative treatment vs. conservative treatment	Four centres: the University of Minnesota, Columbia University Medical Center, the Mayo Clinic in Rochester, Minnesota in USA and the Min Sheng General Hospital in Taiwan	April 2008 to December 2011; not reported	Assessed: n= 2649 Randomised: n= 120 RYGB: n= 2 * conservative: n= 3*	At 3, 6, 9 12, 15, 18, 21 and 24 months; 24 months	Last follow-up: RYGB: 93% (56/60) Conservative: 90% (54/60) At 2 years: RYGB: 93% (56/60) Conservative: 90% (54/60)

Mingrone 2012	Individuals with a BMI of ≥35 kg/m² and type 2 diabetes; prior body weight loss attempts: n.r.; run-in phase: n.r.	Laparoscopic gastric bypass vs. biliopancreatic diversion vs. conservative treatment	Single centre: Day Hospital of Metabolic Diseases and Diabetology of the Catholic University, Rome, Italy	April 2009 to October 2011; not reported	Assessed: n= 72 Randomised: n= 60 RYGB: n= 0* BPD: n= 0* conservative: n= 2*	At 1, 3, 6, 9, 12 and 24 months; 24 months	Last follow-up: RYGB: 95.0% (19/20) BPD: 95.0% (19/20) Conservative: 90.0% (18/20) At 2 years: RYGB: 95.0% (19/20) BPD: 95.0% (19/20) Conservative: 90.0% (18/20)
O'Brien 2006	Individuals with a BMI of 30-35 kg/m² and an obesity-related comorbid condition; prior body weight loss attempts: over at least 5 years; run-in phase: n.r.	Laparoscopic gastric banding vs. conservative treatment	"patient assessments and outpatient treatments were conducted at a community clinic dedicated to obesity management or in the clinics of a university department of surgery. Surgical procedures were conducted at a private community hospital experienced in the care of bariatric surgical patients."	May 2000 to November 2001; May 2000 to November 2003	Assessed: n= 340 Randomised: n= 80 LAGB: n= 1 * conservative: n= 5*	At least every 6 weeks; 24 months	Last follow-up: LAGB: 97.5% (39/40) Conservative: 82.5% (33/40) At 2 years: LAGB: 97.5% (39/40) Conservative: 82.5% (33/40)

Reis 2010	Morbidly obese men (BMI >40 kg/m²) with erectile dysfunction; prior body weight loss attempts: n.r.; run-in phase: n.r.	Laparoscopic gastric bypass vs. conservative treatment	Single centre: "at University of Campinas", Brazil	First semester of 2007 (author request); not reported	Assessed: n= 117 Randomised: n= 20 not reported	RYGB: weekly during the first year, then monthly; RYGB: 24 months after randomisation (16 months after surgery) Control: 24 months	Last follow-up: not reported At 2 years: not reported
Schauer 2014	Individuals with a BMI of 27-43 kg/m² and type 2 diabetes; prior body weight loss attempts: n.r.; run-in phase: n.r.	Laparoscopic gastric bypass vs. sleeve gastrectomy (each plus conservative treatment) vs. conservative treatment	Single centre; Cleveland Clinic, Cleveland, USA	March 2007 to January 2011; May 2007 to January 2016	Assessed: n= 218 Randomised: n= 150 RYGB: n= 0* SG: n= 1* conservative: n= 7*	At 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 42, 48, 54 and 60 months; 36 months	Last follow-up: RYGB: 96% (48/50) SG: 98% (49/50) Conservative: 80% (40/50) At 2 years: not reported
Wentworth 2014	Individuals with a BMI of 25-30 kg/m² and type 2 diabetes; prior body weight loss attempts: n.r.; run-in phase: n.r.	Laparoscopic gastric banding plus conservative treatment vs. conservative treatment	Single centre; Institution not reported; Melbourne, Australia	November 2009 to June 2011; November 2009 to June 2013	Assessed: n= 1231 Randomised: n= 51 LAGB: n= 2** control: n.r.	Both groups: every 3 months in year 1, and every 6 months in year 2; 24 months	Last follow-up: LAGB: 92.0% (23/25) Conservative: 96.2% (25/26) At 2 years: LAGB: 92.0% (23/25) Conservative: 96.2% (25/26)

*Authors reported number of patients who withdrew
**Authors reported number of patients who refused intervention
Abbreviations:
BMI: body mass index
BPD: Biliopancreatic diversion
LAGB: Laparoscopic adjustable gastric banding
n.r.: not reported
RYGB: Roux-en-Y gastric bypass
SG: sleeve gastrectomy
vs.: versus

3.2.3 Results on risk of bias assessment

The method for the random sequence generation was unclear in one RCT only (**Feigel-Guiller 2015**). Allocation concealment was adequate in three RCTs (low risk for selection bias) (**O'Brien 2006; Reis 2010; Wentworth 2014**), and unclear in five RCTs (unclear risk for selection bias) (**Dixon 2008; Dixon 2012; Feigel-Guiller 2015; Ikramuddin 2015; Schauer 2014**). Performance bias due to lack of blinding was rated as high in six (**Dixon 2008; Ikramuddin 2015; Mingrone 2012; O'Brien 2006; Schauer 2014; Wentwort 2014**) and as unclear in four RCTs (**Courcoulas 2015; Dixon 2012; Feigel-Guiller 2015; Reis 2010**). The risk for performance bias was not judged differently depending on outcome. Risk of detection bias was rated as low in one RCT (**Ikramuddin 2015**), as high in six (**Dixon 2008; Dixon 2012; Mingrone 2012; O'Brien 2006; Schauer 2014; Wentwort 2014**) and as unclear in three RCTs (**Courcoulas 2015; Feigel-Guiller 2015; Reis 2010**). The risk for detection bias was not judged differently depending on outcome here. However, for the quality of evidence assessment the aspect of detection bias was taken into consideration differently depending on outcome. Attrition bias for continuous outcome data was rated as low in five (**Dixon 2012; Ikramuddin 2015; Mingrone 2012; Reis 2010; Wentwort 2014**) and as high in five RCTs (**Courcoulas 2015; Dixon 2008; Feigel-Guiller 2015; O'Brien 2006; Schauer 2014).** Attrition bias for binary outcome data was rated as low in five (**Dixon 2012; Ikramuddin 2015; Mingrone 2012; Reis 2010; Wentwort 2014**) and as high in four RCTs (**Courcoulas 2015; Dixon 2008; Feigel-Guiller 2015; Schauer 2014**). One RCT did not report any binary outcomes (**O'Brien 2006**). Reporting bias was rated as low in all ten RCTs. A summarized overview of the risk of bias assessment is shown in Table 6. A detailed description of the risk of bias assessment including the support for judgement is provided in Appendix 5. Risk of bias results for studies with maximal 1 year follow-up are presented in Appendix 8.

Table 6 Risk of bias – 2 year studies

Study	**Random sequence generation (selection bias)**	**Allocation concealment (selection bias)**	**Blinding of participants and personnel (performance bias)**	**Blinding of outcome assessment, judgement did not differ among outcomes (detection bias)**	**Incomplete continuous outcome data (attrition bias)**	**Incomplete binary data (attrition bias)**	**Selective reporting (reporting bias)**
Courcoulas 2015	low	high	unclear	unclear	high	high	low
Dixon 2008	low	unclear	high	high	high	high	low
Dixon 2012	low	unclear	unclear	high	low	low	low
Feigel-Guiller 2015	unclear	unclear	unclear	unclear	high	high	low
Ikramuddin 2015	low	unclear	high	low	low	low	low
Mingrone 2012	low	high	high	high	low	low	low
O'Brien 2006	low	low	high	high	high	not applicable*	low
Reis 2010	Low	low	unclear	unclear	low	low	low
Schauer 2014	Low	unclear	high	high	high	high	low
Wentworth 2014	Low	low	high	high	low	low	low
*no relevant binary outcome identified							

3.2.4 Results on clinical effectiveness, safety and quality of evidence

The results on clinical effectiveness, safety and quality of evidence are limited to ten RCTs reporting outcomes after a follow-up period of 2 or 3 years. The duration of follow-up for the reported results was 2 years in seven RCTs, 3 years in one RCT, and 2 and 3 years in two RCTs (only the results at 3 years were considered). In the following section the results for the critical and important outcomes are summarized. In addition, results were stratified by the different surgical techniques and, if applicable, the results of the sensitivity analyses are presented. Though results were stratified by surgical technique, effect estimates were not compared statistically by subgroup analyses because of the small number of studies within strata and the inclusion of 3-arm studies.

The pooled results of the 6-months to 1-year results are presented in Appendix 9.

3.2.4.1 **Critical outcomes**

3.2.4.1.1 Body weight

Bariatric surgery (overall effect irrespective of type of surgery)

Percent change of body weight was reported by eight RCTs at 2 or 3 years (**Courcoulas 2015; Dixon 2008; Dixon 2012; Feigel-Guiller 2015; Ikramuddin 2015; Mingrone 2012; O'Brien 2006; Schauer 2014**). Body weight decreased more in the bariatric surgery than in the conservative treatment group. The mean difference (MD) in percent change of body weight was -17.9%-points (95% CI -21.4 to -14.5%-points). The heterogeneity was considerable (I^2=81%). A sensitivity analysis identified **Mingrone 2012** as the main source of heterogeneity. Excluding Mingrone 2012 from the analysis resulted in a similar effect size and reduced heterogeneity. Heterogeneity might no longer be important (MD -16.7%-points, 95% CI –18.4 to -14.9%-points, I^2=15%).

Quality of the evidence according to GRADE: The quality of evidence was moderate. The quality of evidence was downgraded because we are considering the risk of bias for this summary as substantial. We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.[14]

Gastric bypass

Four RCTs **(Courcoulas 2015; Ikramuddin 2015; Mingrone 2012; Schauer 2014)** compared gastric bypass with conservative treatment and reported percent change of body weight (Figure 2). Body weight decreased more in the gastric bypass than in the conservative treatment groups. The mean difference in percent body weight change was -21.2%-points (95% CI -26.0 to -16.4%-points). Heterogeneity was considerable (I^2=78%). Similar to the results described above, a sensitivity analysis identified **Mingrone 2012** as main source of heterogeneity. Excluding Mingrone 2012 from the analysis resulted in a similar effect size and reduced heterogeneity. Heterogeneity was no longer detected (MD -19.1%-points, 95% CI –21.6 to -16.6%, I^2=0%).

Adjustable gastric banding

Five RCTs (**Courcoulas 2015; Dixon 2008; Dixon 2012; Feigel-Guiller 2015; O'Brien 2006**) compared adjustable gastric banding with conservative treatment and reported percent change of body weight (Figure 2). Body weight decreased more in the adjustable gastric banding than in the conservative treatment group. The mean difference in percent change of body weight was -15.0%-points (95% CI -18.2 to -11.8%-points). The heterogeneity was moderate to substantial (I^2=59%). Sensitivity analyses did not identify a particular cause for heterogeneity.

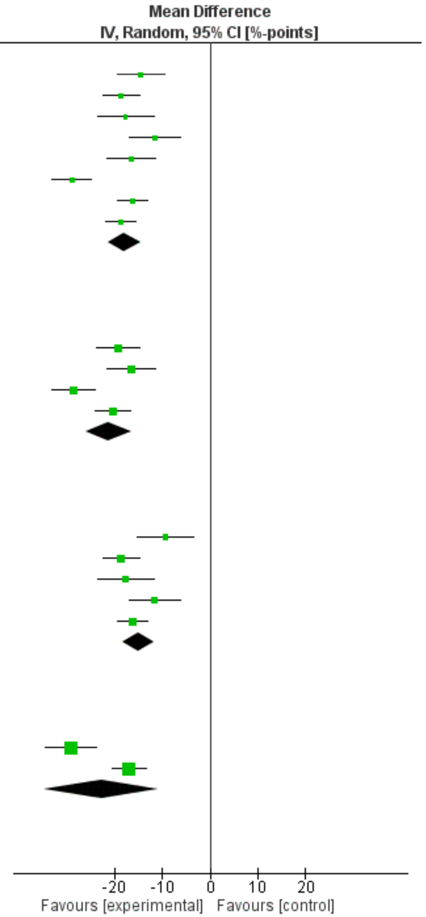
Other surgical interventions

In two RCTs (**Mingrone 2012; Schauer 2014**) other interventions, namely biliopancreatic diversion and sleeve gastrectomy, were compared to conservative treatment (Figure 2). Body weight decreased more in the group "other surgical interventions" than in the conservative treatment group. The mean difference was -22.8%-points (95% CI -34.8 to -10.9%-points). Heterogeneity was considerable (I^2=93%). Due to the limited number of studies in this stratum, heterogeneity was not investigated in subgroup analyses.

Results of the meta-analysis are presented in Figure 2 and the quality of evidence is summarized in Table 7.

Meta-analysis after 1 year is presented in Appendix 9.

	Bariatric surgery			Control				Mean Difference
Study or Subgroup	Mean [%-points]	SD [%-points]	Total	Mean [%-points]	SD [%-points]	Total	Weight	IV, Random, 95% CI [%-points]
3.1.1 Bariatric surgery								
Courcoulas 2015 (1)	-20.2	8.2	46	-5.7	10.8	23	11.9%	-14.50 [-19.51, -9.49]
Dixon 2008	-20	9.4	30	-1.4	4.9	30	13.3%	-18.60 [-22.39, -14.81]
Dixon 2012	-20.6	13.8	30	-2.9	9	30	10.8%	-17.70 [-23.60, -11.80]
Feigel-Guiller 2015 (2)	-14.3	9	22	-2.8	9.3	24	11.5%	-11.50 [-16.79, -6.21]
Ikramuddin 2015	-23.8	13.9	60	-7.3	14.9	59	11.6%	-16.50 [-21.68, -11.32]
Mingrone 2012	-33.6	9	38	-4.7	6.4	18	12.9%	-28.90 [-33.01, -24.79]
O'Brien 2006 (3)	-21.6	7.1	39	-5.5	6.4	31	14.0%	-16.10 [-19.27, -12.93]
Schauer 2014 (4)	-22.8	9.1	97	-4.2	8.3	40	14.0%	-18.60 [-21.75, -15.45]
Subtotal (95% CI)			**362**			**255**	**100.0%**	**-17.94 [-21.36, -14.51]**
Heterogeneity: Tau^2 = 19.19; Chi^2 = 36.46, df = 7 (P < 0.00001); I^2 = 81%								
Test for overall effect: Z = 10.27 (P < 0.00001)								
3.1.2 Gastric bypass								
Courcoulas 2015 (5)	-25	2.8	24	-5.7	10.8	23	24.9%	-19.30 [-23.85, -14.75]
Ikramuddin 2015	-23.8	13.9	60	-7.3	14.9	59	23.3%	-16.50 [-21.68, -11.32]
Mingrone 2012	-33.3	7.9	19	-4.7	6.4	18	24.7%	-28.60 [-33.22, -23.98]
Schauer 2014 (6)	-24.5	9.1	48	-4.2	8.3	40	27.1%	-20.30 [-23.94, -16.66]
Subtotal (95% CI)			**151**			**140**	**100.0%**	**-21.22 [-26.01, -16.42]**
Heterogeneity: Tau^2 = 18.69; Chi^2 = 13.91, df = 3 (P = 0.003); I^2 = 78%								
Test for overall effect: Z = 8.67 (P < 0.00001)								
3.1.3 Adjustable gastric banding								
Courcoulas 2015 (7)	-15	9.1	22	-5.7	10.8	23	16.3%	-9.30 [-15.13, -3.47]
Dixon 2008	-20	9.4	30	-1.4	4.9	30	23.5%	-18.60 [-22.39, -14.81]
Dixon 2012	-20.6	13.8	30	-2.9	9	30	16.1%	-17.70 [-23.60, -11.80]
Feigel-Guiller 2015 (8)	-14.3	9	22	-2.8	9.3	24	18.0%	-11.50 [-16.79, -6.21]
O'Brien 2006 (9)	-21.6	7.1	39	-5.5	6.4	31	26.1%	-16.10 [-19.27, -12.93]
Subtotal (95% CI)			**143**			**138**	**100.0%**	**-15.01 [-18.23, -11.80]**
Heterogeneity: Tau^2 = 7.69; Chi^2 = 9.78, df = 4 (P = 0.04); I^2 = 59%								
Test for overall effect: Z = 9.15 (P < 0.00001)								
3.1.4 Other surgical techniques								
Mingrone 2012	-33.8	10.2	19	-4.7	6.37	18	48.5%	-29.10 [-34.55, -23.65]
Schauer 2014 (10)	-21.1	8.9	49	-4.2	8.3	40	51.5%	-16.90 [-20.48, -13.32]
Subtotal (95% CI)			**68**			**58**	**100.0%**	**-22.82 [-34.77, -10.87]**
Heterogeneity: Tau^2 = 68.89; Chi^2 = 13.45, df = 1 (P = 0.0002); I^2 = 93%								
Test for overall effect: Z = 3.74 (P = 0.0002)								



Footnotes
(1) Data from 3 years follow-up
(2) Data from 3 years follow-up, Author requested data 07.09.2015
(3) Extracted from text, O'Brien 2006
(4) Data from 3 years follow-up
(5) Data from 3 years follow-up
(6) Data from 3 years follow-up
(7) Data from 3 years follow-up
(8) Data from 3 years follow-up, Author requested data 07.09.2015
(9) Extracted from text, O'Brien 2006
(10) Data from 3 years follow-up

Figure 2 Percent change of body weight – 2 years results

3.2.4.1.2 Quality of life

Quality of life was reported by three RCTs at 2 years after intervention (**Dixon 2012; Wentworth 2014; Schauer 2014**). The data of two RCTs, which used the SF-36 questionnaire, were pooled. Both RCTs compared adjustable gastric banding with conservative treatment. Because no overall summary measure of the SF-36 questionnaire was reported, the summary component measures of mental health and physical health were pooled. The third RCT (**Schauer 2014**) reporting on quality of life at 2 years used the RAND-36 questionnaire. The results deriving from the RAND-36 questionnaire could not be pooled with the other two RCTs because no summary component measures were given; i.e. only results on the individual questions.

The pooled results showed that physical health improved significantly more in the adjustable gastric banding than in conservative treatment groups. The mean difference (MD) was 9.4 (95% CI 5.7 to 13.0). Heterogeneity might not be important (I^2=0%). There was no significant differential effect on mental health between adjustable gastric banding and conservative treatment (MD -0.1, 95% CI -3.9 to 3.8). Heterogeneity might not be important (I^2=0%).

The RCT (**Schauer 2014**) reporting results on RAND-36 showed significant improvements after gastric bypass for the physical health components: physical function, role limitations due to physical health, bodily pain and general health, but after sleeve gastrectomy only for general health. Schauer 2014 further showed significant improvements after gastric bypass for the mental health components: energy/fatigue and emotional well-being, but not for role limitations due to emotional problems or social role functioning. After sleeve gastrectomy there were significant improvements for the component energy/fatigue.

Quality of the evidence according to GRADE: The quality of evidence for physical health was low, because of risk of bias. It was not downgraded for imprecision, because the lower end of the 95% CI indicated a clinically important effect (mean difference of > 5). The quality of evidence for mental health was very low due to serious risk of bias and serious imprecision. Our confidence in the effect estimate for physical health is limited: The true effect may be substantially different from the estimate of the effect.[14] Moreover, we have very little confidence in the effect estimate for mental health: The true effect is likely to be substantially different from the estimate of effect.[14]

The results of the meta-analysis are presented in Figure 3. The quality of evidence is summarized in Table 7.

Pooled results after 1 year are presented in Appendix 9.

Study or Subgroup	Bariatric surgery Mean [SF36 score]	Bariatric surgery SD [SF36 score]	Bariatric surgery Total	Control Mean [SF36 score]	Control SD [SF36 score]	Control Total	Weight	Mean Difference IV, Random, 95% CI [SF36 score]
3.3.1 Bariatric surgery, QoL SF-36 physical function >= 2 years								
Dixon 2012	12.6	14.2	30	3.4	13.4	30	27.0%	9.20 [2.21, 16.19]
Wentworth 2014	7.7	6.2	23	-1.7	8.7	25	73.0%	9.40 [5.15, 13.65]
Subtotal (95% CI)			**53**			**55**	**100.0%**	**9.35 [5.72, 12.98]**
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 0.96); I² = 0%								
Test for overall effect: Z = 5.05 (P < 0.00001)								
3.3.2 Adjustable gastric banding, QoL SF-36 physical funktion >= 2 years								
Dixon 2012	12.6	14.2	30	3.4	13.4	30	27.0%	9.20 [2.21, 16.19]
Wentworth 2014	7.7	6.2	23	-1.7	8.7	25	73.0%	9.40 [5.15, 13.65]
Subtotal (95% CI)			**53**			**55**	**100.0%**	**9.35 [5.72, 12.98]**
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 0.96); I² = 0%								
Test for overall effect: Z = 5.05 (P < 0.00001)								
3.3.3 Bariatric surgery, QoL SF-36 mental health >=2 years								
Dixon 2012	0.5	9.4	30	0.8	8	30	74.6%	-0.30 [-4.72, 4.12]
Wentworth 2014	-0.13	15.4	23	-0.82	10.7	25	25.4%	0.69 [-6.87, 8.25]
Subtotal (95% CI)			**53**			**55**	**100.0%**	**-0.05 [-3.86, 3.77]**
Heterogeneity: Tau² = 0.00; Chi² = 0.05, df = 1 (P = 0.82); I² = 0%								
Test for overall effect: Z = 0.02 (P = 0.98)								
3.3.4 Adjustable gastric banding, QoL SF-36 mental health >=2 years								
Dixon 2012	0.5	9.4	30	0.8	8	30	74.6%	-0.30 [-4.72, 4.12]
Wentworth 2014	-0.13	15.4	23	-0.82	10.7	25	25.4%	0.69 [-6.87, 8.25]
Subtotal (95% CI)			**53**			**55**	**100.0%**	**-0.05 [-3.86, 3.77]**
Heterogeneity: Tau² = 0.00; Chi² = 0.05, df = 1 (P = 0.82); I² = 0%								
Test for overall effect: Z = 0.02 (P = 0.98)								

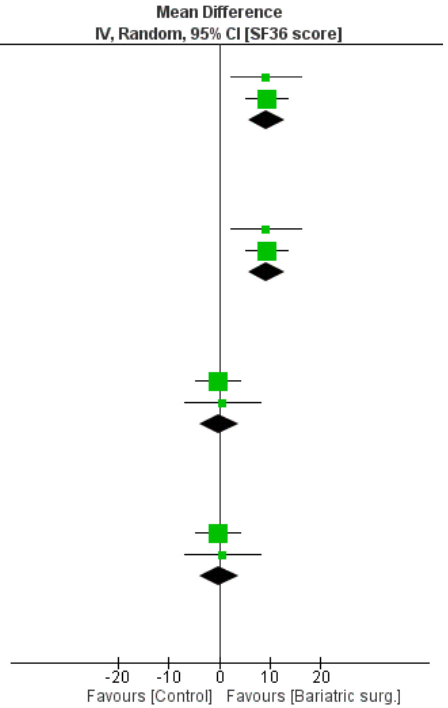


Figure 3 Quality of life, physical health and mental health component - 2 year results

3.2.4.1.3 HbA1c

Bariatric surgery

Change from baseline in HbA1c was pooled for seven RCTs at 2 or 3 years (**Courcoulas 2015; Dixon 2008; Dixon 2012; Ikramuddin 2015; Mingrone 2012; Schauer 2014; Wentworth 2014**). HbA1c decreased more in the bariatric surgery group than in the conservative treatment group. The mean difference was -1.40%-points (95% CI -1.92 to -0.87%-points). The heterogeneity was substantial to considerable (I^2=76%). A sensitivity analysis identified **Dixon 2012** (only RCT in the analysis that was not done in individuals with type 2 diabetes) and **Wentworth 2014** (only study including overweight, but not obese individuals) as main sources of heterogeneity. After excluding Dixon 2012 and Wentworth 2014, the mean difference was -1.58%-points (95% CI –2.03 to -1.12%-points) and heterogeneity was no longer important (I^2=8%). The result of this sensitivity analysis was used for grading the quality of evidence.

Quality of the evidence according to GRADE: The quality of evidence was low because of serious risk of bias and serious indirectness. Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.[14] Indirectness was graded as serious, because the results for absolute effects apply only to obese individuals with type 2 diabetes but not to the general population of obese or overweight individuals as defined by the PICO-question of the present report. Inconsistency was judged as not serious because heterogeneity is explained by Dixon 2012 and Wentworth 2014. It was not downgraded for imprecision because the 95% confidence interval excluded zero (=no effect).

Gastric bypass

Four RCTs (**Courcoulas 2015; Ikramuddin 2015; Mingrone 2012; Schauer 2014**) compared gastric bypass with conservative treatment (Figure 4). HbA1c decreased more in the gastric bypass than in the conservative treatment group. The mean difference was -1.94%-points (95% CI –2.32 to –1.56%-points). Heterogeneity might not be important (I^2=0%).

Adjustable gastric banding

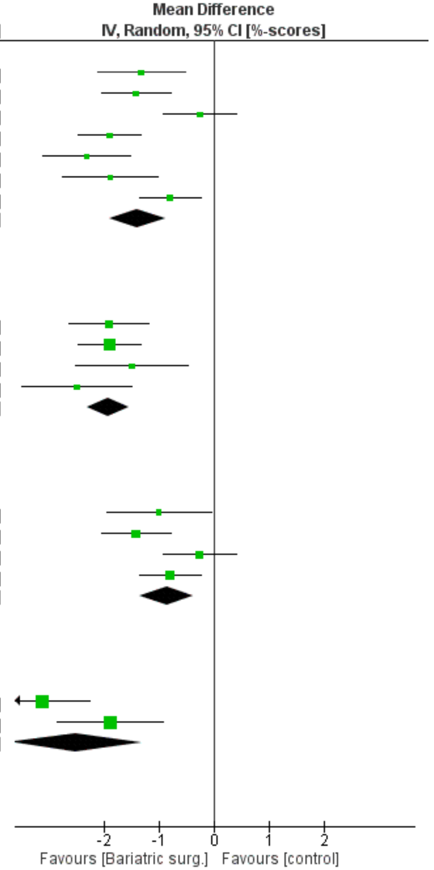
Four RCTs (**Courcoulas 2015; Dixon 2008; Dixon 2012; Wentworth 2014**) compared adjustable gastric banding with conservative treatment (Figure 4). HbA1c decreased more in the adjustable gastric banding than in the conservative treatment group. The mean difference was -0.87%-points (95% CI –1.36 to -0.37%-points). Heterogeneity was moderate (I^2=52%). Similar to the results described above, the sensitivity analysis identified **Dixon 2012** and **Wentworth 2014** as main source of heterogeneity. After excluding Dixon 2012 and Wentworth 2014 from the analysis, the mean difference was -1.30%-points (95% CI –1.83 to -0.77%-points) and Heterogeneity might not be important anymore (I^2=8%).

Other surgical interventions

In two RCTs **(Mingrone 2012; Schauer 2014)** other interventions, namely biliopancreatic diversion and sleeve gastrectomy, were compared to conservative treatment (Figure 4). HbA1c decreased more in the group "other surgical interventions" than in the conservative treatment group. The mean difference was -2.54%-points (95% CI -3.75 to -1.32%-points). The heterogeneity was substantial (I^2=71%). Due to the limited number of studies in this stratum, sources of heterogeneity were not investigated.

The results of the meta-analysis are presented in Figure 4 and the quality of evidence is summarized in Table 7. Pooled results after 1 year are presented in Appendix 9.

Study or Subgroup	Bariatric surgery Mean [%-scores]	Bariatric surgery SD [%-scores]	Bariatric surgery Total	Control Mean [%-scores]	Control SD [%-scores]	Control Total	Weight	Mean Difference IV, Random, 95% CI [%-scores]
3.5.1 Bariatric surgery								
Courcoulas 2015 (1)	-1.12	1.06	46	0.21	1.79	23	13.3%	-1.33 [-2.12, -0.54]
Dixon 2008	-1.81	1.24	30	-0.38	1.26	30	15.0%	-1.43 [-2.06, -0.80]
Dixon 2012	-0.27	1.63	30	-0.01	0.95	30	14.5%	-0.26 [-0.94, 0.42]
Ikramuddin 2015 (2)	-3.1	1.2	60	-1.2	1.9	59	15.6%	-1.90 [-2.47, -1.33]
Mingrone 2012 (3)	-3.14	1.98	38	-0.82	1.07	18	13.2%	-2.32 [-3.12, -1.52]
Schauer 2014 (4)	-2.5	1.99	97	-0.6	2.5	40	12.5%	-1.90 [-2.77, -1.03]
Wentworth 2014	-0.8	0.69	23	0	1.21	25	15.8%	-0.80 [-1.35, -0.25]
Subtotal (95% CI)			**324**			**225**	**100.0%**	**-1.40 [-1.92, -0.87]**
Heterogeneity: Tau² = 0.37; Chi² = 24.62, df = 6 (P = 0.0004); I² = 76%								
Test for overall effect: Z = 5.24 (P < 0.00001)								
3.5.2 Gastric bypass								
Courcoulas 2015 (5)	-1.71	0.17	24	0.21	1.79	23	27.1%	-1.92 [-2.65, -1.19]
Ikramuddin 2015 (6)	-3.1	1.2	60	-1.2	1.9	59	44.7%	-1.90 [-2.47, -1.33]
Mingrone 2012 (7)	-2.32	2.02	19	-0.82	1.07	18	13.7%	-1.50 [-2.53, -0.47]
Schauer 2014 (8)	-3.1	1.38	18	-0.6	2.5	40	14.5%	-2.50 [-3.50, -1.50]
Subtotal (95% CI)			**121**			**140**	**100.0%**	**-1.94 [-2.32, -1.56]**
Heterogeneity: Tau² = 0.00; Chi² = 1.91, df = 3 (P = 0.59); I² = 0%								
Test for overall effect: Z = 9.93 (P < 0.00001)								
3.5.3 Adjustable gastric banding								
Courcoulas 2015 (9)	-0.8	1.47	22	0.21	1.79	23	17.3%	-1.01 [-1.97, -0.05]
Dixon 2008	-1.81	1.24	30	-0.38	1.26	30	27.0%	-1.43 [-2.06, -0.80]
Dixon 2012	-0.27	1.63	30	-0.01	0.95	30	25.5%	-0.26 [-0.94, 0.42]
Wentworth 2014	-0.8	0.69	23	0	1.21	25	30.2%	-0.80 [-1.35, -0.25]
Subtotal (95% CI)			**105**			**108**	**100.0%**	**-0.87 [-1.36, -0.37]**
Heterogeneity: Tau² = 0.13; Chi² = 6.29, df = 3 (P = 0.10); I² = 52%								
Test for overall effect: Z = 3.45 (P = 0.0006)								
3.5.4 Other surgical techniques								
Mingrone 2012 (10)	-3.96	1.61	19	-0.82	1.07	18	51.5%	-3.14 [-4.02, -2.26]
Schauer 2014 (11)	-2.5	2.1	49	-0.6	2.5	40	48.5%	-1.90 [-2.87, -0.93]
Subtotal (95% CI)			**68**			**58**	**100.0%**	**-2.54 [-3.75, -1.32]**
Heterogeneity: Tau² = 0.55; Chi² = 3.45, df = 1 (P = 0.06); I² = 71%								
Test for overall effect: Z = 4.10 (P < 0.0001)								



Footnotes
(1) Data from 3 years follow-up
(2) Extracted from figure 2 and table 2 from Ikramuddin 2015
(3) Calculations based on data received from author inquiry 14/9/2013
(4) Data from 3 years follow-up
(5) Data from 3 years follow-up
(6) Extracted from figure 2 and table 2 from Ikramuddin 2015
(7) Calculations based on data received from author inquiry 14/9/2013
(8) Data from 3 years follow-up
(9) Data from 3 years follow-up
(10) Calculations based on data received from author inquiry 14/9/2013
(11) Data from 3 years follow-up

Figure 4 HbA1c – 2 years results

3.2.4.1.4 Stroke

Stroke was reported by one RCT with a follow-up of 3 years (**Schauer 2014**) comparing sleeve gastrectomy with conservative treatment. There was one stroke after bariatric surgery (1 of 99) and none after conservative treatment (0 of 43) (Figure 5).

Quality of the evidence according to GRADE: The quality of evidence was very low because of serious risk of bias, serious indirectness (effect estimate is limited to obese population with type 2 diabetes) and serious imprecision (very few events were observed in the surgical and conservative treatment groups). We have very little confidence in the effect estimate of stroke: The true effect is likely to be substantially different from the estimate of effect.[14]

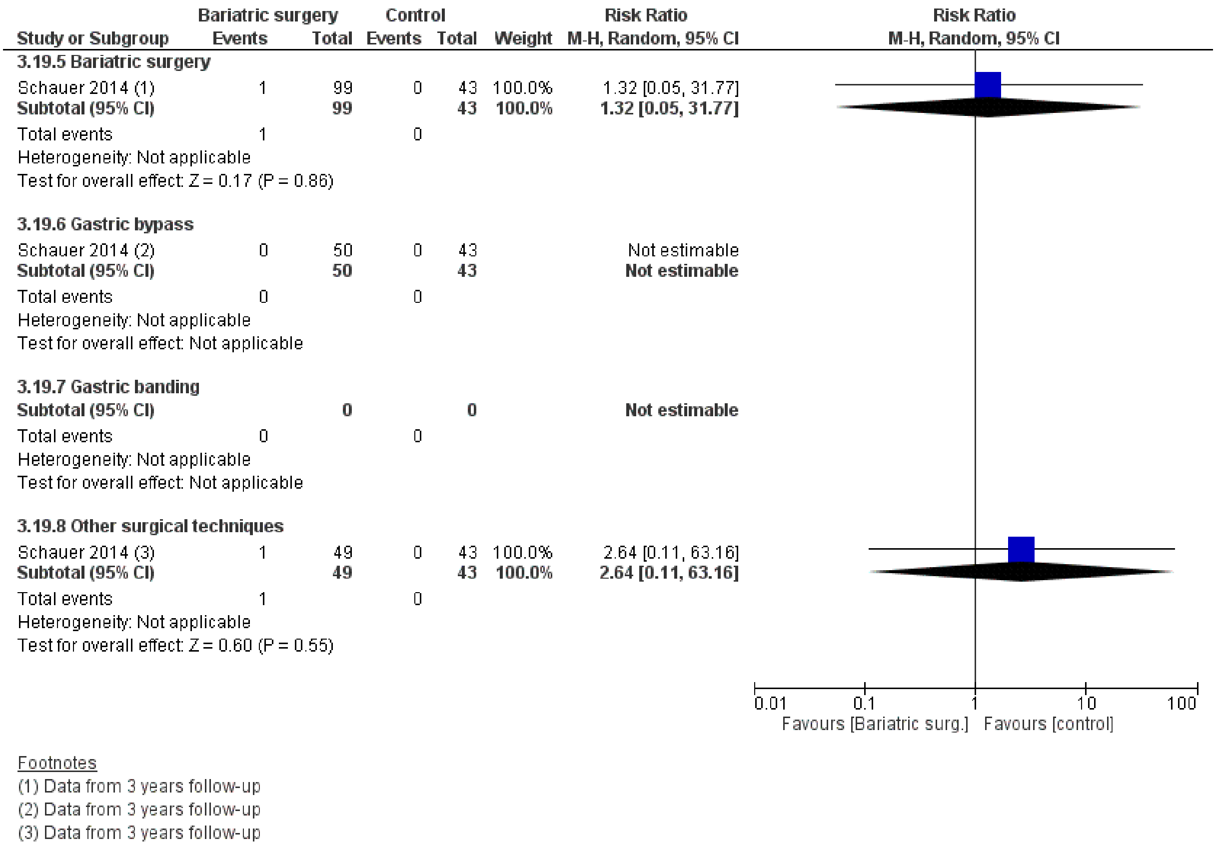


Figure 5 Stroke – 3 years results

3.2.4.1.5 Myocardial infarction

Myocardial infarction was not reported in the RCTs with a follow-up of at least 2 years.

3.2.4.2 **Important outcomes**

3.2.4.2.1 All-cause mortality

Bariatric surgery

All-cause mortality was reported by five RCTs at 2 or 3 years (**Courcoulas 2015; Dixon 2012; Ikramuddin 2015; Reis 2010; Schauer 2014**). One death (from pancreatic cancer) occurred in one of the five RCTs (**Ikramuddin 2015)** in the conservative treatment group. There was no significant difference between the surgical and conservative treatment group (RR 0.33, 95% CI 0.01 to 8.0). It was not possible to estimate heterogeneity.

Quality of the evidence according to GRADE: The quality of evidence was very low because of serious risk of bias and serious imprecision (very few events were observed in the surgical and conservative treatment groups). We have very little confidence in the estimated effect estimate of on all-cause mortality: The true effect is likely to be substantially different from the estimated of effect.[14]

Gastric bypass

Four studies (**Courcoulas 2015; Ikramuddin 2015; Reis 2010; Schauer 2014**) compared gastric bypass with conservative treatment and reported mortality (Figure 6). One death among these studies was reported, and hence, the estimated effect does not differ from the one reported above.

Adjustable gastric banding

Two studies (**Courcoulas 2015; Dixon 2012**) compared adjustable gastric banding with conservative treatment and reported mortality (Figure 6). No death was reported and hence the effect could not be estimated.

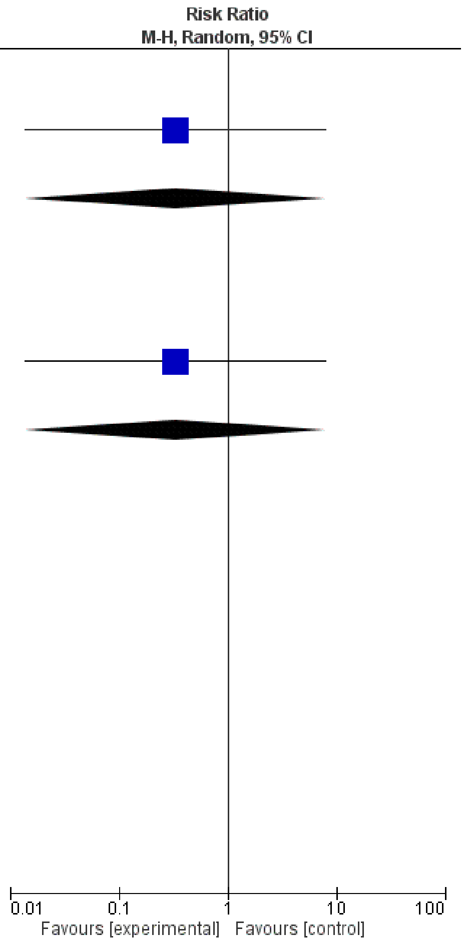
Other surgical interventions

One study (**Schauer 2014**) compared other surgical interventions (sleeve gastrectomy) with conservative treatment and reported no death during the study observation period (Figure 6).

The results of the meta-analysis are presented in Figure 6 and the quality of evidence is summarized in Table 7.

Pooled results after 1 year are presented in Appendix 9.

Study or Subgroup	Bariatric surgery Events	Bariatric surgery Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
3.6.1 Bariatric surgery						
Courcoulas 2015 (1)	0	50	0	25		Not estimable
Dixon 2012	0	30	0	30		Not estimable
Ikramuddin 2015	0	60	1	60	100.0%	0.33 [0.01, 8.02]
Reis 2010	0	10	0	10		Not estimable
Schauer 2014 (2)	0	99	0	43		Not estimable
Subtotal (95% CI)		**249**		**168**	**100.0%**	**0.33 [0.01, 8.02]**
Total events	0		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.68 (P = 0.50)						
3.6.2 Gastric bypass						
Courcoulas 2015 (3)	0	25	0	25		Not estimable
Ikramuddin 2015	0	60	1	60	100.0%	0.33 [0.01, 8.02]
Reis 2010	0	10	0	10		Not estimable
Schauer 2014 (4)	0	50	0	43		Not estimable
Subtotal (95% CI)		**145**		**138**	**100.0%**	**0.33 [0.01, 8.02]**
Total events	0		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.68 (P = 0.50)						
3.6.3 Gastric banding						
Courcoulas 2015 (5)	0	25	0	25		Not estimable
Dixon 2012	0	30	0	30		Not estimable
Subtotal (95% CI)		**55**		**55**		**Not estimable**
Total events	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
3.6.4 Other surgical techniques						
Schauer 2014 (6)	0	49	0	43		Not estimable
Subtotal (95% CI)		**49**		**43**		**Not estimable**
Total events	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						



Footnotes
(1) Author requested data 7.9.2015
(2) Data from 3 years follow-up
(3) Author requested data 7.9.2015
(4) Data from 3 years follow-up
(5) Author requested data 7.9.2015
(6) Data from 3 years follow-up

Figure 6 All-cause mortality – 2 year results

3.2.4.2.2 Serious adverse events including reoperations

Bariatric surgery

Serious adverse events (SAE) including reoperations were reported by three RCTs (**Courcoulas 2015; Dixon 2012; Feigel-Guiller 2015**) at 2 or 3 years. There was no significant difference between the surgical and conservative treatment group (RR 2.4, 95% CI 0.6 to 9.8). Heterogeneity might not be important (I^2=29%).

Quality of the evidence according to GRADE: The quality of evidence was very low because of serious risk of bias and serious imprecision. Imprecision was judged to be serious because of the low event rate and the small number of patients included in the present meta-analysis.

Gastric bypass

One study compared gastric bypass with conservative treatment and reported serious adverse events including reoperations (Figure 7). This study reported serious adverse events including reoperation at 2 and 3 years, and only the 3-year results were used for the analysis. For the outcome serious adverse event including reoperation, one event was reported. The relative risk to suffer a serious adverse event including reoperation after gastric bypass was not significantly different compared to conservative treatment (RR 3.0, 95% CI 0.1 to 69.5). Heterogeneity could not be estimated.

Adjustable gastric banding

Three studies compared adjustable gastric banding with conservative treatment (Figure 7). The relative risk to suffer a serious adverse event including reoperation after adjustable gastric banding was 3.2 (95% CI 0.6 to 17.7) and was not significantly different from conservative treatment. Heterogeneity was moderate to substantial (I^2=47%).

Other surgical interventions

No study comparing other surgical interventions with conservative treatment reported serious adverse events including reoperation.

The results of the meta-analysis are presented in Figure 7, and the quality of evidence is presented as summary in Table 7.

The combined outcome serious adverse events (SAE) and reoperations had two limitations. Firstly, reoperations can only occur in the bariatric surgery group and secondly, several studies reported reoperations without reporting SAE. Therefore, SAE and reoperations were analysed separately in the following sensitivity analysis.

Reoperations – sensitivity analysis

Seven studies reported reoperations. Five reported reoperations until the follow-up at 2 years and two reported reoperations until the follow-up at 3 years. Two-hundred twenty three patients were assigned to the bariatric surgery arms. Of those, 7.6% (17/223) had a reoperation. The absolute risk to be reoperated was 7% (95% CI 3 to 11%). When stratifying by surgical technique, reoperation rates were 2.6% (1 of 39) after gastric bypass; 9.1% (15 of 165) after gastric banding and 5.3% (1 of 19) after biliopancreatic diversion.

Serious adverse events – sensitivity analysis

Three studies reported SAE, one study reported SAE at 2 years and one study at 3 years. Reoperations were excluded from the analysis, even though they were reported as SAE. One study reported only reoperations as SAE, hence no events were extracted for this study. Based on the pooled result of two studies, there was no significant difference between bariatric surgery and conservative treatment (RR 1.40, 95% CI 0.23 to 8.52).

Pooled results after 1 year are presented in Appendix 9.

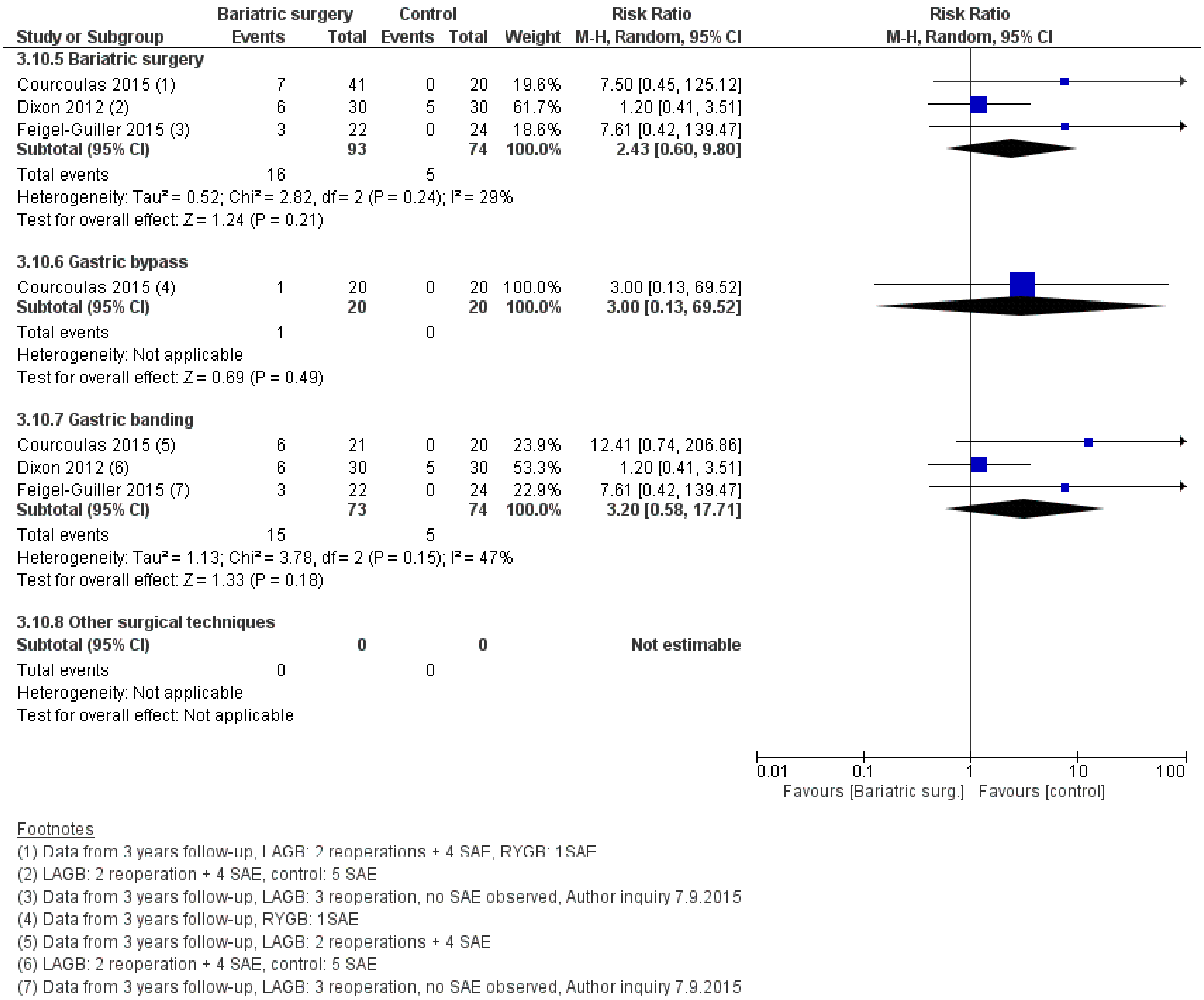


Figure 7 Serious adverse events/reoperations - 2 years results

3.2.4.2.3 Diabetes remission

Bariatric surgery

Diabetes remission was reported by six RCTs at 2 or 3 years. The relative risk to achieve diabetes remission was 10 times higher in the surgical than in the conservative treatment group (RR 10.2, 95% CI 4.4 to 24.0). Heterogeneity might not be important (I^2=21%). The definitions used for diabetes remission are summarized in Appendix 10.

Quality of the evidence according to GRADE: The quality of evidence was low because of very serious risk of bias and serious indirectness. It was downgraded because of indirectness, because results apply only to obese individuals with type 2 diabetes but not to the general population of obese or overweight individuals.

Gastric bypass

Four studies compared gastric bypass with conservative treatment and reported diabetes remission (Figure 8). The relative risk to achieve diabetes remission was 31 times higher in the gastric bypass group than in the conservative treatment group (RR 31.5, 95% CI 7.9 to 126.0). Heterogeneity might not be important (I^2=0%).

Adjustable gastric banding

Three studies compared adjustable gastric banding with conservative treatment and reported diabetes remission (Figure 8). The relative risk to achieve diabetes was 6 times higher (RR 6.2, 95% CI 2.9 to 13.0) in the adjustable gastric banding group than in the conservative treatment group. Heterogeneity might not be important (I^2=0%).

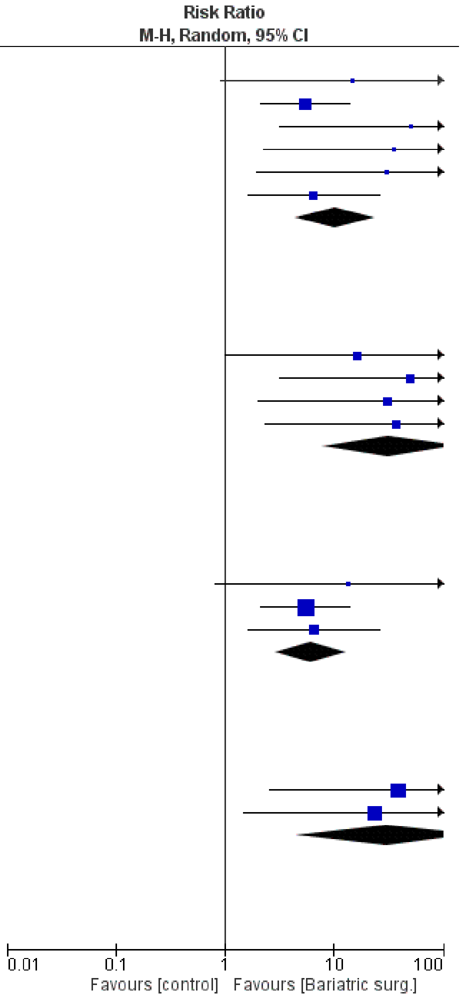
Other surgical interventions

In two studies other surgical interventions, namely biliopancreatic diversion and sleeve gastrectomy, were compared to conservative treatment (Figure 8). The relative risk to achieve diabetes was 30 times higher (RR 30.6, 95% CI 4.3 to 216.0) in the other surgical intervention group than in the conservative treatment group. Heterogeneity might not be important (I^2=0%).

The results of the meta-analysis are presented in Figure 8 and the quality of evidence is presented as summary in Table 7.

Pooled results after 1 year are presented in Appendix 9.

Study or Subgroup	Bariatric surgery Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
3.7.5 Bariatric surgery						
Courcoulas 2015 (1)	14	46	0	23	8.4%	14.81 [0.92, 237.74]
Dixon 2008 (2)	22	30	4	30	40.4%	5.50 [2.15, 14.04]
Ikramuddin 2015 (3)	25	60	0	60	8.4%	51.00 [3.18, 819.00]
Mingrone 2012 (4)	34	40	0	20	8.6%	35.34 [2.28, 548.34]
Schauer 2014 (5)	36	97	0	40	8.5%	30.54 [1.92, 485.84]
Wentworth 2014 (6)	12	23	2	25	25.6%	6.52 [1.63, 26.07]
Subtotal (95% CI)		**296**		**198**	**100.0%**	**10.23 [4.36, 24.02]**
Total events	143		6			
Heterogeneity: Tau² = 0.24; Chi² = 6.32, df = 5 (P = 0.28); I² = 21%						
Test for overall effect: Z = 5.34 (P < 0.00001)						
3.7.6 Gastric bypass						
Courcoulas 2015 (7)	8	24	0	23	24.6%	16.32 [1.00, 267.47]
Ikramuddin 2015 (8)	25	60	0	60	25.0%	51.00 [3.18, 819.00]
Mingrone 2012 (9)	15	20	0	20	25.4%	31.00 [1.98, 485.13]
Schauer 2014 (10)	22	48	0	40	25.0%	37.65 [2.36, 601.83]
Subtotal (95% CI)		**152**		**143**	**100.0%**	**31.47 [7.86, 125.95]**
Total events	70		0			
Heterogeneity: Tau² = 0.00; Chi² = 0.36, df = 3 (P = 0.95); I² = 0%						
Test for overall effect: Z = 4.87 (P < 0.00001)						
3.7.7 Gastric banding						
Courcoulas 2015 (11)	6	22	0	23	7.1%	13.57 [0.81, 227.36]
Dixon 2008 (12)	22	30	4	30	63.8%	5.50 [2.15, 14.04]
Wentworth 2014 (13)	12	23	2	25	29.2%	6.52 [1.63, 26.07]
Subtotal (95% CI)		**75**		**78**	**100.0%**	**6.16 [2.91, 13.02]**
Total events	40		6			
Heterogeneity: Tau² = 0.00; Chi² = 0.38, df = 2 (P = 0.83); I² = 0%						
Test for overall effect: Z = 4.76 (P < 0.00001)						
3.7.8 Other surgical techniques						
Mingrone 2012 (14)	19	20	0	20	50.9%	39.00 [2.52, 604.70]
Schauer 2014 (15)	14	49	0	40	49.1%	23.78 [1.46, 386.69]
Subtotal (95% CI)		**69**		**60**	**100.0%**	**30.58 [4.33, 216.02]**
Total events	33		0			
Heterogeneity: Tau² = 0.00; Chi² = 0.06, df = 1 (P = 0.80); I² = 0%						
Test for overall effect: Z = 3.43 (P = 0.0006)						



Footnotes
(1) Data from 3 years follow-up, HBA1c ≤ 6.5% without diabetes medication
(2) fasting plasma glucose ≤ 7mmol/L (126 mg/dL) and HbA1c ≤ 6.2% without medication
(3) HbA1c ≤ 6.5% at every sampling from months 12 to 24 with no medicines for hyperglycaemia
(4) fasting plasma glucose ≤ 5.6mmol/L and HbA1c ≤ 6.5% for one year without medication;
(5) Data from 3 years follow-up, HBA1c ≤ 6.5% without diabetes medication
(6) fasting plasma glucose < 7mmol/L and < 11.1mmol/L 2h after oral glucose at least 2 days after stopping glucose lowering drugs
(7) Data from 3 years follow-up, HBA1c ≤ 6.5% without diabetes medication
(8) HbA1c ≤ 6.5% at every sampling from months 12 to 24 with no medicines for hyperglycaemia
(9) fasting plasma glucose ≤ 5.6mmol/L and HbA1c ≤ 6.5% for one year without medication;
(10) Data from 3 years follow-up, HBA1c ≤ 6.5% without diabetes medication
(11) Data from 3 years follow-up, HBA1c ≤ 6.5% without diabetes medication
(12) fasting plasma glucose ≤ 7mmol/L (126 mg/dL) and HbA1c ≤ 6.2% without medication
(13) fasting plasma glucose < 7mmol/L and < 11.1mmol/L 2h after oral glucose at least 2 days after stopping glucose lowering drugs
(14) fasting plasma glucose ≤ 5.6mmol/L and HbA1c ≤ 6.5% for one year without medication;
(15) Data from 3 years follow-up, HBA1c ≤ 6.5% without diabetes medication

Figure 8 Diabetes remission – 2 years results

3.2.4.2.4 Hypertension

Hypertension was reported by two RCTs at 2 years. The relative risk was significantly lower in the surgical than in the conservative treatment group (RR 0.68, 95% CI 0.49 to 0.95). Heterogeneity might not be important (I^2=0%).

Quality of the evidence according to GRADE: The quality of evidence was very low because of serious risk of bias, serious indirectness and serious imprecision.

Gastric bypass

One RCT compared gastric bypass with conservative treatment and reported hypertension at 2 years follow-up (Figure 9). The relative risk was significantly lower after gastric bypass (RR 0.68, 95% CI 0.47 to 0.97).

Adjustable gastric banding

One RCT compared adjustable gastric banding with conservative treatment and reported hypertension at 2 years follow-up (Figure 9). The risk to suffer hypertension was lower after gastric bypass (RR 0.72, 95% CI 0.31 to 1.72).

Other surgical interventions

No RCT comparing other surgical interventions with conservative treatment reported hypertension at 2 years follow-up.

The results of the meta-analysis are presented in Figure 9 and the quality of evidence is summarized in Table 7.

Pooled results after 1 year are presented in Appendix 9.

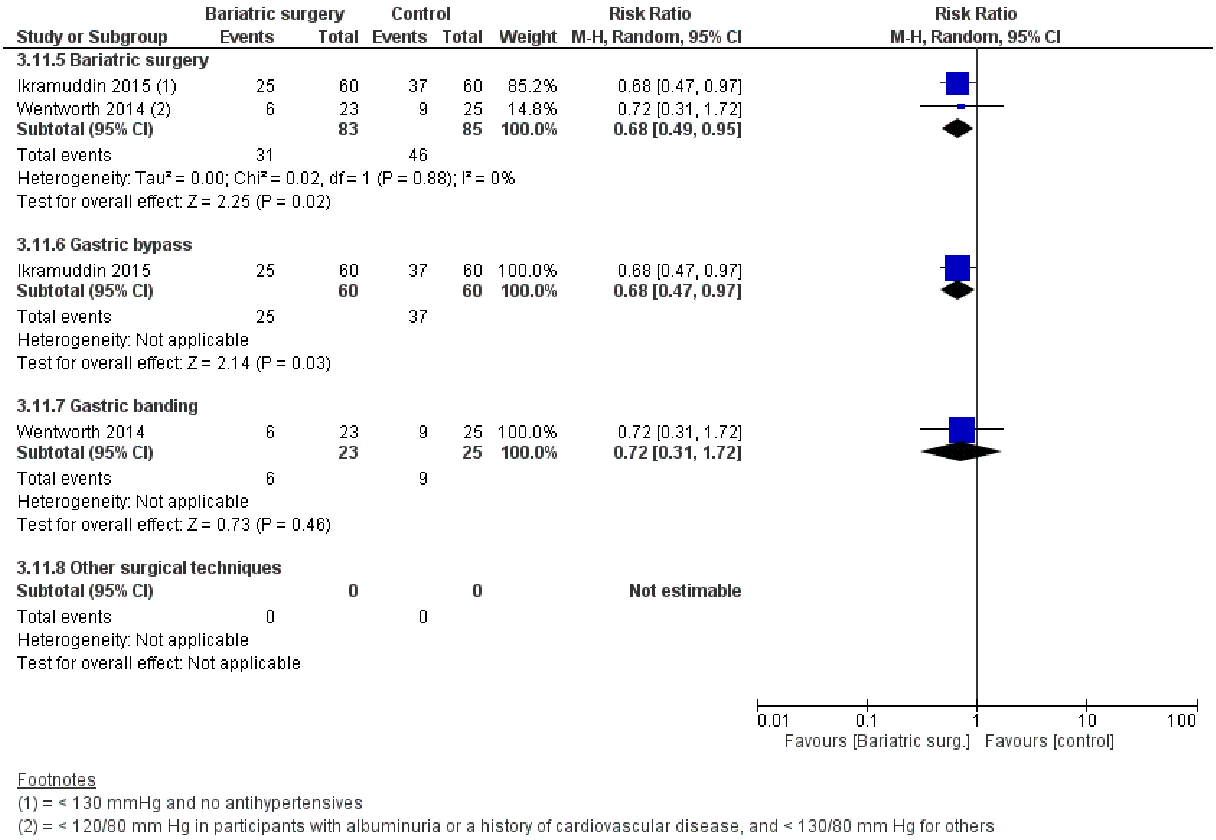


Figure 9 Hypertension – 2 years results

3.2.4.2.5 Dyslipidaemia

One RCT reported on dyslipidaemia (defined as "patient receives medicine against dyslipidaemia") at 2 years. Bariatric surgery reduced the risk for dyslipidaemia (RR 0.63, 95% CI 0.46 to 0.87).

Quality of the evidence according to GRADE: The quality of evidence was low, because of risk of bias and indirectness.

Pooled results after 1 year are presented in Appendix 9.

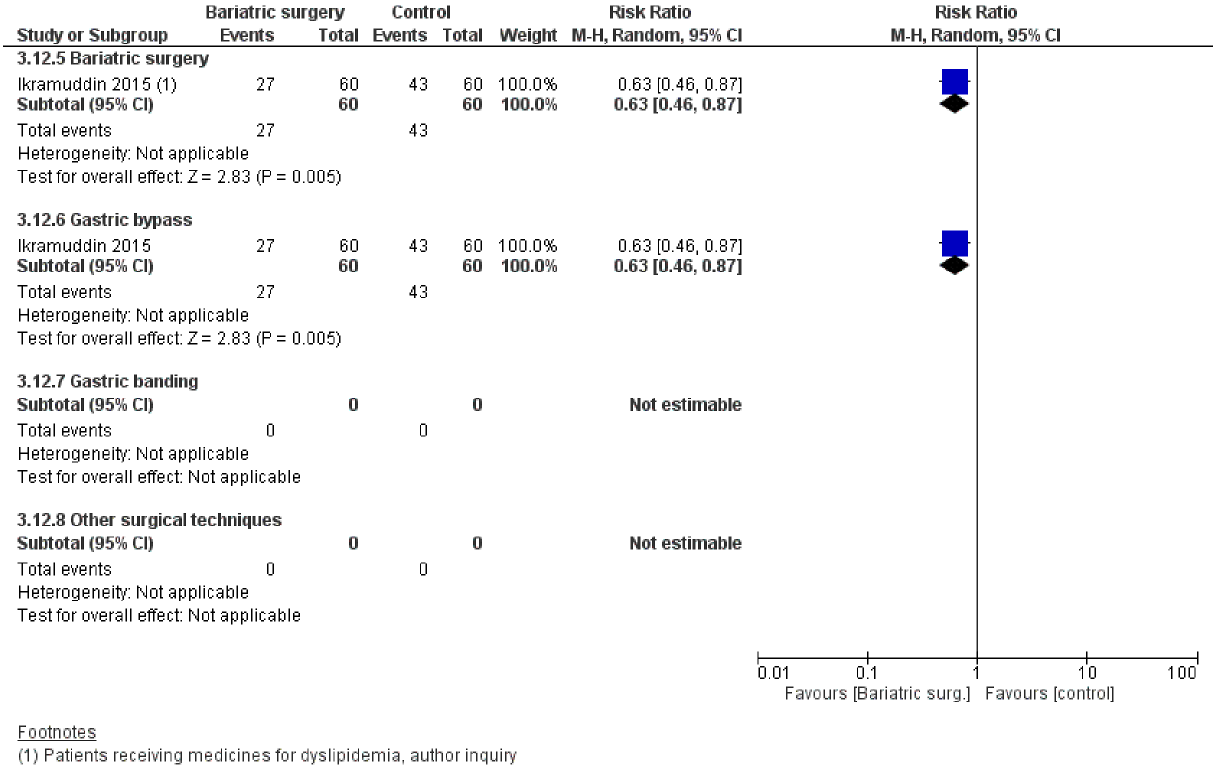


Figure 10 Dyslipidaemia - 2 years results

3.2.4.2.6 Sleep apnoea

Sleep apnoea was reported by one RCT at 3 years. This RCT recruited patients with sleep apnoea and observed in each, the surgical and conservative treatment group, one sleep apnoea remission (<5 events/h). There was no significant difference between the groups; the relative risk was 1.0 (95% CI 0.88 to 1.13).

Quality of the evidence according to GRADE: The quality of evidence was very low because of very serious risk of bias and serious imprecision (very few patients included for the meta-analysis) and serious indirectness (results are limited to obese individuals with sleep apnoea but not to the general population of obese or overweight individuals).

Gastric banding

Results are the same as above, because the one RCT reporting on sleep apnoea used gastric banding (Figure 11).

The results of the meta-analysis are presented in Figure 11 and the quality of evidence is summarized in Table 7.

Pooled results after 1 year are presented in Appendix 9.

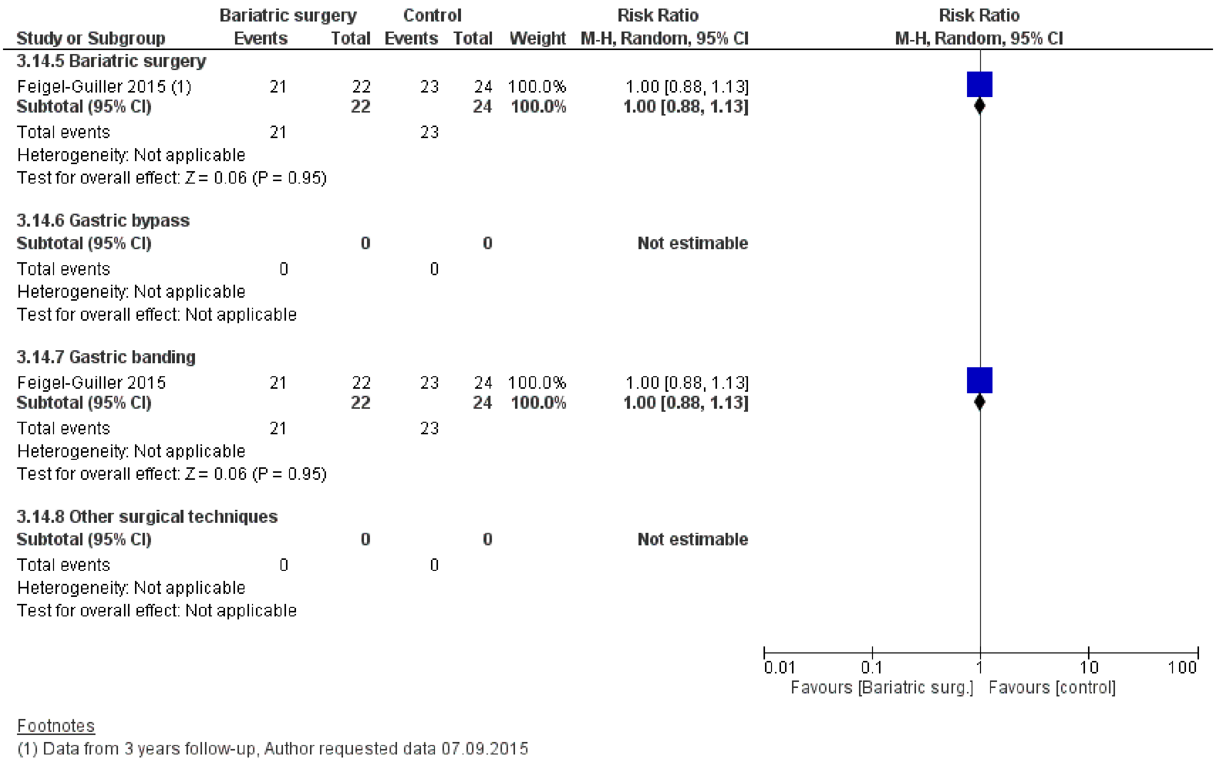


Figure 11 Sleep apnoea – 2 years results

3.2.4.2.7 Cancer

Cancer was reported by one RCT at 2 years and by one RCT at 3 years. There was no significant difference between the surgical and conservative treatment group (RR 0.69, 95% CI 0.16 to 2.95). Heterogeneity might not be important (I^2=0%).

Quality of the evidence according to GRADE: The quality of evidence was very low because of serious risk of bias, serious indirectness and serious imprecision.

Gastric bypass

Two RCTs compared gastric bypass with conservative treatment and reported cancer at 2 years and at 3 years follow-up (Figure 12). There was no significant difference between the gastric bypass and conservative treatment group (RR 0.63, 95% CI 0.12 to 3.18). Heterogeneity might not be important (I^2=0%).

Other surgical interventions

One RCT compared sleeve gastrectomy with conservative treatment and reported cancer at 3 years follow-up (Figure 12). There was no significant difference between the sleeve gastrectomy and conservative treatment group (RR 1.00, 95% CI 0.15 to 6.82). Heterogeneity could not be estimated.

The results of the meta-analysis are presented in Figure 12 and the quality of evidence is summarised in Table 7.

Pooled results after 1 year are presented in Appendix 9.

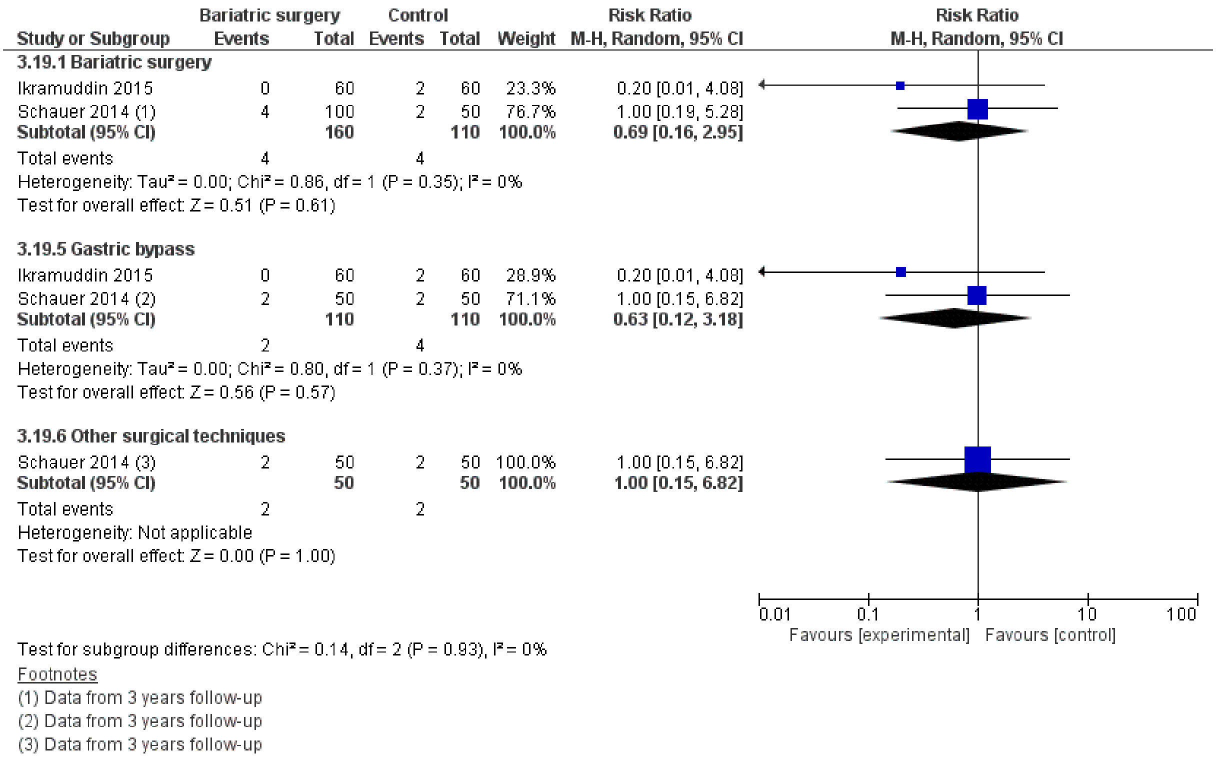


Figure 12 Cancer – 2 years results

3.2.4.2.8 Revision rates

Revision rates were not reported in any RCT.

3.2.4.3 **Additional outcomes**

3.2.4.3.1 Adverse events

An overview of all adverse events reported by each study is presented in Appendix 11.

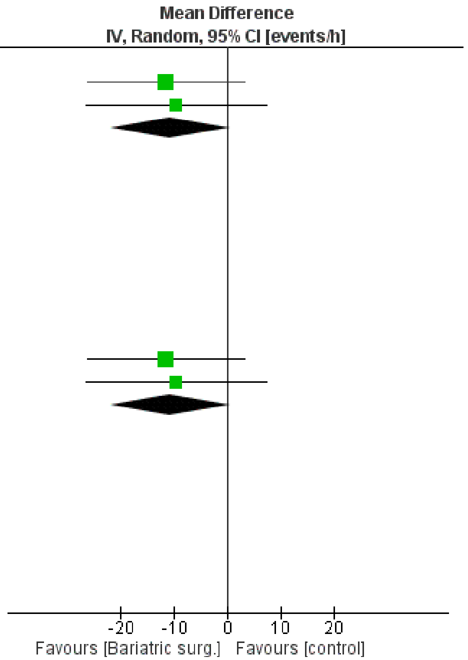
3.2.4.3.2 Apnoea-hypopnea index

Change from baseline in apnoea-hypopnea index (AHI) was reported by two RCTs. One RCT reported change in AHI at 2 years and one RCT at 3 years. There was no significant difference between bariatric surgery and conservative treatment. The mean difference was -10.7 events/h (95% CI -21.9 to 0.4 events/h). Heterogeneity might not be important (I^2=0%).

The results of the meta-analysis are presented in Figure 13 and the quality of evidence is summarized in Table 7.

Pooled results after 1 year are presented in Appendix 9.

	Bariatric surgery			Control				Mean Difference
Study or Subgroup	**Mean [events/h]**	**SD [events/h]**	**Total**	**Mean [events/h]**	**SD [events/h]**	**Total**	**Weight**	**IV, Random, 95% CI [events/h]**
3.12.1 Bariatric surgery								
Dixon 2012	-25.5	30.1	30	-14	28.5	30	56.6%	-11.50 [-26.33, 3.33]
Feigel-Guiller 2015 (1)	-14.2	30.3	22	-4.5	28.1	24	43.4%	-9.70 [-26.63, 7.23]
Subtotal (95% CI)			**52**			**54**	**100.0%**	**-10.72 [-21.88, 0.44]**
Heterogeneity: Tau² = 0.00; Chi² = 0.02, df = 1 (P = 0.88); I² = 0%								
Test for overall effect: Z = 1.88 (P = 0.06)								
3.12.2 Gastric bypass								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
3.12.3 Adjustable gastric banding								
Dixon 2012	-25.5	30.1	30	-14	28.5	30	56.6%	-11.50 [-26.33, 3.33]
Feigel-Guiller 2015 (2)	-14.2	30.3	22	-4.5	28.1	24	43.4%	-9.70 [-26.63, 7.23]
Subtotal (95% CI)			**52**			**54**	**100.0%**	**-10.72 [-21.88, 0.44]**
Heterogeneity: Tau² = 0.00; Chi² = 0.02, df = 1 (P = 0.88); I² = 0%								
Test for overall effect: Z = 1.88 (P = 0.06)								
3.12.4 Other surgical techniques								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								



Footnotes
(1) Data from 3 years follow-up
(2) Data from 3 years follow-up

Figure 13 Apnoea-hypopnea index – 2 years results

3.2.5 Quality of evidence summary

Table 7 Summary of findings (GRADE)

Bariatric surgery compared to conservative treatment for overweight and obesity

Patient or population: overweight and obesity
Settings:
Intervention: Bariatric surgery
Comparison: Conservative treatment

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Conservative treatment	**Bariatric surgery**				
body weight (%-change from baseline)		The mean body weight (%-change from baseline) in the intervention groups was **17.9% lower** (21.4 to 14.5 lower)		617 (8 studies)	⊕⊕⊕⊖ **moderate**[1,2,3]	
Quality of life, physical health SF-36 questionnaire		The mean quality of life, physical health in the intervention groups was **9.4 scores higher** (5.7 to 13 higher)		108 (2 studies)	⊕⊕⊖⊖ **low**[4,5]	
Quality of life, mental health SF-36 questionnaire		The mean quality of life, mental health in the intervention groups was **0.1 scores lower** (3.9 lower to 3.8 higher)		108 (2 studies)	⊕⊖⊖⊖ **very low**[4,6,7]	
HbA1c		The mean HbA1c in the intervention groups was		549 (7 studies)	⊕⊕⊖⊖ **low**[8,9,10,11]	

		1.40%-points lower (1.92 to 0.25 lower)				
Stroke			**RR 1.3** (0.05 to 31.8)	142 (1 study)	⊕⊖⊖⊖ **very low**[10,12,13]	Only one RCT reported this outcome.
Myocardial infarction	See comment	See comment	Not estimable	0 (0)	See comment	No RCT reported this outcome.
All-cause mortality	**6 per 1000**	**2 per 1000** (0 to 48)	**RR 0.33** (0.01 to 8.02)	417 (5 studies)	⊕⊖⊖⊖ **very low**[13,14]	
Serious adverse events including reoperations	**68 per 1000**	**162 per 1000** (41 to 662)	**RR 2.4** (0.6 to 9.8)	167 (3 studies)	⊕⊖⊖⊖ **very low**[13,15,16]	
Diabetes remission	**30 per 1000**	**309 per 1000** (133 to 727)	**RR 10.2** (4.4 to 24)	494 (6 studies)	⊕⊕⊖⊖ **low**[10,17]	
Hypertension	**541 per 1000**	**368 per 1000** (265 to 514)	**RR 0.68** (0.49 to 0.95)	168 (2 studies)	⊕⊖⊖⊖ **very low**[10,16,18]	
Dyslipidaemia	**1467 per 1000**	**924 per 1000** (675 to 1000)	**RR 0.63** (0.46 to 0.87)	120 (1 study)	⊕⊕⊖⊖ **low**[10,19]	Only one RCT reported this outcome.
Sleep apnoea	**42 per 1000**	**42 per 1000** (37 to 47)	**RR 1** (0.88 to 1.13)	46 (1 study)	⊕⊖⊖⊖ **very low**[13,16,20,21]	Only one RCT reported this outcome.
Cancer - not reported	**18 per 1000**	**13 per 1000** (3 to 54)	**RR 0.69** (0.16 to 2.95)	270 (2 studies)	⊕⊖⊖⊖ **very low**[10,13,19]	
Revision rates - not reported	See comment	See comment	Not estimable	-	See comment	No RCT reported this outcome.

* The **illustrative comparative risks** (and its 95% confidence interval) are based on the mean baseline risks from the studies included in the meta-analyses. **CI:** Confidence interval; **RR:** Risk ratio; **RCT:** randomized controlled trial

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

[1] performance bias was high or unclear in all RCTs; attrition bias was high in 5 RCTs.

[2] It was not downgraded for inconsistency because heterogeneity measured by I^2 might have been misleadingly high. Conceptually, I^2 represents the ratio of the variability between studies to the variability within studies and is independent from the number of included studies. When the variability within studies is small—represented by narrow confidence intervals—the variability between studies can also be small and the I^2 nevertheless high and inflated. By sensitivity analyses Mingrone 2012 was identified as source of detected high heterogeneity, but could not identify any specific characteristic (e.g. the surgical technique) that made it different from all other trials, except for the fact that the effect on %body weight loss was much higher than in all other trials.

[3] It was not rated down for imprecision because current guidelines define 5-10% weight reductions as clinically important.

[4] risk for selection bias (allocation concealment) was unclear in 1 RCT; performance bias was high in 1 and unclear in 1 RCT; detection bias was high in 2 RCTs.

[5] not downgraded because the lower 95% CI was greater 5 . 3-5 or more was considered as minimally clinically important difference.

[6] Although the studies did not report overall quality of life scores, it was not downgraded for indirectness because specific summary components (physical health or mental health) of the same quality of life tool were reported and analysed

[7] 95% CI of the effect included 3 and -3 and thus a possible clinically important difference, i.e. 95% CI includes appreciable benefit or harm. Scores between 3-5 or more were considered as minimally clinically important difference.

[8] attrition bias was high in 3 RCTs;

[9] Inconsistency was not downgraded because heterogeneity is explained by Dixon 2012 (only study with a non-diabetic population) and Wentworth 2014 (only study that exclusively included an overweight population, BMI 25-29.9 kg/m^2).

[10] effect estimate limited to an obese population with type 2 diabetes, but not to the general obese or overweight population (as defined in the PICO question of the report).

[11] No clinically important difference was considered to judge imprecision.

[12] risk for selection bias (allocation concealment) was unclear; the risk for performance bias was high; the risk for attrition bias was high

[13] very low event rate

[14] attrition bias was high in 2 RCTs;

[15] performance bias was unclear in all RCTs; detection bias was unclear or high in all RCTs; attrition bias was high in 2 RCTs;

[16] number of participants in meta-analysis is low
[17] attrition bias was high in 3 RCTs;
[18] risk for selection bias (allocation concealment) was unclear in 1 RCT; performance bias was high in all 2 RCTs;
[19] risk for selection bias (allocation concealment) was unclear; performance bias was high;
[20] risk for selection bias (allocation concealment) was unclear; attrition bias was high; performance bias was unclear; detection bias was unclear.
[21] effect estimate limited to an obese population with obstructive sleep apnoea, but not to the general obese or overweight population (as defined in the PICO question of the report).

3.2.6 Additional stratified results

Subgroup analyses were not done, i.e. subgroups were not compared statistically, because the number of studies available in each subgroup was too small and 3–arm studies were included. However, results were stratified by certain study characteristics. Results stratified by surgical techniques are presented in the previous section for each outcome. Results stratified by BMI are presented in the following. Methodological issues, e.g. adequate vs. inadequate concealment of treatment allocation, were not thought of as source of heterogeneity and consequently no stratified or sensitivity analyses were done because of methodological issues; another reason was the small number of studies

3.2.6.1 **Individuals with a BMI ≥35 kg/m² vs. individuals with a BMI of 25-34.9 kg/m²**

Four RCTs included individuals with BMI ≥35 kg/m²and two RCTs included individuals with a BMI of 25 to 34.9 kg/m² (Table 5). Two of the four RCTs including individuals with BMI ≥35 kg/m² compared gastric bypass with conservative treatment and the two others compared adjustable gastric banding with conservative treatment. The two RCTs with BMI of 25 to 34.9 kg/m² compared both adjustable gastric banding with conservative treatment. Comparison was possible for four outcomes, namely percent change of body weight, quality of life, HbA1c and diabetes remission.

3.2.6.1.1 Body weight

Three RCTs included individuals with BMI ≥35 kg/m² and one study included individuals with BMI of 25 to 34.9 kg/m² reported percent change of body weight. Results were similar. The mean difference in percent change of body weight was -19.5% (95% CI -30.4 to -8.6%) in individuals with BMI ≥35 kg/m² and -16.1%-points (95% CI -19.3 to -12.9%-points) in individuals with BMI of 25 to 34.9 kg/m².

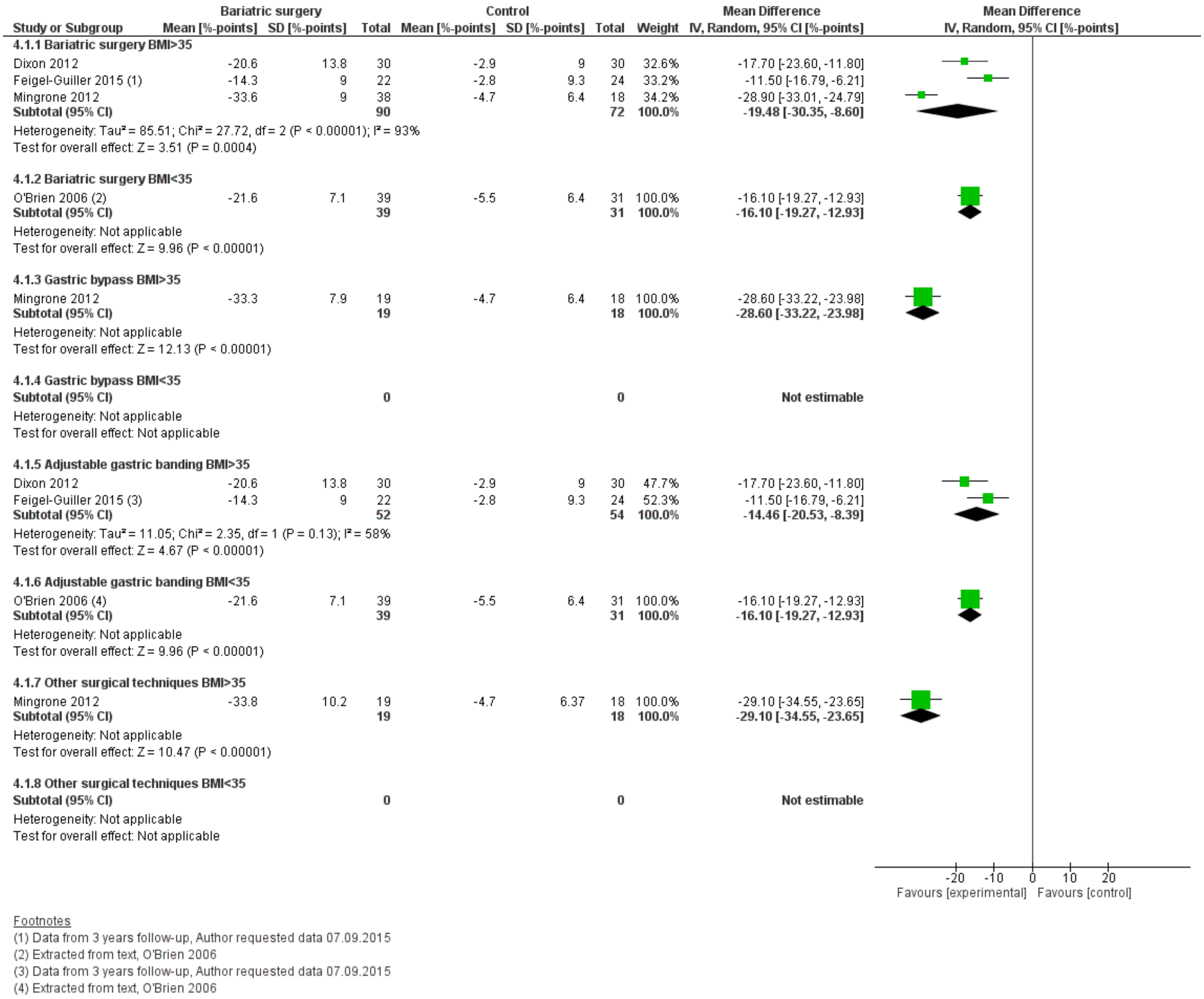


Figure 14 Body weight, sub-group analysis: Individuals with a BMI ≥35 kg/m² vs. individuals with a BMI 25-34.9 kg/m²

3.2.6.1.2 Quality of life

One study included individuals with BMI ≥35 kg/m² and one study included individuals with BMI of 25 to 34.9 kg/m² reported quality of life summary components. Results were similar. The mean differences in physical health were 9.2 (95% CI 2.2 to 16.2 scores) in individuals with BMI ≥35 kg/m² and 9.4 (95% CI 5.1 to 13.7 scores) in individuals with BMI of 25 to 34.9 kg/m² (Figure 15).

Results were also similar for mental health. The mean differences were -0.3 (95% CI -4.7 to 4.1 scores) in individuals with BMI ≥35 kg/m² and 0.7 (95% CI -6.9 to 8.2 scores) in individuals with BMI of 25 to 34.9 kg/m² (Figure 15).

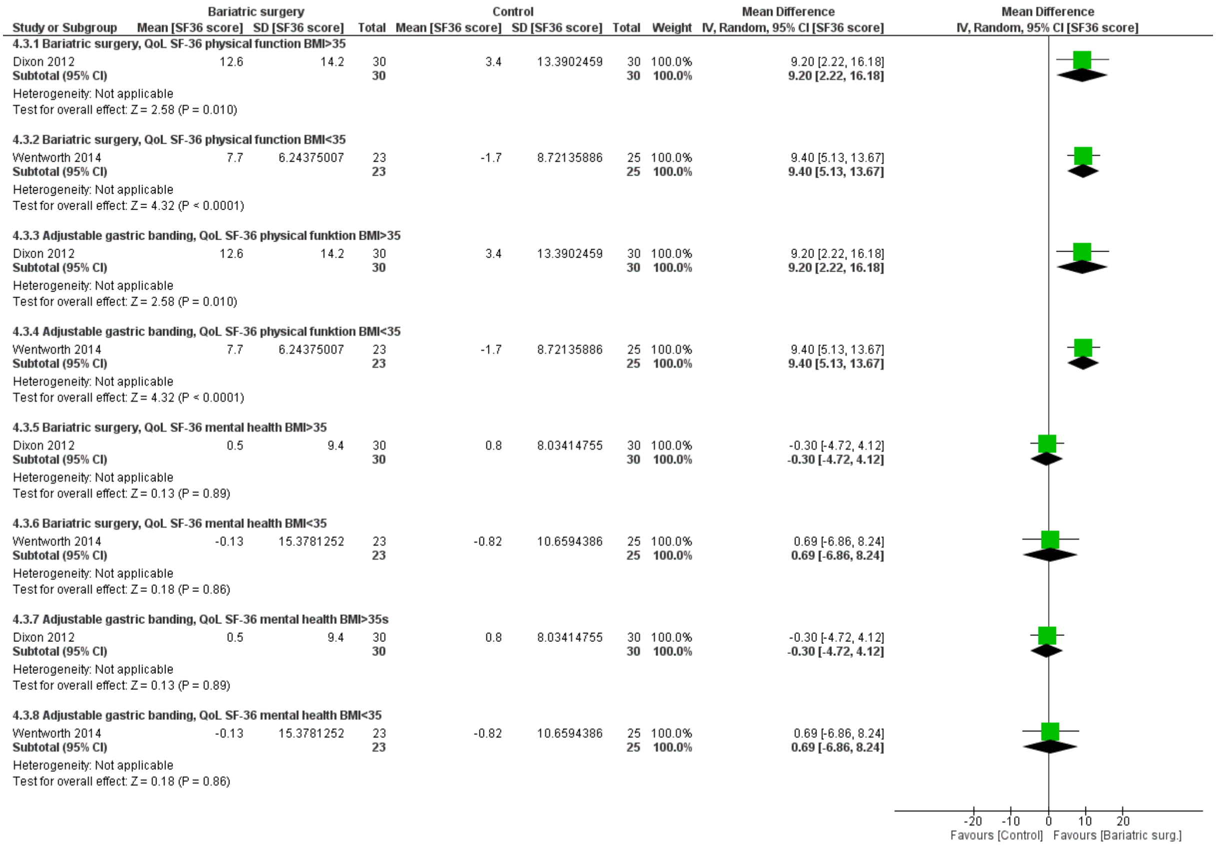
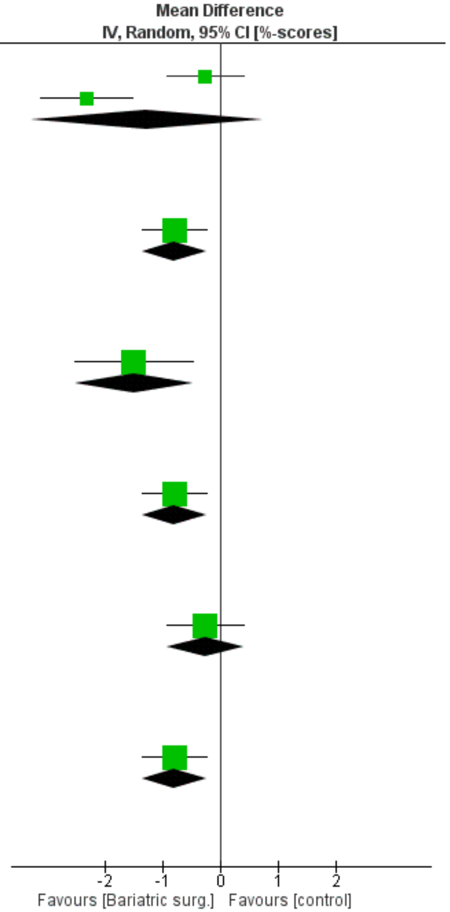


Figure 15 Quality of life, sub-group analysis: Individuals with a BMI ≥35 kg/m² vs. individuals with a BMI 25-34.9 kg/m²

3.2.6.1.3 HbA1c

Two studies included individuals with BMI ≥35 kg/m² and one study included individuals with BMI of 25 to 34.9 kg/m² reported HbA1c as change from baseline. The mean differences for change in HbA1c were -1.28% (95% CI -3.30 to 0.74%) in individuals with BMI ≥35 kg/m² and -0.80% (95% CI -1.35 to -0.25 scores) in individuals with BMI of 25 to 34.9 kg/m² (Figure 16).

Study or Subgroup	Bariatric surgery Mean [%-scores]	Bariatric surgery SD [%-scores]	Bariatric surgery Total	Control Mean [%-scores]	Control SD [%-scores]	Control Total	Weight	Mean Difference IV, Random, 95% CI [%-scores]
4.5.1 Bariatric surgery BMI>35								
Dixon 2012	-0.27	1.63	30	-0.01	0.95	30	50.6%	-0.26 [-0.94, 0.42]
Mingrone 2012 (1)	-3.14	1.98	38	-0.82	1.07	18	49.4%	-2.32 [-3.12, -1.52]
Subtotal (95% CI)			**68**			**48**	**100.0%**	**-1.28 [-3.30, 0.74]**
Heterogeneity: Tau² = 1.98; Chi² = 14.87, df = 1 (P = 0.0001); I² = 93%								
Test for overall effect: Z = 1.24 (P = 0.21)								
4.5.2 Bariatric surgery BMI<35								
Wentworth 2014	-0.8	0.69	23	0	1.21	25	100.0%	-0.80 [-1.35, -0.25]
Subtotal (95% CI)			**23**			**25**	**100.0%**	**-0.80 [-1.35, -0.25]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.84 (P = 0.004)								
4.5.3 Bariatric surgery, only T2DM patients BMI>35								
Mingrone 2012 (2)	-2.32	2.02	19	-0.82	1.07	18	100.0%	-1.50 [-2.53, -0.47]
Subtotal (95% CI)			**19**			**18**	**100.0%**	**-1.50 [-2.53, -0.47]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.84 (P = 0.004)								
4.5.4 Bariatric surgery, only T2DM patients BMI<35								
Wentworth 2014	-0.8	0.69	23	0	1.21	25	100.0%	-0.80 [-1.35, -0.25]
Subtotal (95% CI)			**23**			**25**	**100.0%**	**-0.80 [-1.35, -0.25]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.84 (P = 0.004)								
4.5.5 Adjustable gastric banding BMI>35								
Dixon 2012	-0.27	1.63	30	-0.01	0.95	30	100.0%	-0.26 [-0.94, 0.42]
Subtotal (95% CI)			**30**			**30**	**100.0%**	**-0.26 [-0.94, 0.42]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.75 (P = 0.45)								
4.5.6 Adjustable gastric banding BMI<35								
Wentworth 2014	-0.8	0.69	23	0	1.21	25	100.0%	-0.80 [-1.35, -0.25]
Subtotal (95% CI)			**23**			**25**	**100.0%**	**-0.80 [-1.35, -0.25]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.84 (P = 0.004)								

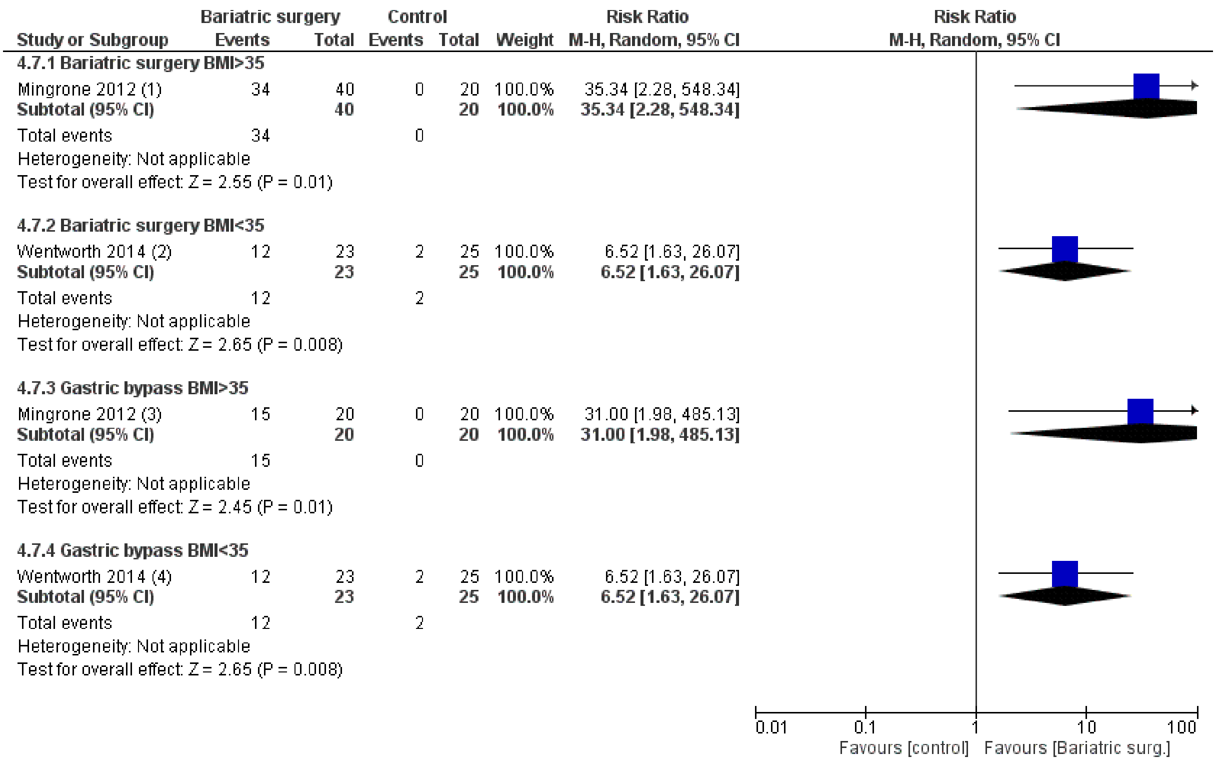


Footnotes
(1) Calculations based on data received from author inquiry 14/9/2013
(2) Calculations based on data received from author inquiry 14/9/2013

Figure 16 HbA1c, sub-group analysis: Individuals with a BMI ≥35 kg/m² vs. individuals with a BMI 25-34.9 kg/m²

3.2.6.1.4 Diabetes remission

One study included individuals with BMI ≥35 kg/m² and one study included individuals with BMI of 25 to 34.9 kg/m² reported diabetes remission. The relative risk to achieve diabetes remission was 35.3 (95% CI 2.3 to 548.3) in individuals with BMI ≥35 kg/m² and 6.5 (95% CI 1.6 to 26.1) in individuals with BMI of 25 to 34.9 kg/m² (Figure 17).



Footnotes
(1) fasting plasma glucose ≤ 5.6mmol/L and HbA1c ≤ 6.5% for one year without medication;
(2) fasting plasma glucose < 7mmol/L and < 11.1mmol/L 2h after oral glucose at least 2 days after stopping glucose lowering drugs
(3) fasting plasma glucose ≤ 5.6mmol/L and HbA1c ≤ 6.5% for one year without medication;
(4) fasting plasma glucose < 7mmol/L and < 11.1mmol/L 2h after oral glucose at least 2 days after stopping glucose lowering drugs

Figure 17 Diabetes remission, sub-group analysis: Individuals with a BMI ≥35 kg/m² vs. individuals with a BMI 25-34.9 kg/m²

3.2.6.2 Perioperative (≤30 days after surgery) vs. longer term (>30 days after surgery) serious adverse events including reoperations

One study (**Courcoulas 2015**) differentiated between perioperative and longer term serious adverse events/reoperations. In this study no serious adverse event including reoperation occurred within ≤30 days after surgery.

3.3 Discussion

3.3.1 General points

The scope of this report follows quite closely the question as it was initially submitted to the SMB. The SMB was asked to evaluate bariatric surgery and gastric bypass in particular regarding effectiveness and costs, taking into account and including possible complications. Conservative treatment was suggested as the comparator. The scope of this report and the hierarchy of the outcomes assessed were elaborated in collaboration with clinical experts.

According to the scope, it was sufficient if the primary RCTs included overweight or obese adults. Overweight and obese people have a markedly higher risk for a number of different comorbidities[54] and accordingly a high prevalence of comorbidities was expected in these populations. RCTs were eligible for the inclusion in the present report whether or not they defined certain comorbidities as additional inclusion criteria for their patients. Most of the RCTs were done in obese individuals except for one in overweight individuals (**Wentworth 2014**). It was possible to stratify results by RCTs done in individuals with a BMI ≥35 kg/m^2 and RCTs done in individuals with a BMI 25-34.9 kg/m^2. Results are discussed later on.

Only the RCTs reporting results after 2 or 3 years were included in the main meta-analyses, grading the evidence (GRADE) and are matter of the discussion.

In most of the included RCTs patients had to have at least one comorbidity in order to be included. Of the ten RCTs reporting results after 2 or 3 years of follow-up, only **Reis 2010** did not specify the presence of comorbidity as an inclusion criterion. The majority of RCTs included overweight or obese patients with type 2 diabetes, only (2 to 3 year data: **Courcoulas 2015, Dixon 2008, Ikramuddin 2015, Mingrone 2012, Schauer 2014, Wentworth 2014;** ≤1 year data: **Ding 2015, Halperin 2014, Liang 2013**). Apart from **Courcoulas 2015** and **Schauer 2014** the authors also pre-specified how long the patients needed to be affected with diabetes. Two RCTs with 2 year data only included patients with obstructive sleep apnoea (**Dixon 2012, Feigel-Guiller 2015**). The RCT by **O'Brien 2006** requested at least one obesity-related comorbidity, such as diabetes, obstructive sleep apnoea, hypertension, gastro-oesophageal reflux etc., as inclusion criterion. One RCT with 2 year data (**Reis 2010**) and 3 RCTs with 1 year data (**Heindorff 1997, Mingrone 2002, Parikh 2014**) did not define comorbidity as inclusion criterion. The prevalence of comorbidities at baseline has not been described in these latter RCTs. Differences in the prevalence of comorbidities at baseline (differences in baseline risk) can affect absolute effect estimates. Furthermore, the presence of one comorbidity can also affect the prevalence or the risk to develop other comorbidities. For instance diabetes can also affect the prevalence of hypertension or myocardial infarction and thereby the expected absolute effects. The variation of effects for an outcome could also depend on the presence or absence of comorbidity if they are mediated via different biological mechanisms. In this case not only the absolute effects would vary between populations, but – due to effect modification – also the relative effects. Thus, the absolute effects in a population without or with a lower prevalence of diabetes could not simply be calculated based on the relative effects in a population of obese diabetics.

Accordingly, some of the pooled absolute effects in the GRADE tables cannot be directly transferred to the Swiss overweight and obese population, where the prevalence of comorbidities is likely to be lower. Hence, the quality of evidence was downgraded due to indirectness. As none of the included trials investigated whether bariatric surgery can reduce the incidence or delay the occurrence of diabetes or other comorbidities, it was not possible to judge the expected effects in a population without comorbidities.

When applying the evidence in clinical practice it needs to be kept in mind that among different ethnicities comorbidities associated with overweight and obesity may occur at different BMIs. In Asians the typical comorbidities associated with overweight and obesity occur at lower BMIs than in Europeans. Accordingly, the WHO has developed a separate classification with different cut-offs for overweight and obesity in Asians. [55] In one of the included studies (**Ikramuddin 2015**) nearly 30% of the population were of Asian origin.

Currently gastric bypass is the most frequently performed surgical intervention in Switzerland and 5 of the included studies provided 2 year data for this intervention (**Courcoulas 2015, Ikramuddin 2015, Mingrone 2012, Reis 2010, Schauer 2014**). The original HTA report for the BAG[7] from 2006, which assessed a variety of conservative and surgical treatment options for obesity, only included the RCT by **Mingrone 2012** for the comparison of bariatric surgery vs. conservative treatment. Based on this, the BAG came to the conclusion that there is evidence from a trial of low methodological quality suggesting that gastric bypass results in a “substantial weight loss in patients with morbid obesity” compared to conservative treatment.[7]

The types of conservative treatment were slightly different among the included studies. The type of conservative treatment may have an influence on the effect size, but it was not possible to stratify results based on differences in the conservative treatment.

As the decision for bariatric surgery or conservative treatment depends to a large extent on patient preferences, the rate of patients refusing the intervention they were randomised to and the adherence rate (individuals undergoing the planned intervention and returning at follow-up) were calculated. More patients in the conservative treatment arm refused the intervention, and less patients in the conservative treatment arm adhered to the intervention. Whereas the first one hints at early disappointment with the treatment received, the latter one might express disappointment with the treatment success later on and could mean that conservative treatment requires stronger commitment by the individual than bariatric surgery.

Furthermore, adherence rate here might not be equivalent to compliance because patients may have received the conservative intervention in the form of counselling, advice, check-ups and have attended the follow-up visit but without following the advice regarding diet, physical activity and life style.

The main interest lies in assessing long-term outcomes. In in the present analysis the results for 6 months to 1 year and results for 2-3 years were pooled separately and focussed on the latter time point, expecting it to be more representative of the long term effects of bariatric surgery. Weight loss is considered to be a critical outcome, because maintaining a reduced body weight in the longer term is likely to translate into beneficial health outcomes. For two of the RCTs included in the present meta-analysis follow-up data beyond 2 years were published. Data were not included in the present report because results were published beyond the date of the literature search (no update search was planned) in case of the RCT by **Mingrone 2012**. The

RCT by **Feigel-Guiller 2015** ended after 3 years but the authors decided to collect additional data after 10 years, Mingrone 2012 reported modest weight re-gain between year 2 and 5 after surgical treatment, but not after conservative treatment. Compared to year 2 less surgically treated individuals were still in diabetes remission at year 5 (approx. 50% vs. 75%). In the conservative treatment group still nobody was in diabetes remission. Individuals in the RCT from **Feigel-Guiller 2015** seemed to maintain their reduced body weight 10 years after treatment in both the surgical and the conservative treatment arm. In both RCTs most weight was lost in the first year.

3.3.2 Evidence by outcome

In the following, results for the critical and important outcomes will be summarized, with a focus on gastric bypass. Though results are presented separately for gastric bypass effect estimates were not compared statistically to other surgical techniques because three RCTs were three-armed and number of remaining RCTs within strata was small.

3.3.2.1 **Critical outcomes**

3.3.2.1.1 Body weight

Body weight decreased more in the bariatric surgery than in the conservative treatment group. Because heterogeneity among the pooled effect size was substantial, a sensitivity analysis was done and identified **Mingrone 2012** as main source of heterogeneity. The individual effect size of this RCT was higher than in all other studies. Excluding Mingrone 2012 from the analysis resulted in a similar effect size and no longer important heterogeneity.

Similarly gastric bypass yielded better results for percent change of body weight than conservative treatment.

The quality of evidence was moderate.

3.3.2.1.2 Quality of life

Because no overall summary measure for quality of life was reported, but the summary measures of the mental health component and physical health component, mental health and physical health were pooled individually.

The pooled results showed that physical health improved significantly more in the bariatric surgery than in conservative treatment groups. The effect was clinically significant.

The quality of evidence for physical health was low.

There was no significant differential effect on mental health between bariatric surgery and conservative treatment and the confidence interval included both clinically important benefit and harm. The quality of evidence for mental health was very low.

3.3.2.1.3 HbA1c (critical) and remission from diabetes (important)

Although HbA1c was classified as a critical and remission from diabetes as an important outcome, they will be discussed together as they are both relevant for the evaluation of the effect of bariatric surgery on diabetes. HbA1c decreased significantly more in the bariatric surgery, as well as in the gastric bypass group than in the conservative treatment group. The relative risk to achieve diabetes remission was statistically higher in the bariatric surgery, as well as in the gastric bypass group, than in conservative treatment arm.

For both outcomes HbA1c and diabetes remission the quality of evidence was low.

The definitions of diabetes remission used by study authors differed. Because of the limited number of studies it was not possible to assess whether effects differed depending on the definition used. Most studies defined diabetes remission as an HbA1c below 6.5% without diabetes medication. The American Diabetes Association defines partial remission as the absence of any medication for diabetes medication in patients with a HbA1c of less than 6.5% and a fasting plasma glucose level of ≤125 mg/dL, complete remission is the absence of any medication for diabetes medication in patients with a HbA1c of less than 5.7% and a fasting plasma glucose level of ≤100 mg/dL.[33]

From the primary RCTs it is not clear how the effect of diabetes remission is maintained over time. **Schauer 2014** reports for example that the percentage of patients with diabetes remission after one year with a HbA1c <6% who did not maintain that level was 80% in the conservative treatment arm, 24% in the gastric bypass arm and 50% in the sleeve gastrectomy arm. **Courcoulas 2015** reports that for the last 2 years of the 3 year follow-up the remission remained stable in only 45% of the patients with Roux-en-Y gastric bypass and in 29% of patients laparoscopic adjustable gastric banding. While the (partial or complete) remission rates declined in the Roux-en-Y gastric bypass group declined from 60% in year one to 45% in year two and 40% in year 3 the effect for laparoscopic adjustable gastric banding remained stable at 29%. In the conservative treatment arm no patient went into remission. Nonetheless, the differences between the three treatment arms were not statistically different. The numbers involved in those assessments are very low (mostly below 10 patients per group). The Swedish Obese Subject trial – a prospective controlled intervention study also reported a 50% relapse after initial remission from diabetes in the surgical arm. However, the investigators did not have any data on this endpoint for the different types of surgery performed (gastric bypass, banding, vertical banded gastroplasty).[56]

Several RCTs investigated and discussed possible predictors for diabetes remission. **Ding 2015** discussed that a reason for the low remission rate after adjustable gastric banding compared **Dixon 2008** and **Courcoulas 2015** might be that they didn't apply inclusion criteria like duration of diabetes and insulin or oral management that may be seen as proxies for disease severity. They also report the inclusion of patients with a lower BMI (30-35 kg/m^2). **Dixon 2008** found in their population with recent onset diabetes (<2 years) that greater percentage of weight loss and lower baseline HbA1c were both independently associated with diabetes remission - most of the variance was explained by weight loss alone. **Wentworth 2014** (patients with less than 5 years duration of diabetes) found the same two predictors in their regression analysis as **Dixon 2008** but also baseline fasting glucose. Yet they point out that in their gastric banding group they also had 3 patients with a weight loss of more than 20% who didn't achieve diabetes remission. In the study by **Schauer 2014** weight loss and a shorter duration of diabetes were mentioned as predictors of having an HbA1c of 6.0% or less after surgery. **Mingrone 2012,** though, found no correlation between normalization of fasting glucose levels and weight loss after gastric bypass and biliopancreatic diversion. The identification of predictors was not the primary aim in these studies and the analyses performed are therefore just exploratory.

3.3.2.1.4 Stroke

One study (**Schauer 2014**) reported on one stroke after gastric bypass (1/99) and none after conservative treatment (0/43). The quality of the evidence was very low. We have very little confidence in the effect estimate of all-cause mortality: The true effect is likely to be substantially different from the estimate of effect (definition according to Balshem et al, JCE 2011).[14] Given the scarcity of events the effect cannot be reliably estimated.

3.3.2.2 **Important outcomes**

3.3.2.2.1 All-cause mortality

One death occurred in one of the five RCTs (**Ikramuddin 2015**) in the conservative treatment group. The quality of evidence was very low. Given the scarcity of events, the effects cannot be reliably estimated.

In addition to long term mortality, perioperative mortality is an important safety outcome. None of the RCTs reported perioperative deaths. Perioperative mortality might be generally lower within the setting of RCTs. According to the evidence statement in clinical guideline by NICE (CG43) learning curves are associated with bariatric surgery and individual procedures.[57] As some of the studies included in this report were run in specialized centres they may have lower rates of perioperative mortality due to the greater experience of the surgeons than non-specialized centres. This may not necessarily affect the transferability of the results to Switzerland as the guideline of the Swiss Society for the Study of morbid Obesity and metabolic Disorders (SMOB) states that surgery should be performed in SMOB approved centres specialized in the treatment of obesity and which should fulfil a series of requirements.[3] Most of the studies included patients with pre-defined co-morbidities but excluded patients with severe co-morbidities like active cardiovascular diseases **(Halperin 2014, Ikramuddin 2015**). It is unclear how comparable the risk profiles are to those of the general population of obese and overweight patients and how they would in the end affect perioperative mortality. Outside such highly controlled settings mortality rates may be higher. For the Swedish Obese Subjects (SOS) study, a prospective study including about 4,000 obese individuals, postoperative mortality was 0.25%.[56] The most frequent surgical intervention in the SOS study was vertical banded gastroplasty. During the same time 0.1% deaths occurred in the control group. On the other hand the SOS study also reports on a significant reduction of long term overall mortality for surgery compared to conservative treatment during a follow-up of 15 years HR=0.71, 95% CI 0.54-0.92).

3.3.2.2.2 Serious adverse events including reoperations

There was no significant difference between the surgical and conservative treatment group for SAEs including or excluding reoperations. The quality of evidence for the endpoint SAE including reoperations was very low.

One serious adverse event including reoperation was reported in the gastric bypass study. The risk to suffer a serious adverse event including reoperation after gastric bypass was not significantly different compared to conservative treatment.

The combined outcome serious adverse events (SAE) and reoperations had two limitations. Firstly, reoperations can only occur in the bariatric surgery group and secondly, several studies reported reoperations without reporting SAE. Therefore, SAE and reoperations were analysed separately in the following sensitivity analysis.

After bariatric surgery 7.6% (17/223) of the individuals were re-operated. In two other prospective studies rates were 14% and 17%.[58 59] In the Swedish Obese Subjects (SOS) study the rate of serious complications leading to reoperations was 3%.[56]

When stratifying by surgical technique, reoperation rate was 2.6% (1/39) after gastric bypass.

The relative risk to suffer a serious adverse event after bariatric surgery was not significantly different compared to conservative treatment.

Reporting of SAE was overall rare and not consistent, i.e., in some instances events classified as SAE in earlier publications were not reported as such in later publications or the designation of incidents was unclear, so that it was not possible to differentiate adverse events and SAEs.

Besides SAE, post hoc data on adverse events was extracted (see Appendix 11). Adverse events were not consistently reported by all RCTs and differed depending on the surgical technique. Therefore the quantification of rates of adverse events was difficult. It seems that after bariatric surgery, hypoglycaemic episodes are a frequent event as well as iron deficiency anaemia, vitamin-deficiencies, dehydration, neuropathy or kidney stones. And it seems that after conservative treatment, hypoglycaemic episodes as well nausea are frequent events. The SOS trial reports that pulmonary adverse events were most frequent, followed by vomiting, wound infection, haemorrhage and anastomotic leak.[56]

The rate for adverse events may be underestimated as the follow-up periods were rather short and it is known that complications can occur after several years.[60-62] According to the systematic review by Puzziferri et al from 2014, the percentage of long-term complications that had to be treated tended to be relatively small though (≤3% after gastric bypass, ≤6% after gastric banding).[62] Similarly to the outcome mortality, it is unclear how the risk profile of the individuals included in the RCTs affects the rates of adverse events.

Due to the rather small number of participants in the included studies, the estimates of SAE-rates are imprecise and the studies probably lack the power to detect differences between the two interventions. Therefore, no conclusions about adverse events in a generalized population, especially in the long-term, can be drawn.

3.3.2.2.3 Hypertension

The risk for hypertension was significantly lower in the surgical than in the conservative treatment group. The quality of evidence was low.

Similarly, after gastric bypass the relative risk was significantly lower.

3.3.2.2.4 Dyslipidaemia

Bariatric surgery reduced the risk for dyslipidaemia compared to conservative treatment. The quality of evidence was low.

3.3.2.2.5 Sleep apnoea and AHI

One RCT recruited patients with sleep apnoea and observed in each, the surgical and conservative treatment group, one sleep apnoea remission (<5 events/h). The quality of evidence was very low.

In addition, the apnoea-hypopnoea index (AHI) was assessed. There was no significant difference between bariatric surgery and conservative treatment.

While **Feigel-Guiller 2015** found no relationship between AHI and reduction in body weight, **Dixon 2012** found an association between weight loss and AHI though they found this effect to be statistically relevant only in the conservative treatment arm. Adherence to therapy i.e. CPAP/ NIV (**c**ontinuous **p**ositive **a**irway **p**ressure/ **n**on-**i**nvasive **v**entilation) was poor in both trials.

3.3.2.2.6 Cancer
One RCT reported patients with cancer after two years and a second RCT reported cancer after three years of follow-up. The occurrence of cancer was not different between the surgical or conservative treatment groups. The quality of evidence was very low.

3.3.2.2.7 Myocardial infarction and revision rates
None of the studies reported on the results after a follow-up of 2 years for the critical outcome myocardial infarction and on the important outcome revision rates.

3.3.2.3 Gastric bypass compared to other surgical techniques

The comparison of different surgical techniques was not part of the scope of this report. Due to the small number of available studies the evaluation was limited to a stratification by surgical technique rather than formal subgroup analyses.

Meta-analyses which included RCTs directly comparing surgical techniques found that gastric bypass resulted in significantly greater body weight loss compared with gastric banding.[63] Colquitt et al. report that gastric bypass was associated with a greater duration of hospitalisation in 2 RCTs and a greater number of late complications in one RCT. Laparoscopic adjustable gastric banding required high rates of reoperation for band removal in one RCT.[64]

Colquitt et al. report that compared to sleeve gastrectomy no significant difference regarding weight loss was found for gastric bypass and there was no consistent picture across trials as to which procedure performed better regarding weight loss. Likewise, quality of life, mortality, reoperations, comorbidities, complications and additional surgical procedures were not significantly different between the two surgical techniques. The rate of serious adverse events were significantly higher in the laparoscopic Roux-en-Y gastric bypass group than after laparoscopic sleeve gastrectomy in one RCT.[64]

Colquitt et al. report on 2 RCTs that compared gastric bypass and biliopancreatic diversion with duodenal switch. In both studies the included patients had high BMIs (BMI > 48 kg/m^2 and BMI 50-60 kg/m^2 respectively). Weight loss was lower in the gastric bypass group than in the biliopancreatic diversion with duodenal switch.[64] Three years after surgery 82% of patients with gastric bypass and 100% of patients with biliopancreatic diversion with duodenal switch had a HbA1c <5%. The results for quality of life were similar on most domains for the two interventions. Reoperations were less frequent in the gastric bypass group and one death occurred in the biliopancreatic diversion with duodenal switch group.

The authors judged the quality of the evidence to be low or very low for most outcomes apart from mortality and – for the comparison with adjustable gastric banding and biliopancreatic diversion with duodenal switch- for the BMI, where it was judged to be moderate. The judgment of moderate quality of the evidence for mortality seems high given the extremely low event rates with just one event (death) per comparison.

3.3.2.4 **Additional analyses**

3.3.2.4.1 BMI

Stratified analysis of individuals with a BMI of ≥35 kg/m^2 and 25 to 34.9 kg/m^2 showed that effects on body weight loss, quality of life, HbA1c and diabetes remission were consistent in both strata. Differences in the magnitude of the effects could not be quantified, due to the small number of studies in each stratum. The number of studies included in these analyses was even smaller than for the non-stratified analyses because not all of the studies could be categorized into one or the other BMI stratum.

3.3.2.5 **Comparison with other evidence syntheses**

Results of this report are consistent with two recent systematic reviews.[64 65] Only one of them did a meta-analysis[65]. This meta-analysis by Gloy et al. included one trial more compared to the analysis in the present report and is due to the combined pooling of 1- and 2-year data in the former. The other systematic review[64] included 7 studies. Similarly to the results of the present report they found that bariatric surgery led to greater body weight loss, improvements in some aspects of quality of life and a higher remission rate of type 2 diabetes compared to conservative treatment; no deaths were reported in either of these systematic reviews.

The National Institute for Health and Care Excellence (NICE) in the UK has produced two clinical guidelines on the management of obesity. The older one, CG 43, had a very broad scope investigating amongst other things preventive and curative interventions in adults and in children in both clinical and non-clinical settings. Relative comparisons for the curative interventions included both the comparison of surgical and conservative interventions; but also the comparison of different surgical techniques with each other and the comparison of different conservative interventions with each other.[66]

The clinical guideline CG189 from 2014 is a partial up-date of the NICE guideline CG 43 and investigates amongst other things the effectiveness of bariatric surgery compared to conservative treatment in people with recent onset type 2 diabetes. Their definition of recent-onset type 2 diabetes was a duration of type 2 diabetes of 10 years or less. In their evidence statements they came to the conclusion that "the surgical intervention is more clinically effective than non-surgical treatment at": increasing percentage weight loss from baseline (very low quality), reduction in use of diabetic medication (moderate quality), increasing the remission of diabetes (very low quality), improving glycaemic control (low quality), and reducing weight as measured by BMI or kg (both low quality). These changes were deemed to constitute clinically important differences. It was reported that the mortality rates didn't appear to differ between the groups (very low quality) and that one study (Schauer 2014) reported greater improvements in quality of life for 5 of 8 domains of the RAND-36 questionnaire after gastric bypass, and 2 of 8 domains after sleeve gastrectomy.[67]

3.3.3 Conclusion

Sixteen RCTs fulfilled the inclusion criteria, of which 10 reported 2 or 3 year data. Most of the RCTs included patients with specific comorbidities (type 2 diabetes (n=6), sleep apnoea (n=2), mix of possible comorbidities (n=1)), only one RCT did not specify a comorbidity as an inclusion criterion.

The effects observed in the present meta-analyses were very consistent with a statistical significant benefit in the surgical arm compared to the conservative treatment for the critical outcomes percent body weight change, physical health (quality of life summary component score), HbA1c, and diabetes remission, and the important outcomes hypertension and dyslipidaemia.

No statistically significant effect was shown for mental health (quality of life summary component), stroke, all-cause mortality, SAE including reoperations, and sleep apnoea.

Data was extremely sparse (max. 1 event) or not existing for the critical outcomes stroke and the important outcomes mortality, myocardial infarction, cancer and revision rates, therefore no conclusions can be drawn for these outcomes.

In the stratified analyses for the different surgical techniques or on BMI (BMI 25-34.9 kg/m^2 vs. BMI ≥35 kg/m^2), the direction of the observed effects – when they were present – consistently showed a benefit for bariatric surgery compared to conservative treatment. However, fewer RCTs were available within such strata.

The quality of evidence was moderate for the critical outcome percentage body weight change and low for change in HbA1c. The overall quality of the evidence is judged to be very low because of the very low quality of evidence for the critical outcomes mental health and stroke; as well as for the absence of data from RCTs for myocardial infarction.

Significant uncertainties remain regarding outcomes in the long-term (>3 years) and the overweight population, where little or no data were available.

4 Health economic analyses

4.1 Methods

While there is a relatively broad international literature, there are only few studies for Switzerland which examine the cost (economic burden) of obesity. There exists no Swiss study of the cost-effectiveness of bariatric surgery or gastric bypass for overweight (BMI 25-29 kg/m^2) and obese (BMI ≥30 kg/m^2) adults. Due to a lack of suitable Swiss data, e.g. relating to healthcare resource use, we have not performed a “de novo” health economic model, Instead, we undertook a systematic review of the international economic literature to understand the impact of bariatric surgery in terms of cost-effectiveness. We aim to gain an understanding of the cost-effectiveness of bariatric surgery from a 'KVG perspective' (third party perspective, taking into account the direct medical costs of all health care services covered by the Swiss statutory health insurance, irrespective of actual payer) and from a societal perspective.

Eligible studies were assessed for quality and transferability to the Swiss situation; a number of adaptations of costs and incremental cost-effectiveness ratios have been undertaken. Given the literature-based approach, we had multiple time horizons and perspectives in the available studies. We made this transparent and took it into account as a factor potentially explaining ICER differences. Budget impact has been addressed on this basis and using Swiss epidemiological data and cost data.

The patient population of interest, the intervention and comparator strategies ('PIC' of the PICO) correspond to those used in the other parts of this HTA report. Health economic endpoints considered included costs, life year gained, QALYs, and incremental cost-effectiveness ratio (ICER, i.e. costs per life year gained or QALY gained). To the extent possible, patient groups with a BMI ≥35 kg/m^2 vs. patients with a BMI 25-34 kg/m^2 were regarded separately.

The health economic analysis included the following steps, which are detailed in the subsequent sections.

1. Literature search.
2. Screening of the search results to identify eligible cost-effectiveness studies and of studies that may be of secondary interest (e.g. providing relevant cost parameters).
3. Extraction of information and quality assessment of the eligible cost-effectiveness studies.
4. Assessment of the eligible cost-effectiveness studies in terms of transferability to Switzerland.
5. For the studies found to be transferable, adaptation of cost-effectiveness results to Switzerland.
6. Comparison of studies and synopsis of findings, with a focus on differences.

4.1.1 Literature search

The aim of the literature search has been to identify literature on the cost and cost-effectiveness of bariatric surgery, with a focus on gastric bypass surgery as comparator, for overweight and obese, adult patients. All types of economic evaluation were considered: cost-effectiveness analyses, cost-benefit analyses, cost-utility analyses and cost minimisation analyses. A search strategy has been developed to identify all relevant literature in the following electronic databases: Medline and Embase databases including abstracts by using OvidSP (including Ovid MEDLINE(R), Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily Update, Embase), the Cochrane Library and the Centre for Review and Dissemination (CRD) database including the Database of Abstracts of Reviews of Effects (DARE), Cochrane reviews, Health Technology Assessments (HTA) and the Economic Evaluation Database from the UK National Health Service (NHS EED). Search strings for additional databases were not developed because the selection of databases described above has been described as both sufficient and very efficient.[68] The economic search string was obtained from the NHS R&D programme, which performed a HTA of "the clinical effectiveness and cost-effectiveness of surgery for people with morbid obesity: a systematic review and economic evaluation".[6 69] The search strategy included keywords and medical subjects headings (MeSH terms) concerning cost, cost-effectiveness and health economic studies.

In the Cochrane library we used the option to search for Technology Assessments and Economic Evaluations. Search results were restricted by selecting human studies only, and removing duplicates. We used a restricted set of keywords to focus our search on health economic studies.

The search strategy and MeSH terms used for each database are described in the Appendix 12. The first search was performed on March 20th, 2015 and was updated on August 6th, 2015.

4.1.2 Screening of the search results

The screening of the literature was divided in three phases. In the first phase all results of the literature search was screened by title. Titles containing relevant keywords such as bariatric surgery, costs, value, cost-effectiveness, cost-utility, quality of life, and burden were considered as potentially relevant.

All papers with potentially relevant titles then proceeded to the second phase, the screening by abstract. In this phase, abstracts (and in particular the results sections of the abstracts) were screened for relevant numbers (e.g. costs, life year gained, QALYs, or ICERs) or for sentences suggesting potentially interesting content in the full text version.

Potentially relevant abstracts then proceeded to the third phase, in which full texts were screened. Articles were then classified as being relevant, partially relevant, or irrelevant.

- Relevant articles needed to meet the following criteria:
    - The article reported a full-scale incremental cost-effectiveness analysis, ideally but not necessarily with an endpoint of cost per QALY gained or life year gained
    - The 'PIC' of the PICO corresponded to the one defined in the scoping document and used in the systematic review part of this HTA report.
    - The analysis was performed for a jurisdiction with broadly similar socioeconomic characteristics as Switzerland. Studies for North, Central and Western European countries (including Italy, Spain and Portugal), the USA, Canada, Australia and New Zealand were considered.
- Partially relevant articles were defined as not meeting the criteria for the 'relevant' category but potentially containing useful additional information concerning effectiveness or costs (e.g. just the costs of bariatric surgery but no effectiveness information or ICER). Depending on the quality and quantity of information available from relevant articles, some partially relevant articles were used as an additional source of information and for comparison.
- The remainder of articles were classified as irrelevant

4.1.3 Extraction of information and quality assessment

For the eligible cost-effectiveness studies (i.e. relevant articles as defined above), data extraction was performed, covering the following information:

- Study population (including country, age and BMI range of the patients)
- Intervention
- Comparator(s)
- Setting and perspective of the study

- Cost types included and cost year
- Type of model
- Time horizon
- Discount rate
- Approach to sensitivity analysis
- Effectiveness
- Costs
- Incremental cost-effectiveness ratio (ICER)

A brief, qualitative characterization of each study was prepared in the results section, covering methodological approaches taken, main data sources, particularities, methodological issues and potential meaningfulness of the results for Switzerland.

Quality of reporting was assessed against the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) 24-item checklist, recommended by the ISPOR Health Economic Evaluations Publication Guidelines Task Force. For each study, each item of the checklist received a score of either "0", "0.5", "1" or "not applicable" ("NA"), depending on the level of transparency of reporting.

Issues (i.e., criterion not met) were reported by CHEERS criterion, across studies and study by study. In addition, selected CHEERS criteria were defined which needed to be met in order to make a study transferable and thus suitable for numerical adaptation of ICER results to Switzerland (see next section). Of note, it was expected that CHEERS criteria 4 (population), 7 (intervention / comparator[s]) and 10 (outcome measures) will be met by all studies assessed because studies failing any of these criteria would not meet the eligibility criteria defined in section Assessment of transferability. Figure 18 illustrates how studies were selected and how transferability was determined.

4.1.4 Assessment of transferability

Some international cost-effectiveness studies were expected to be 'qualitatively transferable' to Switzerland. A variety of authors have worked on criteria for assessing such transferability between jurisdictions.[70] Relevant, related aspects have been published early by O'Brien et al.[71] Welte at al. and Drummond et al. have suggested related multistep procedures.[72 73] In the present study, we perform a modified approach as described below and summarised schematically in Figure 18. One key reason for using this modified approach is that we did not have available any of the original models underlying the eligible cost-effectiveness studies. Therefore, actual model recalculations on the basis of localized input parameters such as e.g. unit costs, were not possible.

- The most top-level criteria are covered by the eligibility criteria. Essentially, this step excluded studies which were not full scale health economic evaluation studies assessing incremental cost-effectiveness, did not meet the 'PIC', or were performed for countries very different from Switzerland in terms of socioeconomic characteristics. All remaining studies had to meet CHEERS criteria 4, 7 and 10.
- Studies not meeting CHEERS items 5, 6, 8, 13, 14 and 19 were regarded as not transferable due to lack of key information. In relation to item 19, the availability of costs and outcomes of interest for both the intervention and the comparator strategies was considered fundamental. Where articles only reported ICERs, the underlying study was considered non-transferable.
- The remaining studies were considered qualitatively transferable, and underwent numerical adaptation of cost-effectiveness results, if scrutiny of transferability factors taken from O'Brien et al. and Welte et al. did not preclude this for a specific reason.[71 72] In all other cases, the results of the scrutiny of transferability factors were used qualitatively.

The following transferability factors were considered:
Methodological characteristics:

- Perspective of cost assessment
- Discount rate
- Medical cost approach
- Productivity cost approach

Healthcare system characteristics:

- Absolute and relative prices in healthcare
- Clinical practice variation; differences in resource use, incentives and regulations for health-care providers
- Technology availability

Population characteristics:

- Demography
- Disease incidence and prevalence
- Case-mix
- Life expectancy
- Health-status preferences
- Acceptance, compliance, incentives to the patients
- Productivity and work-loss time

For the majority of cost-effectiveness studies meeting the general eligibility criteria we did not expect severe transferability problems since methodological and population characteristic were expected to be similar to Switzerland. Regarding healthcare system characteristics, we do not expect big differences in availability of technology. Absolute prices in healthcare were adapted numerically (see next section), while differences in relative prices were expected to pose certain issues. With respect to clinical practice, bariatric surgery procedures should be well comparable across countries. In contrast, standard care (i.e. the comparator strategy) was expected to vary across studies and countries. Therefore, we assessed to the extent possible to what kind of clinical practice the comparator groups were exposed.

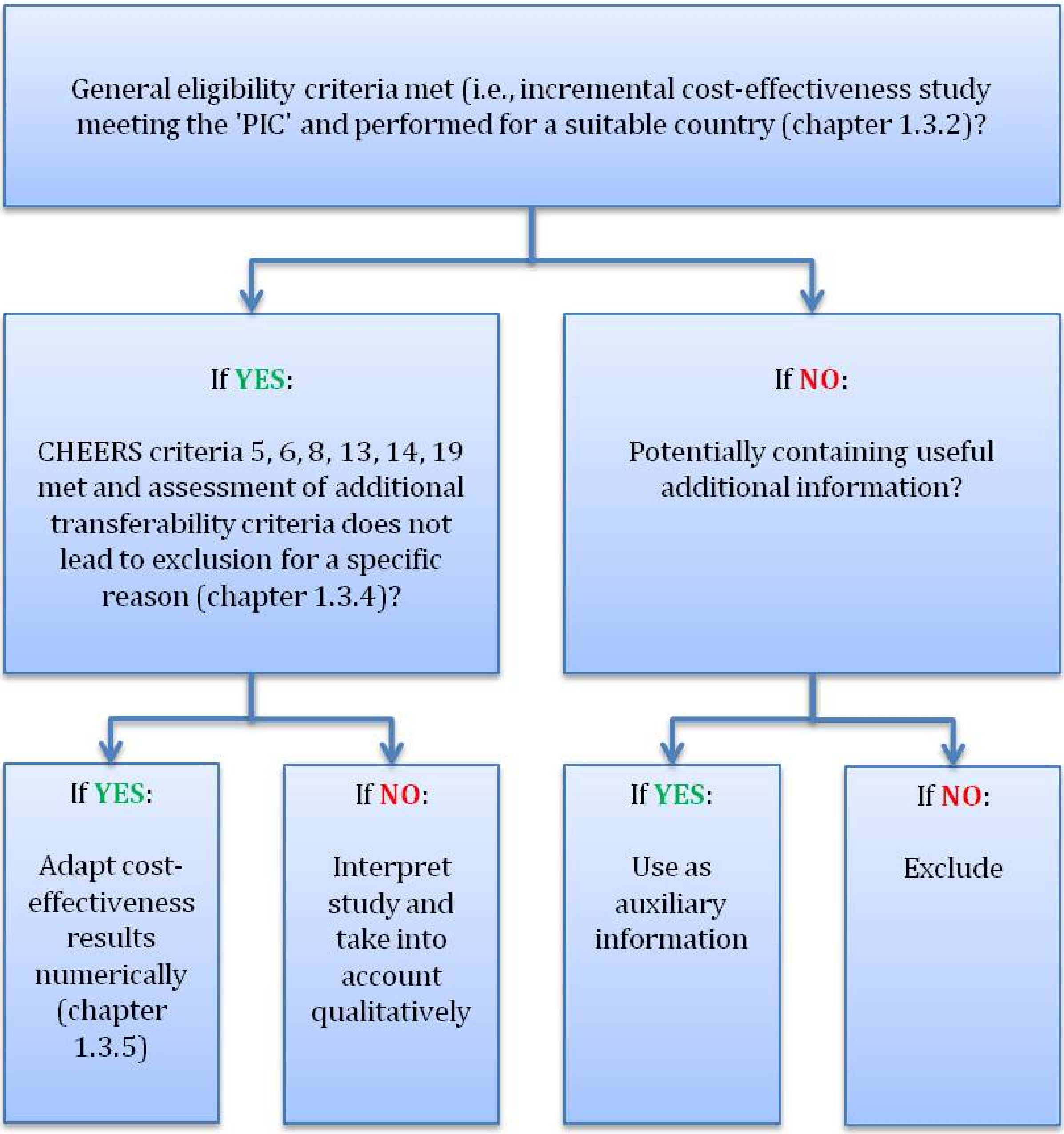


Figure 18 Steps for study selection and determination of transferability to Switzerland.

Studies falling in the lower left box regarded as being qualitatively transferable.

4.1.5 Adaptation of cost-effectiveness results to Switzerland

The adaptation of cost data representing direct medical costs for Switzerland was performed in three distinct steps: correction for different levels of resource utilisation, for different prices of healthcare services, and for change in level of resource utilisation and prices over time. Subsequently, adapted ICERs were calculated. This process cannot be interpreted as achieving realistic ICERs for Switzerland, but may have achieved a certain approximation of cost-effectiveness levels to be expected for Switzerland. It has certainly made the results of international CE studies, reported for different countries and in different currencies, more comparable.

1. *Resource utilisation:* The types and quantities of resource utilisation differ between countries. For the same disease, patients in Switzerland often receive more medical treatments than in other countries (i.e. they are treated more intensively for an equivalent diagnosis). Therefore a “quantity correction" is necessary. The quantity correction was based on the Organization for Economic Co-operation and Development (OECD) statistics of healthcare expenses per capita, corrected for purchasing power. A correction for differences in resource utilisation levels (unaffected by price levels) was thus achieved.[74]
2. *Prices of healthcare services:* The price for the same healthcare service or treatment is often different across countries. A “price correction” was achieved by applying correction factors provided by the OECD. Such purchasing power parities represent the proportional costs for identical products in two countries.[75]
3. *Change in costs over time:* Healthcare costs change over time. For eligible cost-effectiveness studies performed in countries other than Switzerland, the two steps described above achieve an adaptation of reported costs. However, the resulting estimates are valid for the same cost year as in the original study. Additional correction for the development of costs over time was necessary. In the case of a specific disease and set of treatment strategies, costs may change over time due to mere price changes but no changes in resource utilisation, or resource utilisation for the treatment of the disease of interest may also change. In our 'base case' approach, we assumed the latter, and that changes in resource utilisation occur with the same cost impact as at the level of total Swiss health care expenditures. The resulting correction was based on the yearly growth rates of total Swiss healthcare expenditures, as reported by the Swiss Federal Office of Statistics.[76] In a sensitivity analysis, we alternatively assumed no change in resource utilisation over time. The resulting, alternative correction was thus based on the change in Swiss price levels. (General instead of healthcare-specific change in price levels was used, as the reporting of the latter may have been influenced by recent changes in the methodology applied by the Swiss Federal Office of Statistics.)

The adaptation of cost data representing indirect costs followed a similar approach. However, the first of the above described steps is irrelevant in the case of indirect costs. The third step was based on the change in Swiss salaries over time.[77]

4.1.6 Sensitivity analyses

In an alternative adaptation of cost results and ICERs, the third step (adjustment for changes in resource use and costs over time) was based on inflation rates instead of the growth rate of health care expenditures, to cover uncertainty in the adaptation algorithm.

4.1.7 Synthesis of findings

The resulting different pieces of information were synthesized. This necessarily involved an element of interpretation but it was an explicit aim to make all related assumptions transparent. Discussion includes a critical review of possible sources of uncertainty. Comparisons of the assumptions and of the data used by the various cost-effectiveness analyses are provided.

4.1.8 Budget impact analyses

Possibilities for budget impact analysis (BIA) were limited given the available data. Considering this, we pursued two approaches. The first was to estimate the potential surgery costs for treating all eligible patients in a single year. The second was to estimate the costs of bariatric surgery over a time period of ten years, from 2011 until 2020, from the perspective of the Swiss healthcare system. We did not model BIA from a societal perspective.

4.1.8.1 Surgery costs

The potential surgery costs for treating all eligible patients with a gastric bypass in a single year were estimated. The Swiss Federal Office of Statistics reported that 10.3% of the Swiss population may have a BMI above 30 kg/m^2.[78] Using the BMI distributions reported in four Swiss studies (the Swiss monitoring of trends and determinants in cardiovascular disease (MONICA); the socio-medical indicators for the population of Switzerland (SOMIPOPS); the Swiss Health Survey (SGB); the National Research Program 1A (NRP1A)), we estimated that of the patients with BMI above 30 kg/m^2, 82.2% have obesity grade I (BMI 30-35 kg/m^2), 14.3% have obesity grade II (BMI 35-40 kg/m^2), and 3.5% have obesity grade III (BMI >40 kg/m^2) (own estimations, based on Buri 2015).[79] Patients with a BMI >40 kg/m^2 are considered directly eligible for surgery, whereas the eligibility of patients with BMI 35-40 kg/m^2 was assumed to depend on the prevalence of relevant comorbidities (e.g. diabetes or cardiovascular disease). For the present calculation, comorbidity prevalence rates ranging from 0% to 30% were assumed. For the immediate costs of surgery, data from the **Swiss Study Group for Morbid Obesity** (SMOB, www.smob.ch) and from the Swiss DRG system (Version 2.0) were used. The cost for one single laparoscopic gastric bypass surgery in Switzerland was estimated at CHF 14,174. (It was conservatively assumed that all bariatric surgery operations would be performed without complex procedures.)

4.1.8.2 Budget impact analyses of bariatric surgery over 10 years

This analysis was performed using several national and international sources. Cost estimates resulting from the adaptation of transferable cost-effectiveness studies to Switzerland were combined with the current use of bariatric surgery in Switzerland, to the extent possible. Sources used and assumptions made were as follows.

The costs of bariatric surgery were estimated over a time period of ten years, from 2011 to 2020. Information on the number of bariatric surgeries performed in Switzerland in the last few years was available from the SMOB. Data from 2011 to 2014 were available and showed that in the last few years, there was a constant increase in the number of bariatric surgeries (from about 2,300 in 2011 to nearly 4,200 in 2014). From 2015 on, we assumed a constant number of surgeries per year (i.e. N=5,000). This number reflects the increase in the number of bariatric surgeries in the last few years. By keeping the number constant for the following years, we conservatively assume that the number of obese patients requiring surgery will not increase further. This is related to the fact that the number of available hospitals and surgeons potentially limits the maximum possible number of yearly performed surgeries. With the actual infrastructures, 5,000 operations per year seem to be realistic.

For the immediate surgery costs, the above-introduced value of CHF 14,174 was used. In addition to the surgery costs, follow-up costs were calculated for a 5-year time horizon, using costs estimations by Ackroyd et al., adapted to Switzerland.[80] The rationale for using this study as basis for our calculations is provided in the discussion. Ackroyd et al. provided the 5-year costs of surgical and conservative treatment strategies in Germany, for patients with and without diabetes. For laparoscopic gastric bypass, mean yearly follow-up costs of CHF 551 per patients were assumed. For conservative treatment, the mean follow-up costs in the first year per patient were assumed to be CHF 3,588 (including six physician consultations, two laboratory assessments, and food substitutes for the whole year). In the follow-up years 2 to 5, yearly costs of 590 were assumed (including one physician consultation and one laboratory assessment). The cumulative assessment of the effectiveness of the intervention was primarily based on the remission from diabetes. For subjects undergoing surgery or receiving conservative treatment, 2.8 or 0.2 diabetes-free-years were estimated, respectively. For patients with diabetes, independent of surgery, yearly costs of CHF 11,414 were assumed (including, hospital care, ambulatory care, anti-diabetic drugs, and other care). The estimated mean yearly costs per patient undergoing surgery or usual care are presented in Table 8.

Table 8 Mean yearly costs per patient undergoing surgery or usual care.

	1 year	2 year	3 year	4 year	5 year
GBP costs (CHF), with diabetes	14,724	551	2,834	11,965	11,965
GBP costs (CHF), without diabetes	14,724	551	551	551	551
Usual care costs (CHF), with diabetes	12,719	12,004	12,004	12,004	12,004
Usual care costs (CHF), without diabetes	3,588	590	590	590	590

For 2011, the starting year in this analysis, only surgery costs were included. For 2012, the follow-up costs of patients operated in 2011 were added to the annual surgery costs. For 2013, the follow-up costs of 2011 and 2012 were added to the annual surgery costs. For 2014, the follow-up costs of 2011, 2012, and 2013 were added to the annual surgery costs. From year 2015 to 2020, the cumulative costs of the yearly surgery and the previous four years were calculated.

Two scenarios were calculated. In the first scenarios, we assumed that all patients undergoing surgery or conservative treatment had diabetes at baseline (i.e. diabetes prevalence = 100%). In contrast, the second scenario assumed that no patients had diabetes at baseline (i.e. diabetes prevalence = 0%). Combining the two scenarios allowed us to estimate the budget impact of bariatric surgery assuming different diabetes prevalence rates.

4.2 Results

4.2.1 Study selection

A total of 1,183 citations were identified from the searched electronic databases. Following the removal of duplicates, 983 full citations were reviewed by two independent investigators (Figure 19). Based on title and abstract, the two investigators excluded 907 citations due to inappropriate comparator, non-comparative design; character of a review or commentary piece; inappropriate outcome measure; or no relevant cost information given. A total of 76 citations were included for full-text review, of which another 55 were excluded for the same categories of reasons as above. The remaining 22 articles fulfilled the inclusion criteria. They were included in the systematic review and assessed using the CHEERS checklist and the algorithm shown in Figure 18. A total of 15 articles fulfilled transferability criteria (see section 4.1.4), which was needed to make a study qualitatively transferable and thus suitable for numerical adaptation of ICER results to Switzerland.

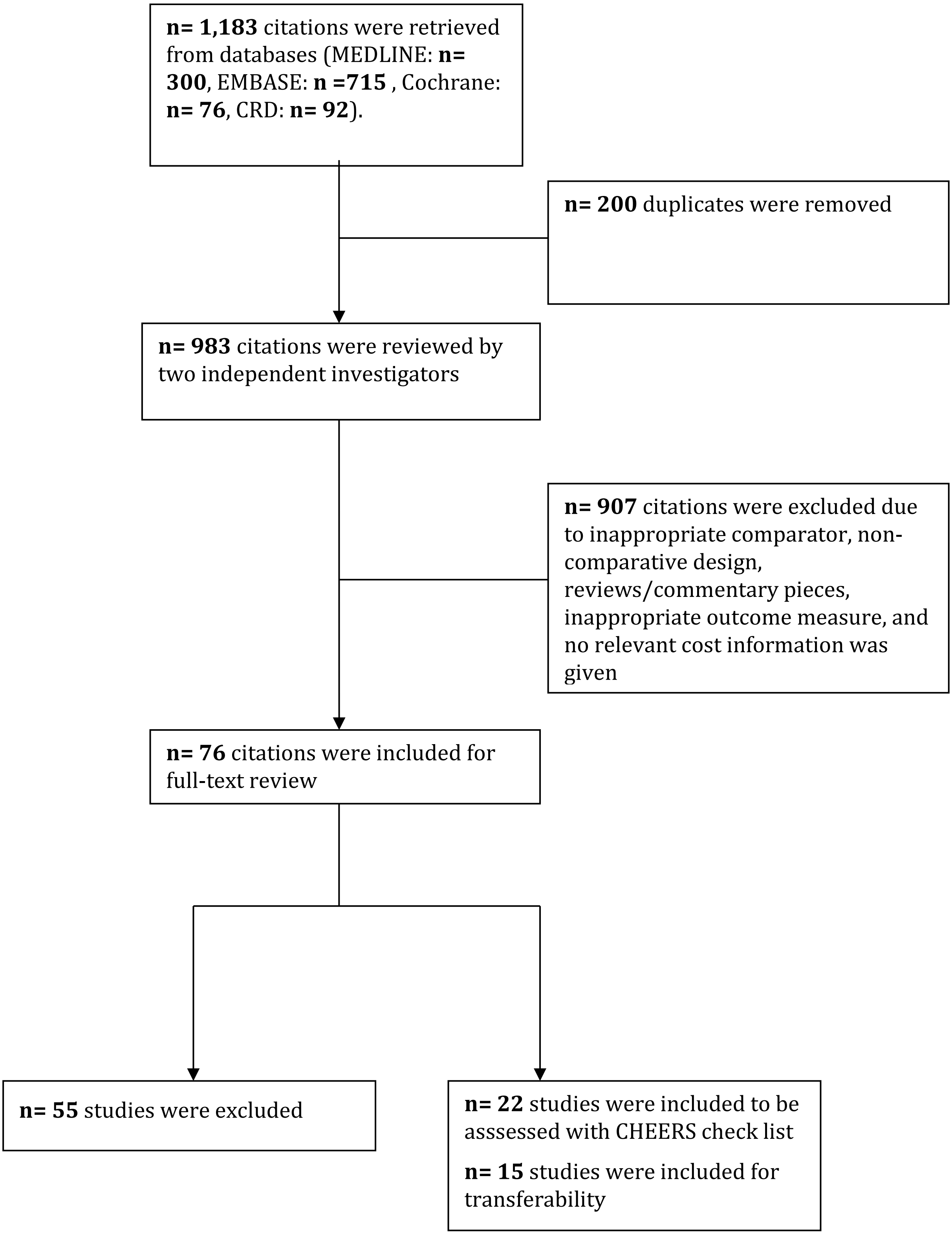


Figure 19 Flow chart describing the systematic process of article selection.

4.2.2 Characteristics and methodology of published cost-effectiveness studies

4.2.2.1 Study and patient characteristics

Of the 22 included studies, nine were from the United States, four from the United Kingdom, two from Australia. The remainder were from continental Europe. All studies were published between 2002 and 2014. Six studies (27%) reported private funding sources, eight studies (36%) reported public funding, seven studies (32%) did not report a source of funding, and one study (Keating 2009) reported mix funding sources.

An overview of the characteristics and demographics of the patient populations studied is provided in Table 9, together with information on intervention, comparator, setting, time horizons of the analyses and type of modelling. Additional details are provided in the study by study description of characteristics and results in section 9.3. Here we give a brief summary.

Seven of the assessed studies incorporated patients with BMI >35 kg/m^2 and type 2 diabetes mellitus (type 2 diabetes) (Ackroyd 2006, Anselmino 2009, Hoerger 2010, Ikramuddin 2009, Keating 2009, Pallock 2013., Picot 2009 and 2012). Hoerger et al. additionally divided the study population into patients with newly diagnosed diabetes and patients with established diabetes (Hoerger 2010).

The rest of the studies incorporated patients with BMI >35 kg/m^2 and with a varying degree obesity related comorbidities such as type 2 diabetes, hypertension, coronary artery disease, stroke, and other life threatening diseases. Mean BMI varied between studies, most of the studies assessed ranges above 35 kg/m^2 (Table 9).

Few studies assessed lower BMI populations, a study by Borg et al. assessed costs and effects of bariatric surgery in a population with BMI 30-34 kg/m^2 (Borg 2014). Similarly, Faria et al. assessed groups of patients with BMI 25-30 kg/m^2 and 30-35 kg/m^2. Keating et al. assessed the cost-effectiveness of LAGB in patients with BMI 30-40 kg/m^2.

Another important component considered in studies was gender and age categories. For example, the study by Borg et al. was very detailed in reporting cost-effectiveness for different ages in the BMI category of 40-44 kg/m^2. Similarly, for age category 45-54 years, results were reported for different BMI categories. All of this information was provided for males and females (Borg 2014). Campbell et al. reported results for females and males in different BMI categories, with a base-case age of 40 years. Craig et al. reported data for age categories 35 years old and 55 years old, for both genders separately, and for BMI categories of 40 kg/m^2 and 50 kg/m^2. The rest of the studies did not consider differences between gender and age.

4.2.2.2 Intervention

The cost-effectiveness analyses evaluated interventions representing surgical treatments for obesity. The majority assessed gastric bypass (Ackroyd 2006, Anselmino 2009, Borg 2014, Castilla 2014, Clegg 2003, Craig 2002, Faria 2013, Horger 2010, Jensen 2005, Mäklin 2011, Picot 2009); in some cases, other terms such as Laparoscopic Roux-en-Y gastric bypass (LRYGB) (Campbell 2010, Ikramuddin 2009, McEwen 2010, Michaud 2012, and Salem 2008) or conventional open Roux-en-Y gastric bypass (ORYGB) were used (Wang 2014). Eight studies explicitly stated that gastric bypass was performed laparoscopically (Ackroyd 2006, Anselmino 2009, Campbell 2010, Clegg 2003, Michaud 2012, Picot 2009, Salem 2008, and Wang 2013). In Craig et al. only open gastric bypass was included (Craig 2002). Castilla et al. assumed 50% laparoscopic and 50% open gastric bypass (Castilla 2014). In the study of McEwen et al. 65% of the patients underwent open Roux-on-Y gastric bypass, whereas 34% underwent LRYGB. In six studies on gastric bypass it was not clearly stated if the intervention was performed laparoscopically or not (Borg 2014, Faria 2013, Hoerger 2010, Ikramuddin 2009, Jensen 2005, Mäklin 2011).

Few studies reported the cost-effectiveness of other surgeries such as adjustable gastric banding, or laparoscopic adjustable gastric banding (Ackroyd 2006, Ananthapavan 2010, Anselmino 2009, Faria 2013, Hoerger 2010, Keating 2009, Mäklin 2011, Picot 2009 and 2012, Pollock 2013, Salem 2008 and Wang 2014), all compared to conservative treatment or no surgery.

Three studies reported on mixed intervention strategies involving gastric bypass and gastric banding surgery (Ackroyd 2006, Horger 2010, and Mäklin 2011). Wang et al reported on LRYGB, ORYGB and LAGB, compared to conservative treatment (Wang 2014).

The studies by Ananthapavan et al., Anselmino et al., Faria et al, Jensen et al, McEwen et al., Padwal et al., and Salem et al., were found to be not qualitatively transferable to Switzerland as they were not providing sufficient information on costs and effects.

4.2.2.3 Comparator

Standard medical management, described as 'conservative therapy' or 'current practice', was typically a combination of diet, exercise, and behavioural modification, and was comparator in 20 studies. Some of these studies were very specific about 'current practice', reflecting that they studied patients with type 2 diabetes (Ackroyd 2006, Hoerger 2010, Ikramuddin 2009, Keating 2009, Picot 2012, and Pollock 2013). For example, usual care in these cases included tight glycaemic control, exercise and other behavioural modification. The two remaining studies described the comparator as 'no intervention' (Campbell 2010 and Craig 2002).

4.2.2.4 Type and characteristics of economic evaluations

Most studies were cost-utility analyses, i.e. cost-effectiveness analyses using QALYs as their benefit measure. The only exception was the study by Michaud et al. (Michaud 2012), which reported life years and costs per life year gained. Decision analytic modelling employing Markov elements or life tables was used in all of the studies.

4.2.2.5 Perspective of studies

The perspective of an economic evaluation is important for the decision maker in order to decide to whom the costs occur. This is very important as an intervention may e.g. be cost-effective from a societal perspective, but not from a health care or third party payer perspective. In this systematic review, 21 out of 22 studies adopted the perspective a health care provider or third party payer perspective, meaning that either all direct medical costs or a subset of these were incorporated in these analyses. One study did not follow recommendations of reporting perspective (McEwen 2010).

Five studies additionally reported costs form the societal perspective (Ananthapavan 2010, Borg 2014, Faria 2013, Jensen 2005, Michaud 2012). However, three of these studies were not transferable due to missing information on the year of cost data (Ananthapavan 2010, Faria 2013, Jensen 2005). Only two studies (Borg 2014 and Michaud 2012) were transferable to Switzerland and incorporated the societal perspective.

4.2.2.6 Time horizon of studies

Selecting an appropriate time horizon for cost-effectiveness analysis is important due to a potential impact on the results. The time horizon should be long enough to measure all clinically relevant costs and benefits, in our example bariatric surgery compared to conservative treatment.

While twelve out of 22 of the reviewed cost-effectiveness studies adopted a lifelong time horizon for the reporting of costs and effects, Pollock 2013 used a 40 years-time horizon and Ikramuddin 2009 used a 35 years-time horizon. Four studies reported 20 years-time horizons, namely Clegg et al., Catilla et al., Picot et al., (Clegg 2003, Catilla 2014, Picot 2009, and Picot

2012), while Michaud et al., (Michaud 2012) used 10 years and Ackroyd et al., and Anselmino et al. (Ackroyd 2006 and Anselmino 2009) used 5 years.

4.2.2.7 **Discounting**

Any economic evaluation where costs and effects occur over more than a year should apply discounting.[81] Discounting ensures that for costs and effects occurring at different points in time, only the present value at the initial intervention time is taken into account. In the current systematic review, a total of 19 out of 22 studies reported discounting. Two thirds (14 studies) used 3.0% discounting for costs and effects (Ananthapavan 2010, Borg 2014, Campbell 2010, Castilla 2014, Craig 2002, Faria 2013, Hoerger 2010, Ikramuddin 2009, Keating 2009, Mäklin 2011, McEwen 2010, Michaud 2012, Salem 2008, Wang 2014). Five used 3.5% discounting of costs and effects, reflecting the UK standard (Ackroyd 2006, Anselmino 2009, Picot 2009 and 2012, and Pollock 2013). One early study discounted costs by 6.0% and effects by 1.5% (Clegg 2003). Jensen et al. did not provide appropriate information on their approach to discounting (Jensen 2005). Padwal et al. was a systematic review (Padwal 2011).

Table 9 Population demographics and characteristics of included cost-effectiveness studies.

Study	Population/ Country	Age	Intervention	Comparator	Perspective /setting	Time frame	Cost Types/year	Discounting	Modelling
Ackroyd 2006[80]	Patients with a BMI of >35 kg/m² and type 2 diabetes in Germany, UK and France.	Not reported	Laparoscopic gastric bypass and adjustable gastric banding.	Conservative treatment (Monitoring or medically guided dieting) -Germany: 6 GP consultations, 2 Lab assessments, 365 Food substitutes. -France: 30 days in institution, 6 specialist consultations, 2 Lab assessments, 2 Food substitutes. - United Kingdom: 4 GP consultations, 4 nurse consultations, 2 dietician consultations, 1 Lab assessment, 56 Food	Payers perspective.	5 years.	Direct costs/ 2005.	3.5%	Decision analytic modelling with incorporation of budget impact analyses.

				substitutes.					
Anathapavan 2010 [82]	28 adolescents who had a gastric banding surgery in Melbourne (Australian). Patients had a BMI of >35 kg/m^2.	Average age 14 -19 years.	Laparoscopic adjustable gastric banding.	Current practice: combination of diet, exercise, and behaviour modification.	Societal perspective.	100 years.	Indirect-Direct costs/2001.	3%	Simulation-modelling .
Anselmino 2009 [83]	Patients with a BMI of ≥35 kg/m^2 and type 2 diabetes mellitus in Austria, Italia and Spain.	Adults .	Laparoscopic gastric bypass and laparoscopic adjustable gastric banding.	Conservative treatment (in Spain includes two low calorie diets), in Austria It includes annual medical check-ups, whereas Italian one is based in assumptions that resources for CT in England's National Health service.	Payer perspective in three Europeans countries: Austria, Italy and Spain.	5 years.	Direct costs/2009.	3.5%	Previously published economic model which incorporates eight output variables.

Borg 2014 [84]	Female and male in Sweden with age below and above 55 years of age. stratum age 45-54 years and BMI 40-44 kg/m² as base case.	Stratum age 45-54 years in the base case.	Gastric bypass (not specified if laparoscopic or open).	The conservative treatment alternative consisted of the prevalent mixture of non-surgical obesity treatments (Sjöstrom 2013).	Societal perspective.	Life time horizon.	Direct costs/2012.	3%	A Markov micro-simulation model.
Campbell 2010 [85]	Patients with BMI >40 kg/m² or BMI >35 kg/m² with comorbid conditions in the US.	18 – 74 years.	Laparoscopic Roux-en-Y gastric bypass and laparoscopic adjustable gastric banding.	No intervention.	Third-party payer.	Life time horizon.	Direct costs/2006.	3%	Markov model of treatment and out-comes for obese adults.
Castilla 2014 [86]	79 patients from Spain with an average BMI of 50.7 kg/m² (range 36.6 to 76.3).	18 - 55 years.	Gastric bypass (50% laparoscopic and 50% open).	Usual care - Diet, exercise, behaviour modification.	Payer perspective.	5/10/15/20 years and Lifetime.	Direct costs/2012.	3%	Discrete-event simulation model.
Clegg 2003 [69]	Hypothetical cohort of 100 patients, 90% females from the UK with an average BMI of 45 kg/m².	40 years.	Laparoscopic gastric bypass and laparoscopic adjustable gastric banding	Non-surgical management: 4 GP visits, 2 dietitians contacts, 2 practice nurse contacts, 2 district nurse contacts, every 3	NHS perspective and a personal social services perspective for the costs	20 Years.	Direct costs/1999-2000.	Costs 6% and QALY 1.5%	Results are based on a systematic review.

				years a VLCD for 12 weeks.	and benefits.				
Craig 2002 [87]	Severe obese >40 kg/m² patients from the US which were non smokers without cardiovascular disease.	35 - 55 years.	Open gastric bypass.	No surgery.	Payer perspective.	Lifetime.	Direct medical costs (surgery, complications, follow-up, comorbidities)(2001).	3%	Decision model.
Faria 2013 [88]	Hypothetical population of 10.,000 samples and 1'000 trials for each sample. Patient had a mean BMI of 49.6 kg/m².	Cohort age at model entry 40 years.	Gastric band and gastric bypass (not specified if laparoscopic or open).	Best medical management.	Societal perspective.	Lifetime.	Direct/indirect costs.	3% after the first cycle	Markov model.
Hoerger 2010 [89]	Patients from the US with BMI ≥35 kg/m² and newly or established diabetes.	Not reported	Gastric bypass and banding surgery (not specified if laparoscopic or open).	Usual diabetes care including tight glycaemic control	Unclear.	Lifetime.	Direct costs (2005) (surgery, complications, follow-up)	3%	Markov model.

Ikramuddin 2009 [90]	Diabetes patients from the US, 22.1% were males with a mean BMI of 48.4 kg/m^2.	Mean 50.1 years.	Laparoscopic Roux-en-Y gastric bypass (LRYGB) (not specified if laparoscopic or open).	Standard medical treatment of obese type 2 diabetes patients	Third-party payer perspective.	35 years.	Direct medical costs (2007) of surgery (surgery, follow-up, complications) + medical management of diabetes.	3%	Standard Markov structure, combined with Monte Carlo simulation and tracker variables.
Jensen 2005 [91]	Hypothetical population of white women from the US. Patients in gastric bypass surgery group had a BMI of 40 (at age 40) and patients in the non surgical group had a BMI of 33 kg/m^2 (at age 18).	Gastric bypass group = 40 years. Non surgical group = 18 years.	Gastric bypass (not specified if laparoscopic or open).	Non surgery i.e. diet and exercise.	Societal perspective.	Lifetime.	All costs were inflated to 2004.	Inflation costs of 2004 were used.	Decision model.
Keating 2009 [92]	Patients with recently diagnosed type 2 diabetes from Australian with a BMI of 30 - 39.9 kg/m^2 . 55% were females.	Mean age of 49 years.	Laparoscopic adjustable gastric banding.	Conservative therapy (diet, physical activity and medical therapy).	Health care perspective.	Lifetime.	Direct costs (2006) of surgery (surgery, complications, follow-up) + costs of diabetes (remission, monitoring,	3%	Markov model.

							treatment).		
Mäklin 2011 [93]	Patients from Finland with a mean BMI of 47 kg/m² (38-59) and 35% were males.	Mean age of 43 years.	Gastric bypass and gastric banding (not specified if laparoscopic or open).	Ordinary treatment (was considered to include a range of interventions from brief advice given by physicians to intensive conservative treatment, as there is no standard medical treatment for obesity in use in Finland).	Health care perspective.	10 years.	Direct costs (2010) (surgery, complications, follow-up) including ordinary treatment, but excluding medication.	3%	Combination of decision tree and Markov model.
McEwen 2010 [94]	221 US patients with BMI>40 kg/m² or BMI >35 kg/m² and two life-threatening conditions. 88% were females.	42 +/- 10 years.	65% open and 34% laparoscopic Roux-en-Y GB.	Usual care before intervention.	Not clear.	2 years and lifetime.	Direct costs.	3%	A simple comparison of costs before and after surgery.

Michaud 2012 [95]	Current eligibility: BMI >40 kg/m^2 or BMI >35 kg/m^2 with high risk comorbidities (Extended eligibility to those with BMI >35 kg/m^2 or BMI >30 kg/m^2 with qualifying comorbidities).	50 years.	Laparoscopic Roux-en-Y gastric bypass (LRYGB).	Baseline management.	Societal perspective.	Lifetime.	Costs of treatment, change in medical costs, change in deadweight loss and change in income (2010).	3%	Future Elderly Model (FEM), an established and well-studied microsimulation model of aging and health.
Padwal 2011 [96]	Systematic review with summary of several studies.	Systematic review with summary of several studies.	Bariatric surgery (i.e., adjustable gastric banding, Roux-en-Y gastric bypass, sleeve gastrectomy).	Another contemporary surgical comparator or a non-surgical treatment.	Payer perspective.	Systematic review with summaries of several studies.	Systematic review with summaries of several studies.	Not reported	No model.
Picot 2009 [1]	Patients with a BMI of >40 kg/m^2 from the UK	40 years.	Laparoscopic gastric bypass and adjustable gastric banding	Non-surgical comparator (primarily monitoring rather than active treatment)	NHS and Personal Social Services.	20 years.	Direct costs (surgery, complications, hospitalisation, comorbidities)	3.5%	Transition model.

Picot 2012 [97]	Patients with Class I and II obesity (BMI 30-40 kg/m²) with type 2 diabetes or Class I obesity (BMI 30-35 kg/m²) in the UK.	Unclear, presumably between 41.8 and 46.6 years.	Laparoscopic adjustable gastric banding (LAGB).	Non-surgical interventions.	UK National Health Service (NHS).	2, 5, and 20 years.	Direct costs (2009/2010) (surgery, complications, hospitalisation, comorbidities).	3.5%	Transition model.
Pollock 2013 [98]	Obese patients from the UK with type 2 diabetes, 46.5% were males with a mean BMI of 42.4 kg/m² (SD 4.5), Duration of diabetes 1 year (SD 0.33).	Age 46.9 years	Laparoscopic adjustable gastric banding (LAGB).	Standard medical management.	National Health Service.	40 years.	Direct costs (2010) (surgery, complications, dietitian visits, clinical psychology consultations, contacts with GP and outpatient visits) + diabetes costs (medications, complications).	3.5%	Computer model of diabetes.
Salem 2008 [99]	The base-case scenarios included morbidly obese patients from the US without obesity-related comorbidities, with a BMI of	35, 45, and 55 years.	Laparoscopic adjustable gastric banding (LAGB) and laparoscopic Roux-en-Y gastric bypass (LRYGB).	Non-operative weight loss interventions.	Payer perspective.	Lifetime.	Medical costs (2004).	3%	Deterministic decision analytic model.

	40, 50, and 60 kg/m^2.								
Wang 2014 [100]	The reference was defined as US females with BMI of 44 kg/m^2 (simulation range 18-70 years).	53 years.	Laparoscopic gastric bypass (LRYGB), conservative (open) Roux-en-Y gastric bypass (ORYGB), and laparoscopic adjustable gastric banding (LAGB).	Non-surgical intervention.	Health care perspective.	Lifetime.	Direct medical costs (2010) (surgery, complications, follow-up).	3%	Two-part linked model using a deterministic approach for the first 5-year period postsurgery and separate empirical forecasts for the natural history of BMI, costs and outcomes in the remaining years.

4.2.2.8 **Measurement of cost and data sources**

The types of costs included in the studies varied, depending on the perspective of the economic evaluation conducted. Overall, our impression is that the direct medical costs of intervention (gastric bypass, GBP; laparoscopic Roux-en-Y gastric bypass, LRYGB; adjustable gastric banding, AGB; laparoscopic adjustable gastric banding, LAGB) and comparator (conservative therapy, usual care) were extensively reported in the selected publications. This implies a potentially good overview of the cost implications of bariatric surgery from a healthcare perspective. In contrast, concerning indirect costs, only two studies reported specific data (Borg 2014, Michaud 2012). For this reason, it wasn't possible to achieve a comprehensive understanding of the cost-effectiveness of bariatric surgery from a societal perspective.

The accuracy in reporting types of costs and the underlying sources varied across studies. In general, all analyses aimed to include the costs of surgery and related diseases. Only one study reported total lifetime medical costs (i.e., disease-specific and disease-nonspecific costs) (Michaud 2012).

Eleven out of 22 studies explicitly reported the inclusion of in-patient costs (Ackroyd 2006, Ananthapavan 2010, Anselmino 2009, Clegg 2003, Jensen 2005, Keating 2009, McEwen 2010, Michaud 2012, Picot 2009 and 2012, Salem 2008). For all other studies we can assume that inpatient costs were at least partially included in the surgery costs. Only five studies reported the inclusion of costs of preoperative assessment (Ackroyd 2006, Anselmino 2009, Campbell 2010, Clegg 2003, and Picot 2009). However, it is possible that in the other studies such costs were implicitly included in the surgery or inpatient costs. About half of the studies reported costs for physicians or specialist visits (Ackroyd 2006, Ananthapavan 2010, Clegg 2003, Hoerger 2010, Ikramuddin 2009, Jensen 2005, McEwen 2010, Picot 2009 and 2012, Pollock 2013), whereas laboratory test costs were explicitly reported in only three studies (Ackroyd 2006, Jensen 2005, and McEwen 2010). The majority of the studies (17 studies) included costs of complications or comorbidities. Only five studies did not explicitly report the inclusion of either complications or comorbidities (in particular diabetes, coronary heart disease, and stroke) (Ananthapavan 2010, Jensen 2005, McEwen 2010, Michaud 2012, and Padwal 2011). Only one study included costs for parent time and travel costs in the analysis (Ananthapavan 2010). Concerning indirect costs, two studies reported details on productivity loss or deadweight costs (i.e. costs arising from inefficient resource allocation) (Borg 2014, Michaud 2012). Table 10 summarizes the main types of cost included in the selected studies.

For the definition of unit costs, a great variety of sources were used, ranging from results of HTAs to expert opinion, from national registries to international studies. The majority of the included articles reported the use of national registries, national surveys, or national statistics (Ackroyd 2006, Ananthapavan 2010, Anselmino 2009, Borg 2014, Campbell 2010, Clegg 2003, Craig 2002, Ikramuddin 2009, Jensen 2005, Keating 2009, Mäklin 2011, Michaud 2012, Picot 2009 Picot 2012, Pollock 2013, Salem 2008, and Wang 2014). Seven studies clearly assessed surgery or in-patient costs using hospital data (Ananthapavan 2010, Anselmino 2009, Borg 2014, Keating 2009, Mäklin 2011, McEwen 2010, and Picot 2012). Six studies stated the use of expert opinion or interviews (Ackroyd 2006, Campbell 2010, Hoerger 2010, Ikramuddin 2009, Jensen 2005, Picot 2009, and Salem 2008). Four US studies used results of the Framingham Heart Study and Third National Health and Nutrition Examination Survey (Thompson 1999) as sources of unit costs (Craig 2002, Jensen 2005, and Salem 1999). Two studies, Ackroyd 2006 and Anselmino 2009, extracted cost information from the Cost of Diabetes in Europe–type 2 (CODE-2) study to estimate the cost of diabetes care in obese patients (Ackroyd 2006, Anselmino 2009). The paper of Castilla et al. was almost exclusively based on published national studies (Castilla 2014). Two articles extracted cost information from previously published cost-effectiveness analyses (Anselmino 2009 used data from Clegg 2002; Pollock 2013 used data from Salem 2008). In the analysis of McEwen et al. the majority of the cost inputs were directly derived from the paid insurance claims of the 221 included patients (McEwen 2010). In the publication of Faria et al. a long list of possible sources was provided (Faria 2013). However, it was not clear which publication was used for which type of cost. The article by Padwal et al. is a review of several cost-effectiveness analyses in which only the ICERs were reported (Padwal 2011). Table 11 illustrates the types of costs and main sources used in the selected cost-effectiveness analyses.

Table 10 Overview of types of costs included in the eligible cost-effectiveness analyses.

Author and publication year	**Direct medical costs**								**Parent time / travel**	**Indirect costs**
	Preoperative assessment	**Surgery**	**Hospitalisation**	**GP visits**	**Specialist visits**	**Complication / death**	**Comorbidities**	**Laboratory tests**		
Ackroyd 2006	Yes	Yes	Yes	Yes	Yes	Yes		Yes		
Ananthapavan 2010		Yes	Yes	Yes					Yes	
Anselmino 2009	Yes	Yes	Yes			Yes				
Borg 2014		Yes				Yes	Yes			Yes
Campbell 2010	Yes	Yes				Yes				
Castilla 2014		Yes					Yes			
Clegg 2003	Yes	Yes	Yes	Yes	Yes	Yes				
Craig 2002		Yes				Yes	Yes			
Faria 2013		Yes				Yes	Yes			
Hoerger 2010		Yes		Yes	Yes	Yes				
Ikramuddin 2009		Yes		Yes	Yes	Yes				
Jensen 2005		Yes	Yes		Yes			Yes		
Keating 2009		Yes	Yes			Yes	Yes			
Mäklin 2011		Yes				Yes				
McEwen 2010		Yes	Yes	Yes				Yes		
Michaud 2012		Yes	Yes							Yes
Padwal 2011										
Picot 2009	Yes	Yes	Yes	Yes	Yes	Yes	Yes			
Picot 2012		Yes	Yes	Yes	Yes	Yes				
Pollock 2013		Yes		Yes	Yes	Yes				
Salem2008	Yes	Yes	Yes			Yes	Yes			
Wang 2014		Yes				Yes				

Table 11 Types of costs and main sources used in the eligible cost-effectiveness analyses

Article	Type of costs	Sources
Ackroyd 2006	Human resources (surgeons, physicians, nurses, nutritionists), imaging and laboratory tests, operating-room overhead, post-surgical recovery room, hospital stay, consultations, complications, implants, and other not specified factors.	National tariffs, registries, publications and interviews when no other source was available. For diabetes: CODE-2 survey published results.
Ananthapavan 2010	Physician and surgical consultations, laboratory tests, hospitalisation and parent time and travel.	Live, Eat and Play (LEAP) pilot study, Medicare Benefits Schedule, Department of Health and Aging, Australian Bureau of Statistics, Royal Automobile Club of Victoria, private hospital (Australian Refined Diagnostic Related Group).
Anselmino 2009	Hospitalisation (incl. surgery), preoperative assessment, complication, and 5 years follow-up costs.	Lombardy diagnostic-related group (DRG) tariff in Italy, average service-based hospital tariff (Leistungsorientierte Krankenanstaltenfinanzierung, LKF) in Austria, and microcosting estimates from two hospitals in Madrid (Hospital Clinico "San Carlos" and Fundación Hospital Alcorcón) in Spain. Depending on the country, difference sources were used (from national to international literature, including in particular the HTA reports published by the National Institute for Clinical Excellence.[6] For diabetes: CODE-2 survey published results.
Borg 2014	Direct costs: intervention cost of the surgical procedure including any plastic surgery required, cost of adverse events, and excess healthcare costs of treating obesity related diseases. Indirect costs: productivity loss.	One of the major organisations performing GBP in Sweden, official hospital prices in the Southern Healthcare Region in Sweden, previous publications, Statistics Sweden.[101 102]
Campbell 2010	Initial procedure cost, complication/death costs, other medical expenditures, and follow-up costs.	Healthcare Cost and Utilization Project (HCUP) Database with International Classification of Diseases, Ninth Revision codes, the guidelines of the American Society for Metabolic and Bariatric Surgery, input from a clinical expert , the Physician Fee Guide, the Red Book (drug costs), and a study of the Medical Expenditure Panel Survey.[103]
Castilla 2014	Direct costs including surgery and comorbidities.	Several (mainly) national publications.

Clegg 2003	Costs included preoperative assessment (visits, specialist consultations), hospitalisation, complications, and 20-years follow-up.	National Health Services in Scotland Information and Statistics Division. Scottish Health Service Costs 1999/2000.
Craig 2002	Medical costs associated with the initial surgery, treatment of complications, follow-up care, and treatment of obesity-related diseases (e.g. coronary heart disease, stroke, diabetes, hypercholesterolemia, and hypertension).	Medical Care Component of the Consumer Price Index for All Urban Consumers, published literature[104], Healthcare Cost and Utilization Project (HCUPnet, Agency for Healthcare Research and Quality, 2000 Drug Topics Red Book (Montvale, NJ), Data File Documentation of the National Health Interview Survey 1997, National Center for Health Statistics.
Faria 2013	Surgery, complications, and comorbidities.	Unclear. Literature whenever available, otherwise from an institutional database.
Hoerger 2010	Surgery costs, complications, 40 years follow-up (care visits, nutritional supplements, long-term complications).	Medstat claims by Eric A. Finkelstein (2008, unpublished data), publication of Parikh 2006[105], UK Prospective Diabetes Study (UKPDS), or the opinion of an expert panel.
Ikramuddin 2009	Direct costs including surgery, management, and complications.	Agency for Healthcare Quality and Research (Healthcare Cost & Utilization Project, HCUP), MAG Mutual Healthcare Solutions Physicians' Fee and Coding Guide, Drug Topics Redbook, DRG Guidebook, and published literature.[105]
Jensen 2005	Direct costs including operating room, nursing, equipment, anaesthesia, pharmaceutical, and diagnostic test costs. For non-surgical treatment 1-hour weekly weight loss program.	Nguyen et al. 2001[106], Thompson 1999[104], Heshka 2003[107], US Bureau of Labor Statistics.
Keating 2009	Intervention, maintenance, complications, diabetes monitoring/remission, and health care costs to treat diabetes.	Private hospital and private medical specialists, Australian Government Department of Health and Ageing (in particular Australian 2006 Medicare Benefits Schedule and Pharmaceutical Benefits Schedule).
Mäklin 2011	Intervention costs and other average annual healthcare costs including complications	Hospital discharge register and hospital benchmarking database from the National Institute for Health and Welfare. Annual healthcare costs were estimated from the Health

	(no medication and productivity loss).	2000 *Health Examination Survey* data.
McEwen 2010	Total direct medical costs were subdivided into seven mutually exclusive categories including outpatient pharmacy, inpatient, outpatient clinic, diagnostic testing, laboratory testing, emergency room, and a residual category for all other costs.	Paid claims of the 221 included patients, except for outpatient medication costs which were calculated from medications dispensed and average wholesale prices (Red Book: Pharmacy's Fundamental Reference, 2006 Edition).
Michaud 2012	Treatment and medication costs, deadweight, and income changes.	Medical Expenditure Panel Survey (MEPS, prior to age 65), the Medicare Current Beneficiary Survey (MCBS, after age 65), and published literature.[108] [109]
Padwal 2011	Only ICERs were reported.	Ackroyd 2006, Anselmino 2009, Campbell 2010, Clegg 2002[6], Craig 2002, Hoerger 2010, Ikramuddin 2009, Jensen 2005, Keating 2009, McEwen 2010, Picot 2009, and Salem 2008.
Picot 2009	Costs included preoperative assessment (visits, specialist consultations), hospitalisation, complications, and 20 years follow-up.	Clegg 2002[6], published literature, discussion with surgical specialists and a costing developed for Aberdeen specialist obesity services (U. Kulkarni, NHS Grampian, 2008, personal communication).
Picot 2012	Costs of visits, surgery, hospitalisation, specialist consultation, physiotherapy, and complications were included for the LAGB. Out-patient visits and medical management for weigh loss program for usual care.	Finance Department of the Southampton University Hospitals NHS Trust (SUHT), Department of Health (NHS Reference Costs 2006–2007), Unit costs of Health and Social Care.
Pollock 2013	Costs of surgery, complications, diabetes, medication, and visits (physician, dietician, psychologist).	Cost-effectiveness analysis in UK patients with type 2 diabetes (Baudet 2011), NHS Electronic Drug Tariff, Health and Social Care Information Centre, NHS National Tariff using Healthcare Resource Group (HRG) code FZ05B, HTA (Picot 2009), a cost-effectiveness analysis (Salem 2008).

Salem 2008	Usual care medical costs associated with the surgery, including procedural fees, treatment of postoperative complications, follow-up care, and treatment of obesity-related diseases, such as coronary heart disease, stroke, type 2 diabetes, hypercholesterolemia, and hypertension.	Framingham Heart Study and Third National Health and Nutrition Examination Survey [104], estimates of nationally representative hospital charges obtained from the Healthcare Cost and Utilization Project and expert opinion, a cost analysis of the Veterans Administration healthcare system.[110]
Wang 2014	Surgery, complications, and follow-up.	Medicare claims database (2004–2008).

4.2.2.9 **Measurement of clinical effects and data sources**

Clinical effectiveness was generally conceptualised in terms of weight loss or BMI reduction, reduced or delayed comorbidities (e.g. diabetes remission), complications, and mortality reduction. Overall, the included studies provided a good overview of the effectiveness parameters to be considered in the context of bariatric surgery. Nevertheless, it is important to mention that the effects included in the analyses varied strongly across studies and were partially influenced by study population and time horizon. Additionally, it is important to mention that reviewed cost-effectiveness studies used changes in BMI rather than changes in body weight as a percentage change from baseline, and this is a limitation.

Table 12 summarises main effectiveness assumptions, whereas Table 13 provides an overview of the inclusion or exclusion of short- to mid-term effectiveness (i.e. no more than 20 years follow-up), long-term effectiveness (i.e. more than 20 years follow-up; usually lifetime), surgical and post-surgical mortality, complications or adverse events, and changes in diabetes remission. Table 14 illustrates details concerning abdominoplasty, reoperation, and revision.Finally, Table 15 shows the main sources of effectiveness and utility data used in the selected studies.

Almost all studies included short-term effects related to bariatric surgery. In three studies, this was unclear (Faria 2013, Ikramuddin 2009, Jensen 2005).

Long-term effects, usually measured in terms of reduced BMI and reduced mortality, were considered in seven studies (Ananthapavan 2010, Borg 2014, Campbell 2010, Castilla 2014, Craig 2002, Keating 2009, and Michaud 2012) and unclear in four studies (Faria 2013, Hoerger 2010, Jensen 2005, and Salem 2008). In some studies with long-term modelling, the authors decided to assume no more benefit after a certain number of years since surgery. For example, Clegg et al. assumed that after the fifth year after surgery, there were no more benefits (i.e. the BMI in the analyses was reset to the baseline value) (Clegg 2003). This assumption was partially revised in two follow-up analyses by the same group (Picot 2009 and Picot 2012). In Wang et al., BMI changes post procedures were derived from Picot et al. (Wang 2014, Picot 2009). Pollock et al. included the benefit of the surgery only for the first year; thereafter a natural course of risk progression was assumed (Pollock 2013).

Almost all studies included surgery-related complications or adverse events in their analyses. However, it was not always clear which complications were taken in consideration. The most frequent complication, mentioned in 11 studies, was reoperation after surgery (Ackroyd 2006, Campbell 2010, Clegg 2003, Craig 2002, Hoerger 2011, Ikramuddin 2009, Keating 2009, Mäklin 2011, Michaud 2012, Picot 2009, and Picot 2012). Hernia was mentioned in four articles

(Ackroyd 2006, Clegg 2003, Craig 2002, Picot 2009), whereas infections were mentioned in six (Ackroyd 2006, Clegg 2003, Craig 2002, Keating 2009, Picot 2009, Picot 2012). Pulmonary embolism was included in three studies (Ackroyd 2006, Craig 2002, and Picot 2009). Four studies mentioned abdominoplasty as a possible complication (Craig 2002, Hoerger 2010, Mäklin 2011, and Picot 2009). Less frequent complications were mentioned in six articles (Ackroyd 2006, Borg 2014, Campbell 2010, Clegg 2003, Michaud 2012, and Picot 2012). In eight studies the included complications were not specified (Ananthapavan 2010, Anselmino 2009, Faria 2013, Jensen 2005, McEwen 2010, Pollock 2013, Salem 2008, Wang 2013).

Several studies put a special focus on the benefit of surgery in terms of diabetes remission (Ackroyd 2006, Anselmino 2009, Hoerger 2010, Keating 2009, Mäklin 2011, Picot 2012, and Pollock 2013). Other commonly reported comorbidities were coronary heart disease (Castilla 2014, Faria 2013, Hoerger 2010, Ikramuddin 2009, Michaud 2012, Picot 2009, Picot 2012, Pollock 2013, and Salem 2008), stroke (Castilla 2014, Hoerger 2010, Michaud 2012, Picot 2009, Picot 2012, Pollock 2013, and Salem 2008), cancer (Castilla 2014, Faria 2013, Michaud 2012), nephropathies (Faria 2013, Hoerger 2010, Ikramuddin 2009, and Pollock 2013), and ulcer (Ikramuddin 2009, Pollock 2013). Other less common comorbidities were mentioned in three articles (Faria 2013, Ikramuddin 2009, and Mäklin 2011).

The majority of the studies included surgical mortality (i.e. mortality during or up to 30 days after surgery) in the effectiveness assessment. Only five studies excluded it from their analyses (Ackroyd 2006, Anselmino 2009, McEwen 2010, Michaud 2012, and Picot 2012). Twelve studies included post-surgical mortality (i.e. death rates after more than 30 days after surgery) (Borg 2014, Campbell 2010, Castilla 2014, Craig 2002, Hoerger 2010, Ikramuddin 2009, Jensen 2005, Keating 2005, Mäklin 2011, McEwen 2010, Michaud 2012, and Wang 2013).

Concerning conservative treatment, the majority of the studies considered no benefits in terms of BMI reduction (Campbell 2010, Castilla 2014, Clegg 2003, Craig 2002, Mäklin 2011, Picot 2009, Salem 2008, and Wang 2014). Few studies, assumed after few years with little BMI reduction, a natural progression of body weight based on the standard population (Ackroyd 2006, Anselmino 2009, Jensen 2005, Picot 2012, and Pollock 2013). Borg et al. assumed for the usual care group an annual increment of the BMI over time (Borg 2014). For six studies the effects of conservative treatment were not reported or unclear (Ananthapavan 2010, Faria 2013, Hoerger 2010, Ikramuddin 2009, McEwen 2010, and Michaud 2012).

Table 12 Effectiveness assumption for bariatric surgery and conservative treatment in the eligible studies.

	Bariatric surgery strategy assumptions	Conservative treatment strategy assumptions
Ackroyd 2006	Baseline: BMI > 35kg/m². Follow-up years 1-5: BMI reduction ranging from 16.1 to 17.7 kg/m², type 2 diabetes prevalence 50%. Mortality was not included in the model.	Follow-up year 1: BMI reduction by 3 kg/m², 80% type 2 diabetes prevalence. Follow-up years 2-5: no BMI reduction, type 2 diabetes prevalence 100%.
Ananthapavan 2010	Baseline: BMI> 35 kg/m². Mean reduction in BMI at 3 years: 13.9 kg/m². Maintained at 4 years and into the future. Surgical mortality was included in the model	Not reported.
Anselmino 2009	See Ackroyd 2006.	See Ackroyd 2006.
Borg 2014	Baseline: BMI 40-44 kg/m². First year average weight loss of 27%. Thereafter only 75% of the reduction is maintained lifelong. Surgical and post-surgical, BMI dependent mortality were included in the model.	Annual BMI increment over time. The increment is +0.12 kg/m² in patients aged < 45 years, +0.07 kg/m² for age 45 to 65 years, and -0.14 kg/m² for age ≥ 65 years, regardless of gender.
Campbell 2010	Baseline: BMI > 40 kg/m². Cumulative BMI reduction: -19.2% after year 1 and -32.0% after 5 years. Thereafter a constant BMI was assumed lifelong. Surgical and post-surgical, BMI dependent mortality were included in the model.	Patients receiving no treatment were assumed to maintain a constant BMI for the duration of the model.
Castilla 2014	Baseline BMI: 50.7 kg/m². Two years after surgery: 37% BMI reduction. Long-term effects, based on the SOS study: sustained 25% BMI reduction. Surgical and post-surgical, BMI dependent mortality were included in the model.	Patients not operated were considered to remain in the same BMI range their whole lifetime.
Clegg 2003	Baseline: BMI 45 kg/m². 1-5 years after surgery: BMI 29 kg/m². 6-20 years after surgery: BMI = baseline (i.e. No benefit). Surgical mortality was included in the model.	Constant for 20 years.
Craig 2002	Mean percentage reduction of excess weight of about 58% five years after surgery (Excess weight was defined as the weight above a body mass index of 22 kg/m²). After 5 years: with successful surgery (93.7% of the cases) lifetime reduction of BMI. Surgical and post-surgical, comorbidity dependent mortality was included in the model.	Lifetime with initial BMI.
Faria 2013	Not clearly reported. Surgical mortality was included in the calculations.	Not reported.

Hoerger 2010	Baseline: BMI >35 kg/m². Excess weight loss of 63.25% and a BMI loss of 16.17 kg/m² were assumed (time unit was unclear). Diabetes remission rate: 80.3% for persons with newly diagnosed diabetes and 40% for persons with established diabetes. Surgical and post-surgical mortality (based on effects on blood pressure, cholesterol, remission or improvement of diabetes) were included in the model.	Not reported.
Ikramuddin 2009	Not reported in the document. Based on the sensitivity analyses, no assumption with respect to weight gain after LRYGB was made (i.e. the effects of surgery on BMI reduction were maintained constant). Surgical and post-surgical, non-specific mortality were included in the model.	Not reported.
Jensen 2005	Baseline BMI >33 kg/m². Data from the SOS study were used but not reported in detail. Surgical and post-surgical, BMI dependent mortality were included in the model.	Weight loss after 2 years commercial weight loss program: -2.9 [6.5] kg. In case of bad/no compliance with the diet/exercise program, weight gain at a rate similar to that of the overall population was assumed.
Keating 2009	Baseline BMI: 37 kg/m². Based on diabetes remission: 11.4 years over a lifetime. Surgical and post-surgical, diabetes dependent mortality was included in the model.	Based on diabetes remission: 2.1 years over a lifetime.
Mäklin 2011	Baseline BMI: 47 kg/m². Excess weight loss over a 10-years horizon: 60%. Diabetes prevalence reduction: 82%. Surgical and post-surgical, BMI dependent mortality were included in the model.	BMI remain constant, based on the SOS study results.
McEwen 2010	In the first year after surgery: reduction from BMI 51 to 31 kg/m² in women and from BMI 59 to 35 kg/m² in men. Long-term assumptions were not reported. Post-surgical, non-specific mortality was included in the models.	Not reported.
Michaud 2012	Baseline BMI >40 kg/m² (or 35-40 kg/m² with comorbidities). A permanent weight reduction of 25% is achieved. Post-surgical mortality was included in the models.	Not reported.
Padwal 2011	Systematic review (no assumptions).	Systematic review (no assumptions).

Picot 2009	Baseline cohort: BMI >40. 5 years after LRYGB: 36% reduction of initial weight (Clegg 2003). From 5 to 10 years after surgery: 17.7% decline in percentage of weight loss (SOS study). Surgical mortality was included in the model.	Baseline cohort: Stable BMI over time.
Picot 2012	Baseline BMI: 33.5 kg/m^2. Excess weight loss at 2-years follow-up: 62.5% (Dixon 2008 [36]) - 87.2% (O'Brien 2006 [111]). Diabetes remission: 70%. Lifetime: weight reduction until 10 years, thereafter baseline values. Mortality was <u>not</u> included in the model.	Baseline BMI: 33.5 kg/m^2. Excess weight loss at 2-years follow-up: 4.3% (Dixon 2008[36]) - 21.8% (O'Brien 2006 [111]). Diabetes remission: 13%. Lifetime: weight reduction until 10 years, thereafter baseline values.
Pollock 2013	Baseline BMI: 37.1 kg/m^2. In the first year, diabetes remission at 73%. Thereafter natural course of risk progression based on UKPDS and Framingham studies. Surgical mortality was included in the model.	Baseline BMI: 37.1 kg/m^2. In the first year, diabetes remission at 13%. Thereafter natural course of risk progression based on UKPDS and Framingham studies.
Salem 2008	Excess weight loss of 71% in the first 3 years. Lifetime assumption not clearly reported. Surgical mortality was included in the model.	Stable BMI over time.
Wang 2014	Baseline BMI: 44 kg/m^2. BMI changes post procedure were derived from Picot et al. (Picot 2009). Surgical and post-surgical, BMI specific mortality was included in the model.	Baseline BMI: 44 kg/m^2. BMI changes post procedure were derived from Picot et al. (Picot 2009).

Table 13 Overview of the inclusion or exclusion of short-term effectiveness, long-term, effectiveness, mortality (surgical and post-surgical), complications or adverse events, and diabetes remission.

	Short-term effectiveness (≤20 years)	Long-term effectiveness (> 20 years)	Surgical mortality	Post-surgical mortality	Complications or adverse events	Diabetes remission
Ackroyd 2006	Yes	No	No	No	Yes	Yes
Ananthapavan 2010	Yes	Yes	Yes	No	Yes	No
Anselmino 2009	Yes	No	No	No	Yes	Yes
Borg 2014	Yes	Yes	Yes	Yes	Yes	No
Campbell 2010	Yes	Yes	Yes	Yes	Yes	No
Castilla 2014	Yes	Yes	Yes	Yes	No	No
Clegg 2003	Yes	No	Yes	No	Yes	No
Craig 2002	Yes	Yes	Yes	Yes	Yes	No
Faria 2013	Unclear	Unclear	Yes	No	Yes	Unclear
Hoerger 2010	Yes	Unclear	Yes	Yes	Yes	Yes

Ikramuddin 2009	Unclear	No	Yes	Yes	Yes	No
Jensen 2005	Unclear	Unclear	Yes	Yes	Yes	No
Keating 2009	Yes	Yes	Yes	Yes	Yes	Yes
Mäklin 2011	Yes	No	Yes	Yes	Yes	Yes
McEwen 2010	Yes	No	No	Yes	Yes	No
Michaud 2012	Yes	Yes	No	Yes	Yes	Yes
Padwal 2011	-	-	-	-	-	-
Picot 2009	Yes	No	Yes	No	Yes	Yes
Picot 2012	Yes	No	No	No	Yes	Yes
Pollock 2013	Yes	No	Yes	No	Yes	Yes
Salem 2008	Yes	Unclear	Yes	No	Yes	No
Wang 2013	Yes	No	Yes	Yes	Yes	No

Table 14 provides additional detail about plastic surgery, reoperations and revisions. Plastic surgery in the included literature was usually called abdominoplasty (other terms like reconstructive surgery, body contouring, body lifting, lifting, lift were not used) and was explicitly mentioned in six studies (Borg 2014, Craig 2002, Hoerger 2010, Mäklin 2011, Picot 2009, Salem 2008). Reoperations or revisions were mentioned in twelve articles (Ackroyd 2006, Campbell 2010, Clegg 2003, Craig 2002, Hoerger 2010, Ikramuddin 2009, Keating 2009, Mäklin 2011, Michaud 2012, Picot 2009, Picot 2012, Salem 2008). In general, costs associated with abdominoplasty, reoperation or revision was included in the complication costs or in the total costs. Since complications were not always described in detail in all articles, it may be possible that abdominoplasty, reoperations or revisions were included even if not mentioned.
Scrutiny of the cost-effectiveness results did not reveal large ICER differences between studies explicitly including plastic surgery versus others. Moreover, a couple of studies reported in their sensitivity analyses that the frequency or costs of abdominoplasty, reoperation and surgery had little impact on the cost-effectiveness of bariatric surgery (e.g. Salem 2008, Hoerger 2010).

Table 14 Overview of abdominoplasty, reoperation, and revision.

Author and publication year	Abdominoplasty	Reoperation	Revision	Comments
Ackroyd 2006	No	Yes	Yes	Reoperation and revision costs included in the complications costs
Ananthapavan 2010	No	No	No	
Anselmino 2009	No	No	No	
Borg 2014	Yes	No	No	Plastic surgery costs included in the intervention costs
Campbell 2010	No	Yes	No	Reoperation costs included in the complication costs
Castilla 2014	No	No	No	
Clegg 2003	No	Yes	Yes	Reoperation and revision costs included in the complications costs
Craig 2002	Yes	Yes	Yes	Abdominoplasty, reoperation, and revision costs included in total costs
Faria 2013	No	No	No	
Hoerger 2010	Yes	No	Yes	Revision and abdominoplasty costs included in follow-up costs
Ikramuddin 2009	No	Yes	No	Reoperation costs included in total costs
Jensen 2005	No	No	No	
Keating 2009	No	Yes	No	Reoperation costs included in complication costs
Mäklin 2011	Yes	Yes	No	Abdominoplasty and reoperation costs included in total costs
McEwen 2010	No	No	No	
Michaud 2012	No	Yes	No	Reoperation costs included in total costs
Padwal 2011	-	-	-	Review
Picot 2009	Yes	Yes	Yes	Abdominoplasty, reoperation, and revision costs included in total costs
Picot 2012	No	No	Yes	Revision costs included in total costs
Pollock 2013	No	No	No	
Salem 2008	Yes	Yes	Yes	Abdominoplasty, reoperation, and revision costs included in total costs
Wang 2014	No	No	No	

The sources of effectiveness estimates varied across studies. However, some sources were used more frequently than others. For example, the results of the Swedish Obese Subjects (SOS) study[56 112-114], a prospective, long-term controlled intervention study including 2,010 obese subjects who underwent bariatric surgery and 2037 matched obese control subjects receiving usual care was used by eleven studies as their main source for effectiveness estimates (Borg 2014, Castilla 2014, Clegg 2003, Craig 2002, Faria 2013, Hoerger 2010, Jensen 2005, Mäklin 2011, Michaud 2012, Padwal 2011, and Picot 2009). Six studies (Campbell 2010, Keating 2009, Michaud 2012, Picot 2009, Picot 2012, and Pollock 2013) extracted effectiveness estimates from Australian studies[36 43 111 115]. Two studies (Hoerger 2010 and Mäklin 2011) used effectiveness information from a meta-analysis published in 2009.[116] Several studies directly extracted effectiveness data from previously published cost-effectiveness analyses (Anselmino 2009, Clegg 2003, Faria 2013, Michaud 2012, Padwal 2011, Picot 2009, and Wang 2014).

In the studies by Ananthapavan et al. and McEwen et al., the effectiveness of the intervention was directly based on trial data (Ananthapavan 2010 and McEwen 2010). In the analysis by Ackroyd et al., effectiveness inputs were provided by several national institutions (i.e. the National Institute for Clinical Excellence (NICE), the "Agence Nationale d'Accréditation et d'Evaluation en Santé" (ANAES), the Australian Safety and Efficacy Register of New Interventional Procedures (ASERNIP-S), the Swedish Council on Technology Assessment in Health Care (SBU) and the "Deutsche Adipositas Gesellschaft" (DGA)) (Ackroyd 2006). Clegg et al. used effectiveness information from a relatively old study published by Pories et al.[6 117] Ikramuddin et al. used unpublished data from the Minnesota cohort, a prospective observational study conducted at an academic medical centre in the United States (Ikramuddin 2009). Data from the Framingham Heart Study were used as basis in the model by Salem et al.[99 104]

It should be noted that only two articles included in the systematic review performed in the first part of this report were used in the cost-effectiveness studies.[36 111] The reason for difference is the publication year. The majority of the RCTs investigating the effectiveness of bariatric surgery were published in the last two years. Therefore, they were not yet available for the above mentioned cost-effectiveness analyses.

Table 15 Effectiveness and utility main sources used in the selected cost-effectiveness studies.

Article	Sources of effectiveness estimates	Sources of utility estimates
Ackroyd 2006	National institutes or registries: National Institute for Clinical Excellence (NICE), the "Agence Nationale d'Accréditation et d'Evaluation en Santé" (ANAES), the Australian Safety and Efficacy Register of New Interventional Procedures (ASERNIP-S), the Swedish Council on Technology Assessment in Health Care (SBU) and the "Deutsche Adipositas Gesellschaft" DGA.	Health Outcomes Data Repository (HODaR) Cardiff Research Consortium .
Ananthapavan 2010	Audit of in-patient and out-patient records of 28 adolescents (2 boys and 26 girls, mean age 18 years, range 15–19), who underwent LAGB at the Centre for Bariatric Surgery, Avenue Hospital (Windsor, VIC, Australia).	Haby2006.[118]
Anselmino 2009	Ackroyd 2006.	Unclear.
Borg 2014	The Swedish Obese Subjects (SOS) trial - a prospective controlled intervention study (Sjöstrom 2013).	Health Outcomes Data Repository (HODaR) Cardiff Research Consortium (Ackroyd 2006).
Campbell 2010	Angrisani 2007[119], O'Brien 2006.[111]	EQ-5D data from the 2000 Medical Expenditure Panel Survey; utilities reported for laparoscopic surgery for hernia repair in the United Kingdom (McCormack 2005).[120]
Castilla 2014	Spanish study of Mar2011 for short-term. The Swedish Obese Subjects (SOS) study for long-term.[56]	National publications (in particular 2011).
Clegg 2003	HTA (Clegg 2002) including 17 RCT and 1 cohort study (the Swedish Obese Subjects (SOS) study).[6]	Economic evaluation of orlistat (Hakim 2002).[121]
Craig 2002	Pories 1995.[117]	Framingham Heart Study.[104]
Faria 2013	Literature whenever available (e.g. Keating 2009, McEwen 2010, Sjöstrom 2012, Craig 2002, Padwal 2011, Pollock 2011) or retrieved from their institutional database.	Unclear. Literature whenever available (e.g. Keating 2009, McEwen 2010, Sjöstrom 2012, Craig 2002, Padwal 2011, Pollock 2011) or retrieved from their institutional database.
Hoerger 2010	Several publications, in particular a meta-analysis (Buchwald 2009) and the Swedish Obese Subjects (SOS) study (Sjöstrom 2004).	Average of five different sources.

Ikramuddin 2009	A prospective observational study conducted at an academic medical center in the United States (Minnesota cohort; unpublished data, University of Minnesota Medical Center, Minneapolis).	Cost of Diabetes in Europe–type 2 (CODE-2) study.[122]
Jensen 2005	Swedish Obese Subjects (SOS) Study.[112]	Economic evaluation of orlistat.[121]
Keating 2009	2-year randomized controlled trial involving 60 obese participants (BMI 30-40 kg/m^2) in Australia.[36]	DiabCo$t study.[123]
Mäklin 2011	Several publications, in particular a meta-analysis[116] (Buchwald et al. 2004 and 2009) and the Swedish Obese Subjects (SOS) study.[114]	Randomized trials and 15D utilities from the Health 2000 Health Examination survey.
McEwen 2010	221 patients included in the study.	221 patients included in the study.
Michaud 2012	A randomised controlled trial[36] (Dixon 2008), a HTA[1], and the Swedish Obese Subjects (SOS) study.[113]	Unclear. Probably Dixon 2008 (a randomised controlled trial), Picot 2009 (a meta-analysis)[1], or Sjöstrom 2004.[113]
Padwal 2011	Ackroyd 2006, Anselmino 2009, Campbell 2010, Clegg 2002, Craig 2002, Hoerger 2010, Ikramuddin 2009, Jensen 2005, Keating 2009, McEwen 2010, Picot 2009, Salem 2008.	Ackroyd 2006, Anselmino 2009, Campbell 2010, Clegg 2002, Craig 2002, Hoerger 2010, Ikramuddin 2009, Jensen 2005, Keating 2009, McEwen 2010, Picot 2009, Salem 2008.
Picot 2009	HTA[6], Swedish Obese Subjects (SOS) study[113], Angrisani 2007[119], Australian studies.[36 111]	Economic evaluation of orlistat[121] , Australian study[115], Currie 2006[124], Lee 2005.[125]
Picot 2012	Australian studies.[36 111]	Economic evaluation of orlistat.[121]
Pollock 2013	Australian study.[36]	UK Prospective Diabetes Study (UKPDS), Cost of Diabetes in Europe–type 2 (CODE-2) study.[122]
Salem 2008	Framingham Heart Study.[104]	Framingham Heart Study[104] and Craig 2002.
Wang 2014	HTA.[1]	Published studies.[126-128]

4.2.2.10 **Measurement of health-related quality of life and data sources**

It should be noted that the measures of health-related quality of life considered in the systematic review part of this HTA report are different from those needed for health economic evaluation. For example, in the systematic review part, quality of life is reported in terms of different dimensions / scales of the SF-36 instrument (e.g. representing physical or mental health). The SF-36 and other ‘conventional’ instruments are typically designed to study quality of life and its components from a clinical perspective, with an interest in relatively subtle changes. In contrast, cost-effectiveness analyses require quality of life measured in the form of utility weights. Utilities are values that reflect an individual's preferences for different health states. They provide an overall summary measure on an interval scale between zero (representing death) and one (representing perfect health). In the literature included in the health economic part of the assessment, only the HTA report by Picot et al. provided quality of life information in terms of SF-36 scores (Picot 2009). These scores were based on the results of the Australian studies published by O'Brien et al. and Dixon et al. (O'Brien 2006, Dixon 2007). The mean SF-36 scores concerning physical function and physical role, reported after two years of follow up, were significantly higher in patients undergoing laparoscopic gastric banding if compared with non-surgical care (mean difference 12.5, $p<0.05$). In contrast, no significant difference was found concerning mental health (difference 4.0, $p>0.05$). These results are consistent with those reported in the clinical section of this report (see Figure 15) and suggest that the utility increase after surgery may be driven by physical changes.

For the measurement of quality-adjusted lifetime, almost all studies combined survival estimates with utilities to generate QALYs as their main summary measure of effectiveness. Only one study assessed disability-adjusted life years (Ananthapavan 2010). One study did not consider quality of life but only duration of survival (Michaud 2012). In Padwal et al., a review of several cost-effectiveness analyses, only ICERs were provided (Padwal 2011).

The sources of utility estimates varied across studies. In four studies, utility information was extracted from previously published cost-effectiveness analyses (Borg 2014, Faria 2013, Michaud 2012, and Padwal 2011). Two studies (Craig 2002 and Salem 2008) used utility data from the Framingham Heart Study (Thompson 1999), whereas two other studies (Ackroyd 2006 and Borg 2014) applied data from the Health Outcomes Data Repository (HODaR) of the Cardiff Research Consortium (Currie 2005). Four studies (Clegg 2003, Jensen 2005, Picot 2009, and Picot 2012) used utility information from an economic evaluation of orlistat, a drug commonly used to treat obesity (Hakim 2002). Two studies (Ikramuddin 2009 and Pollock 2013) extracted utility values from the CODE-2 study.[122] Other cost-effectiveness analyses used utility values from national surveys or publications (Ananthapavan 2010, Campbell 2010, Castilla 2014, Hoerger 2010, Keating 2009, Mäklin 2011, and Wang 2014). McEwen et al. used utility data measured in their study population (McEwen 2010). The utility sources in Anselmino et al., were not provided (Anselmino 2009). As already mentioned, a summary of the sources of utility estimates is provided in section (4.2.2.10 Measurement of health-related quality of life and data sources). More detailed information on health-related quality of life and data sources used for the selected studies is available in section Summary of studies.

4.2.2.11 **Sensitivity analyses within cost-effectiveness studies**

There was a general consistency between studies that relevant parameters with underlying uncertainty were varied, particularly with regard to treatment efficacy, rates of complications, costs, and utilities. It should be noted that 15 out of 21 cost-effectiveness studies did not undertake PSA, and this was not justified. (The 22nd study by Padwal et al. was a systematic review and there were no sensitivity analyses expected (Padwal 2011).)

- Deterministic sensitivity analyses were performed in six out of 22 assessed studies (Hoerger 2010, Mäklin 2011, Picot 2009 and 2012, Salem 2008 and Wang 2014). The values of key input parameters were altered one at a time, while maintaining all other input parameters at base-case values. Ranges of variation were selected based on upper and lower limits of 95% confidence intervals, or by choosing published or arbitrary values.

- Scenario analyses were performed in 15 out of 22 assessed studies (Ackroyd 2006, Anselmino 2009, Borg 2014, Campbell 2010, Castilla 2014, Clegg 2003, Craig 2002, Ikramuddin 2009, Jensen 2005, Keating 2009, Mäklin 2011, McEwen 2010, Pollock 2013, Picot 2009 and Wang 2014). Studies (Anselmino 2009, Borg 2014, Campbell 2010, Clegg

2003, Ikramuddin 2009, Keating 2009, Jensen 2005, Mäklin 2011, McEwen 2010, Pollock 2013, Picot 2009 and Wang 2014) tested parameters including, e.g., hospitalization costs, increased or reduced effectiveness of surgery compared to conservative treatment, discount rates, BMI categories, rates of complications with surgery, reduced costs of conservative therapy.

Ackroyd et al. used two-way sensitivity analyses to vary BMI reduction and type 2 diabetes prevalence by 20% in each arm of the model. Similarly, Craig et al. used two-way sensitivity analyses varying both lifetime medical cost and discount rate (Craig 2002).

Based on the reported sets of deterministic, scenario and two-way sensitivity analyses, the results in the identified economic evaluations of bariatric surgeries appeared to be reasonably robust.

- PSA was used in seven out of 22 assessed studies (Ananthapavan 2010, Campbell 2010, Castilla 2014, Faria 2013, Picot 2009 and 2012, and Pollock 2013). Ikramuddin et al. used non-parametric bootstrapping to explore uncertainty for clinical and costs parameters (Ikramuddin 2009). Probabilistic sensitivity analyses showed that the results were robust.

Overall, the range of sensitivity analyses performed within the studies showed that the derived CE results were robust, and results still remain cost-effective.

4.2.3 Adaptation of economic evaluations results to Switzerland and differences observed between studies

The original results of the 22 included cost-effectiveness studies are summarised in Table 16 and Table 17. Of these 22 studies, 15 were found to be qualitatively transferable to Switzerland and provided sufficient information on costs and effects. The results of these 15 studies were adapted to Switzerland, and are shown in Table 18, Table 19 and Figure 20. Figure 20 graphically illustrates differences in costs and effects between surgery and conservative treatment. The results of the 15 studies are quite consistent. The vast majority of studies indicate bariatric surgery compared to conservative treatment or non-surgical management appears to be a cost saving (dominant) or cost-effective strategy according to the criteria set by the authors. Adapted ICERs were generally below CHF 50,000 per QALY gained.

Included studies were reporting different perspective, for example three out of 15 studies were form a healthcare perspective (Keating 2009, Mäklin 2011, and Wang 2014), ten were from a payer perspective (Ackroyd 2006, Campbell 2010, Castilla 2014, Clegg 2003, Craig 2002, Ikramuddin 2009, Padwal 2011, Picot 2009, Picot 2012 and Pollock 2013). Borg et al and Michaud et al were from a societal perspective (Borg 2014, Michaud 2012).

All studies of patients with BMI values >35 kg/m^2 indicated bariatric surgery to be cost saving, or cost-effective based on the criteria set by the authors. Adapted ICERs for Switzerland were below CHF 50,000 per QALY gained. More specifically, four studies with long-time horizons (10 years to lifetime) showed cost saving results (Table 9) (Borg 2014, Castilla 2014, Keating 2009 and Mäklin 2011). Even a study with a short time horizon (5 years), by Ackroyd et al., indicated cost savings for Germany and France (Ackroyd 2006). (The same study indicated a cost-effective ICER for the UK.). Other studies (Campbell 2010, Clegg 2003, Craig 2002, Hoerger 2010, Ikramuddin 2009, Picot 2009, Picot 2012, Pollock 2013, and Wang 2014) showed cost-effective results, based on the cost-effectiveness threshold applied by the authors. Adapted ICERs for Switzerland ranged from CHF 8,000 to CHF 44,000 ICER per QALY gained (Table 18) (Campbell 2010, Clegg 2003, Craig 2002, Hoerger 2010, Ikramuddin 2009, Picot 2009, Picot 2012, Pollock 2013, and Wang 2014). Ignoring the study by Clegg et al. (Clegg 2003), which can be regarded as replaced by Picot et al. (Picot 2009) from the same group of authors, reduced the ICER range to below CHF 32,000 per QALY gained (Craig 2002). In the study by Michaud et al., where effectiveness outcome of interest was life years gained rather than QALYs gained, the adapted cost-effectiveness ratio was below CHF 9,000 per life year gained.

When studies like Borg et al. adopted a societal perspective, adapted ICERs for females aged 45-54 with BMI 40-44 kg/m^2 were below CHF 3,000 per QALY gained when gastric bypass was compared to conservative treatment. For males aged 45-54 with BMI 40-44 kg/m^2, gastric bypass was cost-saving and thus dominant. Michaud et al. also reported cost-effectiveness from a societal perspective, for US patients with BMI > 40 kg/m^2 or with BMI > 35 kg/m^2 and a high risk of comorbidities (Michaud 2012). The adapted ICER for gastric bypass compared to conservative treatment was CHF 8,158 per life year gained.

For patients in the BMI category <35 kg/m^2, the cost-effectiveness of bariatric surgery compared to conservative treatment was only assessed by Borg et al. and Picot et al.; both applied a lifetime horizon (Borg 2014 and Picot 2012). In the study by Borg et al., the adapted ICER was below CHF 3,000 per QALY gained. In the study by Picot et al, the adapted ICER was below CHF 50,000 per QALY gained.

Similarly, from the societal perspective, for patients in the BMI class <35 kg/m^2, the cost-effectiveness of bariatric surgery compared to conservative treatment was only assessed by Borg et al. (Borg 2014). In the study by Borg et al., the adapted ICER from the societal perspective was CHF 10,458 per QALY gained for females and CHF 12,365 per QALY gained for males. For extended eligibility (BMI >35 or BMI >30 kg/m^2 with qualifying comorbidities), the adopted ICER for gastric bypass compared to conservative treatment was CHF 10,502 per life year gained.

In order to understand the observed differences between studies, we assessed the following potential explanatory factors:

- Age of subjects included in the economic evaluation
- Gender
- BMI class
- Type of surgery
- Time horizon and discounting
- Costs
- Modelling of effectiveness, mortality, clinical events, life years gained and utility

In the assessed studies, there was no consistent relationship between age and the incremental cost-effectiveness of bariatric surgery. Younger patients may have fewer early events but an expected to have longer survival in comparison with older patients. Consistent with this notion, Wang et al. reported a U-shaped relationship (Wang 2014: Figure 3). ICERs reached their minimum at about age 55-60 years. In contrast, studies like those by Borg et al. and Craig et al. showed ICERs per QALY gained that increased with the patients' age.

There was also no consistent trend across studies regarding a dependency of incremental cost-effectiveness on gender. For example, Borg et al. reported bariatric surgery to be less cost saving in women than in men. The study by Campbell et al. yielded similar results for men and women, expect in the BMI category >50 kg/m^2, where the incremental cost associated with surgery was better for women. Craig et al. showed less favourable results for men than for women (Table 18 and Table 19).

However, there was consistency across studies in the sense that bariatric surgery compared to conservative treatment displayed better cost-effectiveness for patients with higher BMI >35 kg/m^2 than for patients with lower BMI <35 kg/m^2 (Borg 2014, Campbell 2010, Craig 2002, and Picot 2012) (Table 18 and Table 19).

There were different types of bariatric surgeries performed in the reviewed studies. The most common ones were gastric bypass (LRYG), and other surgeries were, ORYGB, gastric banding, adjustable gastric banding and laparoscopic adjustable gastric banding (LAGB). Cost-effectiveness results for gastric bypass compared to conservative treatment ranged from a cost saving (dominant) situation to CHF 43,480 per QALY gained (Clegg 2003). Cost-effectiveness results for alternative surgical approaches provided within the same study showed that gastric bypass was better than gastric banding in terms of benefits (QALYs gained, life years gained) (Table 18 and Table 19), but also more expensive (Ackroyd 2006, Hoerger 2010, Mäklin 2011).

Other surgeries, such as open gastric bypass in the study by Craig et al., the cost-effectiveness was better than CHF 32,000 per QALY gained (Craig 2002). On the other hand, for gastric banding compared to conservative treatment, ICERs ranged from cost saving (dominant) to CHF 18,000 per QALY gained (Table 18, Table 19, and Figure 20).

Few studies reported results for different types of gastric bypass surgery. In the study by Wang et al., ORYGB compared to conservative treatment had an ICER of CHF 18,772 per QALY gained, whereas LRYG compared to conservative treatment the ICER was CHF 7,199 per QALY gained. The authors also compared LAGB with conservative treatment, and showed an ICER per QALY gained of CHF 6,757 (Figure 20, Table 18 and Table 19).

There was a consistency between studies with regards to time horizon, i.e. the higher the time horizon, the higher QALY difference between surgical interventions and conservative treatment. Furthermore, most of the studies used 3% or 3.5% discounting for effects, therefore discounting had little impact on the differences between the observed results.

If we assess why some of the studies found cost savings and others found cost-effective results (including studies with long time horizon and BMI > 35 kg/m^2), part of the explanation may lie in the cost items taken in consideration, and other assumptions with regards to which effects of bariatric surgery were modelled, how, and for how long.

For example; in terms of absolute costs, some studies assessed disease specific versus total health care costs. Michaud et al. assessed total health care costs as opposed to condition specific costs, therefore absolute costs in this study were much higher than observed in the other reviewed studies.

Some studies found a high net-difference in costs between bariatric surgery and conservative treatment. Even though there was not a clear pattern between studies, such differences may be because in some studies the estimated costs of conservative treatment were very low and were not accounting for all potential, obesity-related costs during follow-up. This might potentially lead to a high difference in costs observed and therefore a higher ICER per QALY gained. In the study by Ackroyd et al., for the UK, costs for conservative therapy were much lower than for other countries such as Germany and France. Therefore, the ICER per QALY gained was higher.

QALY estimates were another major driver of the cost-effectiveness differences observed between studies. For example, five studies found a low net QALY difference (bariatric surgery versus conservative treatment) of below 1.0 (Picot 2012, Clegg 2003, Ikramuddin 2009, Mäklin 2011, and Pollock 2013). Picot et al. studied a low BMI population. The studies by Ikramuddin et al. and Pollock et al. were both based on the same diabetes model and used a low, regression-estimated utility effect per one unit change in BMI (0.004). In the study by Mäklin et al., utility was based on a population health survey and subsequent regression analysis. There was no mortality difference modelled and the time horizon of the analysis was restricted to 10 years.

There were four studies that reported high QALY differences as compared to other studies, i.e. the differences observed were > 2.5 QALYs (Borg 2014, Castilla 2014, Campbell 2010, and Wang 2014). The study with the highest QALY difference was Castilla et al. The reporting of this study was insufficient, which may have made methodological problems. In particular, event related utilities appeared to be very high and it was not clear for how long these utilities were applied in the model. In all other cases (Borg 2014, Campbell 2010, and Wang 2014) the high QALY difference was at least partially driven by a modelling of condition-related mortality, leading to differences between the bariatric surgery versus conservative management strategies. Some other assumptions, for instance the duration of the BMI effect of bariatric surgery, may also have been optimistic. In the study by Campbell et al., the use of a more conservative set of related assumptions in an alternative analysis led to a lower QALY difference.

All other studies reported QALY differences between 0.9 and 2.0 (Ackroyd 2006, Craig 2002, Hoerger 2010, for patients with established diabetes, and Keating 2009, Picot 2009). In some cases, modelling of mortality was undertaken, but it may have been counteracted by other conservative assumptions (for example, on effect size or duration of BMI change). Keating et al. found a net difference of 1.2 in QALY, which is closer to the lower end of the range of net difference found within this group. A reason for this may be that this study was focused on diabetes and therefore may not have covered all relevant aspects with regards to patient mortality. The study by Michaud et al., did not report QALYs but life expectancy. The model used was very different from those used in the other studies, hindering comparison. However, the survival differences observed were consistent with the QALY differences observed by the other studies.

Some studies assumed a re-increase of BMI after a certain time and then a stable difference after the observation period. However, descriptions were not always entirely clear in the study. If the difference in BMI was maintained for longer, this tended to translate into higher net differences in terms of life years lived and QALYs gained for surgery.

The reviewed studies modelled condition-related mortality to varying degrees, e.g. only due to diabetes; only due to cardiovascular diseases; both combined; or due to overweight in general. The modelled mortality differences were observed in the SOS prospective observational study and may be real. However, they were not observed in randomised clinical trials. Generally, randomised trials did not have long enough follow-up periods.

Another important observation was that assumptions made with regard to duration of diabetes remission was very diverse, e.g. remission for 2 years (Picot 2009) versus annual probability of relapse of 8.3% (Hoerger 2010; SOS-based). Some studies strongly focused their cost-effectiveness models on the effects on diabetes. Except in the case of Hoerger et al., these studies (Ikramuddin 2009, Keating 2009, Pollock 2013) found relatively small QALY differences. The explanation for this may be that focus on diabetes in these studies may have led to potentially incomplete coverage of other effects, which would have yielded lower QALYs.

Table 16 Results of economic evaluations for laparoscopic gastric bypass, as originally reported by the authors

Study	Total Costs of I	Total costs of C	Outcome of I (QALY)	Outcome of C (QALY)	ICER per QALY
	(Difference)		(Difference)		
Ackroyd 2006 - *Gastric bypass versus conservative therapy* Germany	€ 12,166	€ 17,197	3.34	2.0	* Bariatric surgery dominant
	(€ -5031)		(1.34)		
France	€ 13,399	€ 19,276	3.34	2.0	Bariatric surgery dominant
	(€ -5,877)		(1.34)		
United Kingdom	€ 9,121	€7,083	3.34	2.0	Bariatric surgery dominant
	(€ 2038)		(1.34)		

Borg 2014 - *Gastric bypass versus conservative therapy* Male (Age 45-54) **BMI 30-34 kg/m²**	SEK 63,143	SEK 41,795	14.64	11.68	SEK 7,212
	(SEK 21,348)		(2.96)		
Male (Age 45-54) **BMI 35-39 kg/m²**	SEK 87,422	SEK 196,141	13.51	10.17	Bariatric surgery dominant
	(SEK -108,719)		(3.34)		
Male (Age 45-54) **BMI 40-44 kg/m²**	SEK 146,381	SEK 469,978	11.91	8.43	Bariatric surgery dominant
	(SEK -323,597)		(3.48)		
Male (Age 45-54) **BMI 45-49 kg/m²**	SEK 297,941	SEK 888,649	10.6	7.17	Bariatric surgery dominant
	(SEK -590,708)		(3.43)		
Male (**Age 25–34**) BMI 40-44**kg/m²**	SEK 184,961	SEK 568,371	14.84	10.97	Bariatric surgery dominant
	(SEK -383,410)		(3.87)		
Male (**Age 35–44**) BMI 40-44 **kg/m²**	SEK 163,096	SEK 516,825	13.74	9.94	Bariatric surgery dominant
	(SEK -353,729)		(3.8)		
Male (**Age 45–54**) BMI 40-44 **kg/m²**	SEK 146,381	SEK 469,978	11.91	8.43	Bariatric surgery dominant
	(SEK -323,597)		(3.48)		
Male (**Age 55 - 64**) BMI 40-44 **kg/m²**	SEK 131,629	SEK 357,771	9.6	6.59	Bariatric surgery dominant
	(SEK -226,142)		(3.01)		
Male (**Age 65–74**) BMI 40-44 **kg/m²**	SEK 112,050	SEK 226,819	6.78	4.54	SEK 89,958
	(SEK -114,769)		(2.24)		
Female (Age 45-54) **BMI 30–34 kg/m²**	SEK 71,198	SEK 39,063	15.53	12.62	SEK 11,043
	(SEK 32,135)		(2.91)		
Female (Age 45-54) **BMI 35–39 kg/m²**	SEK 95,196	SEK 119,142	14.29	10.95	Bariatric surgery dominant
	(SEK -23,946)		(3.34)		
Female (Age 45-54) **BMI 40–44 kg/m²**	SEK 126,427	SEK 274,136	13.00	9.35	Bariatric surgery dominant
	(SEK -147,709)		(3.65)		
Female (Age 45-54) **BMI 45–49 kg/m²**	SEK 204,017	SEK 50,018	11.46	7.97	Bariatric surgery dominant
	(SEK -296,101)		(3.49)		
Female (**Age 25–34**) BMI 40-44 **kg/m²**	SEK 145,526	SEK 329,448	15.27	11.4	Bariatric surgery dominant
	(SEK -183,922)		(3.87)		
Female (**Age 35–44**) BMI 40-44 **kg/m²**	SEK 135,213	SEK 304,438	14.55	10.72	Bariatric surgery dominant
	(SEK -169,226)		(3.83)		
Female (**Age 45–54**) BMI 40-44 **kg/m²**	SEK 126,427	SEK 274,136	13.00	9.35	Bariatric surgery dominant
	(SEK -147,709)		(3.65)		

Female (**Age 55–64**) BMI 40-44 **kg/m²**	SEK 116,142	SEK 211,766	10.75	7.57	Bariatric surgery dominant
	(SEK -95,624)		(3.18)		
Female (**Age 65–74**) BMI 40-44 **kg/m²**	SEK 102,357	SEK 138,453	7.93	5.35	Bariatric surgery dominant
	(SEK -36,096)		(2.58)		
Campbell 2010 - ***Laparoscopic Roux-en-Y gastric bypass versus conservative therapy*** Aggregate population (Base case - 40 years) – Angrisani et al.	$ 124,811	$ 108,523	19.054	16.55	$ 5,618
	($ 16,288)		(2.9)		
Aggregate population (Base case - 40 years) – O'Brien et al.	$ 129,442	$ 108,523	18.56	16.155	$ 8,698
	($ 20,919)		(2.4)		
Males (Angrisani et al.) **BMI 35-39.9 kg/m²**	$ 117,087	$ 87,943	18.431	16.38	$ 14,210
	($ 29,144)		(2.05)		
Males (Angrisani et al.) **BMI 40-49.9 kg/m²**	$ 120,594	$ 101,778	17.966	14.805	$ 5,953
	($ 18,816)		(3.16)		
Males (Angrisani et al.) **BMI > 50 kg/m²**	$ 122,712	$117,284	17.682	12.835	$ 1,120
	($ 5,428)		(4.85)		
Males (O'Brien et al.) **BMI 35-39.9 kg/m²**	$ 117,776	$ 87,943	18.335	16.38	$ 15,260
	($ 59,833)		(1.96)		
Males (O'Brien et al.) **BMI 40-49.9 kg/m²**	$ 124,687	$ 101,778	17.421	14.805	$ 8,757
	($ 22,909)		(2.62)		
Males (O'Brien et al.) **BMI > 50 kg/m²**	$ 131,959	$ 117,284	16.435	12.835	$ 4,076
	($ 14675)		(3.60)		
Females (Angrisani et al.) **BMI 35-39.9 kg/m²**	$ 122,592	$ 95,334	19.662	17.756	$ 14,301
	($ 27,258)		(1.91)		
Females (Angrisani et al.) **BMI 40-49.9 kg/m²**	$ 126,667	$ 112,316	19.238	16.338	$ 4,952
	($ 14,361)		(2.90)		
Females (Angrisani et al.) **BMI > 50 kg/m²**	$ 129,148	$ 132,033	18.979	14.449	Bariatric surgery dominant
	($ -2,885)		(4.53)		
Females (O'Brien et al.) **BMI 35-39.9 kg/m²**	$ 123,433	$ 95,334	19.567	17.756	$ 15,516
	($ 28,099)		(1.81)		
Females (O'Brien et al.) **BMI 40-49.9 kg/m²**	$ 131,758	$ 122,316	17.728	16.338	$ 13,987
	($19,442)		(1.39)		
Females (O'Brien et al.) **BMI > 50 kg/m²**	$ 140,683	$ 132,033	17.804	14.449	$ 2,578
	($ 8650)		(3.36)		

Castilla 2014 - ***Gastric bypass versus conservative therapy***	€ 17,431	€ 31,425	18.18	12.55	Bariatric surgery dominant
Lifetime time horizon (base case)	(€ -13,994)		(5.63		
Clegg 2003 - ***Gastric bypass versus conservative therapy***	£ 9,764	£ 6,964	11.67	11.22	£ 6,289
Females 90% **BMI 45 kg/m²**	(£ 2,800)		(0.45)		
Hoerger 2010 - ***Gastric bypass versus conservative therapy***	$ 86,665	$ 71,130	11.76	9.55	$ 7,029
Newly diagnosed diabetes	($ 15,535)		(11.76)		
Established diabetes	$ 99,944	$ 79,618	9.38	7.68	$ 11,956
	($ 20,326)		(1.7)		
Ikramuddin 2009 - ***Laparoscopic Roux-en-Y gastric bypass versus conservative therapy***	$ 83,482	$ 63,722	6.782	5.833	$ 21,980
Females 77.9% (Age 50.1) **BMI 48.4 kg/m²**	($ 19,760)		(0.899)		
Mäklin 2011 - ***Gastric bypass versus conservative therapy***	€ 33,379	€ 50,667	7.67	7.04	Bariatric surgery dominant
Females 65% (Age 43) **BMI 47 kg/m²**	(€ -17,288)		(0.63)		
Michaud 2012 - ***Roux-en-Y gastric bypass versus conservative therapy***	$ 369,585	$ 354,234	30.35	28.8	$ 9,904
BMI > 40 or BMI > 35 **kg/m²** with high risk comorbidities	($ 15,351)		(1.55)		
BMI > 35 or BMI > 30 **kg/m²** with qualifying comorbidities	$ 352,244	$ 338,205	29.87	28.87	$ 12,999
	($ 14,039)		(1.08)		
Picot 2009 – ***Gastric bypass versus non-surgical intervention***	£ 19,824	£ 13,561	12.32	10.8	£ 4,120

20 years time horizon (age 40) **BMI ≥ 40 kg/m²**					
	(£ 6,263)		(1.52)		
Wang 2014 - ***Laparoscopic gastric bypass versus conservative therapy***	$ 169,074	$ 150,934	13.4	10.6	$ 6,479
Females 77.9% (Age 50.1) **BMI 48.4 kg/m²**	($ 18,140)		(2.8)		

I=intervention, C-comparator, QALY-quality adjusted life years, BMI-body mass index, Bariatric surgery dominant= cost saving

Table 17 Other bariatric surgeries

Study	Total Costs of I	Total costs of C	Outcome of I (QALY)	Outcome of C (QALY)	ICER per QALY
	(Difference)		(Difference)		
Ackroyd 2006, 2006 – ***Adjustable gastric banding versus conservative therapy***	€ 13,610	€ 17,197	3.03	2.0	* Bariatric surgery dominant
Germany	(€ -3,587)		(1.03)		
France	€ 14,796	€ 19,276	3.03	2.0	Bariatric surgery dominant
	(€ -4,480)		(1.03		
United Kingdom	€ 9,072	€7,083	3.03	2.0	€ 1,931
	(€ 3,203)		(1.03)		
Craig 2002 – Open ***Gastric bypass versus conservative therapy***	$ 68,600	$ 38,500	19.56	18.51	$ 28,667
Males (age 35) **BMI 40 kg/m²**	($ 30,100)		(1.05)		
Males (age 35) **BMI 50 kg/m²**	$ 75,000	$ 53,200	18.87	16.83	$ 10,686
	($ 21,800)		(2.04)		
Males (age 55) **BMI 40 kg/m²**	$ 77,600	$ 47,900	13.32	12.48	$ 35,357
	($ 29,700)		(0.84)		
Males (age 55) **BMI 50 kg/m²**	$ 85,300	$ 63,500	12.81	11.17	$ 13,293
	($ 21,800)		(1.64)		

Females (age 35) **BMI 40 kg/m²**	$ 59,000	$ 35,300	19.82	18.21	$ 14,720
	($ 23,700)		(1.61)		
Females (age 35) **BMI 50 kg/m²**	$ 64,800	$ 48,500	18.88	16.03	$ 5,719
	($ 16,300)		(2.85)		
Females (age 55) **BMI 40 kg/m²**	$ 69,600	$ 84,200	13.94	12.62	$ 1,612
	($ 21,400)		(1.32)		
Females (age 55) **BMI 50 kg/m²**	$ 77,000	$ 64,100	13.23	10.88	$ 5,489
	($ 12,900)		(2.35)		
Hoerger 2010 - ***Banding surgery versus conservative therapy***	$ 89,029	$ 71,130	11.12	9.55	$ 11,401
Newly diagnosed diabetes	($ 17,899)		(1.57)		
Established diabetes	$ 96,921	$ 79,618	9.02	7.68	$ 12,913
	($ 17,303)		(1.34)		
Keating 2009 – ***Laparoscopic adjustable gastric banding versus conservative therapy***	AUD 98,931	AUD 101,376	15.7	14.5	Bariatric surgery dominant
Females 55% (Age 49) **BMI 30-39.9 kg/m²**	(AUD -2,445)		(1.2)		
Mäklin 2011 – ***Gastric banding versus conservative therapy***	€ 34,594	€ 42,070	7.39	7.19	Bariatric surgery dominant
Females 65% (Age 43) **BMI 47 kg/m²**	(€ -7,476)		(0.2)		
Picot 2009 - ***Adjustable gastric banding versus non-surgical intervention***	£ 17,126	£ 13,561	11.72	10.8	£ 3,875
20 years time horizon (age 40) **BMI ≥ 40 kg/m²**	(£ 3,565)		(0.92)		
Picot 2012 – ***Laparoscopic adjustable gastric banding versus conservative therapy***	£ 14,182	£ 11,148	4.09	3.48	£ 4,974
5 years time horizon **BMI ≥30 and <40 kg/m²), with T2D**	(£ 3,034)		(0.61)		
20 years time horizon	£ 35,055	£ 33,262	11.49	10.39	£ 1,630

BMI ≥30 and <40 kg/m²), with T2D	(£ 1,793)		(1.1)		
5 years time horizon **BMI ≥30 and <35 kg/m²)**	£ 9,923	£4,801	4.03	3.74	£ 17,662
	(£ 5,122)		(0.29)		
20 years time horizon **BMI ≥30 and <35 kg/m²)**	£ 15,211	£ 9,750	11.52	11.12	£ 13,653
	(£ 5,461)		(0.4)		
Pollock 2013 – Laparoscopic adjustable gastric banding versus conservative therapy 10 years time horizon Females 53.5% (Age 46.9) **BMI 42.4 kg/m²**	£ 12,584	£ 7,826	5.63	5.35	£ 16,993
	(£ 4,758)		(0.28)		
20 years time horizon Females 53.5% (Age 46.9) **BMI 42.4 kg/m²**	£ 18,089	£ 14,633	8.63	8.05	£ 5,959
	(£ 3,456)		(0.58)		
30 years time horizon Females 53.5% (Age 46.9) **BMI 42.4 kg/m²**	£ 122,203	£ 19,047	9.85	8.99	£ 3,670
	(£3,156)		(0.86)		
Lifetime horizon Females 53.5% (Age 46.9) **BMI 42.4 kg/m²**	£ 23,562	£ 20,263	10.05	9.14	£ 3,625
	(£ 3,299)		(0.91)		
Wang 2014 – Laparoscopic adjustable gastric banding versus conservative therapy Females 77.9% (Age 50.1) **BMI 48.4 kg/m²**	$ 164,313	$ 150,934	12.8	10.6	$ 6,081
	($ 13,379)		(2.2)		
Wang 2014 – conventional open *Roux-en-Y gastric bypass versus conservative* Females 77.9% (mean age 50.1), BMI 48.4 kg/m²	$ 194,858	$ 150,934	13.2	10.6	$ 16,894
	($ 43,924)		(2.6)		

I=intervention, C-comparator, QALY-quality adjusted life years, BMI-body mass index, Bariatric surgery dominant= cost saving

Figure 20 Cost-effectiveness plane, based on costs adapted to Switzerland and original effect estimates

Blue dots represent the cost-effectiveness studies assessing laparoscopic gastric bypass (or other term used in the studies gastric bypass (GB), laparoscopic Roux-en-Y gastric bypass (LRYGB), Roux-en-Y gastric bypass (RYGB), versus conservative treatment. Red dots represent cost-effectiveness studies assessing other bariatric surgeries (adjustable gastric banding (AGBAN), gastric banding (GBAN), laparoscopic adjustable gastric banding (LAGBAN) conservative (open) Roux-en-Y gastric bypass (ORYGB) versus conservative treatment. In the case of studies providing results for different patient groups, results were selected for presentation based on the longest available time horizon (if there was more than one time horizon available), female gender (when available in the study) and BMI >35 kg/m^2. Ackroyd et al. described results based on the three different countries, Cambel et al. for two estimates of treatment effects (by Angrisani and O'Brien), Hoerget et al. for two diabetes groups, Wang et al. for two types bypass surgeries and gastric banding, and Picot et al. for two different banding techniques. From the Borg et al. study, we selected the results for female patients aged 45-54 years, with BMI 40 kg/m^2.

Table 18 Results of economic evaluations, adopted for Switzerland

Study	Total Costs of I	Total costs of C	Outcome of I (QALY)	Outcome of C (QALY)	ICER per QALY
	(Difference)		(Difference)		
Ackroyd 2006 - ***Gastric bypass versus conservative therapy*** Germany	CHF 38,832	CHF 54,891	3.34	2.0	* Bariatric surgery dominant
	(CHF -16,058)		(1.34)		
France	CHF 42,725	CHF 61,465	3.34	2.0	Bariatric surgery dominant
	(CHF -18,740)		(1.34)		
United Kingdom	CHF 44,477	CHF 34,539	3.34	2.0	CHF 7,416
	(CHF 9,938)		(1.34)		
Borg 2014 - ***Gastric bypass versus conservative therapy*** Male (Age 45-54) **BMI 30-34 kg/m²**	CHF 12,975	CHF 8,588	14.64	11.68	CHF 1,482
	(CHF 4,387)		(2.96)		
Male (Age 45-54) **BMI 35-39 kg/m²**	CHF 17,964	CHF 40,305	13.51	10.17	Bariatric surgery dominant
	(CHF -22,340)		(3.34)		
Male (Age 45-54) **BMI 40-44 kg/m²**	CHF 30,080	CHF 96,575	11.91	8.43	Bariatric surgery dominant
	(CHF -66,495)		(3.48)		
Male (Age 45-54) **BMI 45-49 kg/m²**	CHF 61,223	CHF 182,607	10.6	7.17	Bariatric surgery dominant
	(CHF -121,383)		(3.43)		
Male (**Age 25–34**) BMI 40 -44 **kg/m²**	CHF 38,007	CHF 116,793	14.84	10.97	Bariatric surgery dominant
	(CHF -78,786)		(3.87)		
Male (**Age 35–44**) BMI 40 -44 **kg/m²**	CHF 33,514	CHF 106,201	13.74	9.94	Bariatric surgery dominant
	(CHF -72,687)		(3.8)		
Male (**Age 45–54**) BMI 40 -44	CHF 30,080	CHF 96,575	11.91	8.43	Bariatric surgery dominant
	(CHF -66,495)		(3.48)		
Male (**Age 55-64**) BMI 40 -44 **kg/m²**	CHF 27,048	CHF 73,518	9.6	6.59	Bariatric surgery dominant
	(CHF -46,469)		(3.01)		
Male (**Age 65–74**) BMI 40 -44 **kg/m²**	CHF 23,025	CHF 46,609	6.78	4.54	Bariatric surgery dominant
	(CHF -23,584)		(2.24)		
Female (Age 45-54) **BMI 30–34 kg/m²**	CHF 14,630	CHF 8,027	15.53	12.62	CHF 2,269
	(CHF 6,603)		(2.91)		
Female (Age 45-54) **BMI 35–39 kg/m²**	CHF 19,562	CHF 24,482	14.29	10.95	Bariatric surgery
	(CHF -4,921)		(3.34)		

					dominant
Female (Age 45-54) **BMI 40–44 kg/m²**	CHF 25,979	CHF 56,332	13.00	9.35	Bariatric surgery dominant
	(CHF -30,352)		(3.65)		
Female (Age 45-54) **BMI 45–49 kg/m²**	CHF 41,923	CHF 102,768	11.46	7.97	Bariatric surgery dominant
	(CHF -60,845)		(3.49)		
Female (**Age 25–34**) BMI 40-44 **kg/m²**	CHF 29,904	CHF 67,698	15.27	11.4	Bariatric surgery dominant
	(CHF -37,794)		(3.87)		
Female (**Age 35–44**) BMI 40-44 **kg/m²**	CHF 27,785	CHF 62,558	14.55	10.72	Bariatric surgery dominant
	(CHF -34,774)		(3.83)		
Female (**Age 45–54**) BMI 40-44 **kg/m²**	CHF 25,979	CHF 56,332	13.00	9.35	Bariatric surgery dominant
	(CHF -30,352)		(3.65)		
Female (**Age 55–64**) BMI 40-44 **kg/m²**	CHF 23,866	CHF 43,515	10.75	7.57	Bariatric surgery dominant
	(CHF -19,650)		(3.18)		
Female (**Age 65–74**) BMI 40-44 **kg/m²**	CHF 21,033	CHF 28,450	7.93	5.35	Bariatric surgery dominant
	(CHF -7,417)		(2.58)		
Campbell 2010 - *Laparoscopic Roux-en-Y gastric bypass versus conservative therapy*	CHF 166,423	CHF 144,705	19.054	16.55	CHF 7,492
Aggregate population (Base case - 40 years) – Angrisani et al.	(CHF 21,718)		(2.9)		
Aggregate population (Base case - 40 years) – O'Brien et al.	CHF 172,598	CHF 144,705	18.56	16.155	CHF 11,598
	(CHF 27,893)		(2.4)		
Males (Angrisani et al.) **BMI 35-39.9 kg/m²**	CHF 156,124	CHF 117,263	18.431	16.38	CHF 18,947
	(CHF 38,861)		(2.05)		
Males (Angrisani et al.) **BMI 40-49.9 kg/m²**	CHF 160,800	CHF 135,711	17.966	14.805	CHF 7,937
	(CHF 25,089)		(3.16)		
Males (Angrisani et al.) **BMI > 50 kg/m²**	CHF 163,624	CHF 156,387	17.682	12.835	CHF 1,493
	(CHF 7,238)		(4.85)		
Males (O'Brien et al.) **BMI 35-39.9 kg/m²**	CHF 157, 043	CHF 117,263	18.335	16.38	CHF 2,0347
	(CHF 39,779)		(1.96)		
Males (O'Brien et al.) **BMI 40-49.9 kg/m²**	CHF 166,258	CHF 135,711	17.421	14.805	CHF 11,677
	(CHF 30,547)		(2.62)		
Males (O'Brien et al.) **BMI > 50 kg/m²**	CHF 175,954	CHF 156,387	16.435	12.835	CHF 5,435
	(CHF 19,568)		(3.60)		

Females (Angrisani et al.) **BMI 35-39.9 kg/m²**	CHF 163,464	CHF 127,118	19.662	17.756	CHF 16,069
	(CHF 36,346)		(1.91)		
Females (Angrisani et al.) **BMI 40-49.9 kg/m²**	CHF 168,911	CHF 149,762	19.238	16.338	CHF 6,603
	(CHF 19,149)		(2.90)		
Females (Angrisani et al.) **BMI > 50 kg/m²**	CHF 172,206	CHF 176,053	18.979	14.449	Bariatric surgery dominant
	(CHF -3,847)		(4.53)		
Females (O'Brien et al.) **BMI 35-39.9 kg/m²**	CHF 164,586	CHF 127,118	19.567	17.756	CHF 20,689
	(CHF 37,467)		(1.81)		
Females (O'Brien et al.) **BMI 40-49.9 kg/m²**	CHF 175,686	CHF 149,762	17.728	16.338	CHF 18,650
	(CHF 25,924)		(1.39)		
Females (O'Brien et al.) **BMI > 50 kg/m²**	CHF 187,587	CHF 176,053	17.804	14.449	CHF 3,438
	(CHF 11,534)		(3.36)		
Castilla 2014 - *Gastric bypass versus conservative therapy*	CHF 74,396	CHF 134,121	18.18	12.55	Bariatric surgery dominant
5 years time horizon	(CHF -59,725)		(5.63)		
Clegg 2003 - *Gastric bypass versus conservative therapy*	CHF 67,507	CHF 48,148	11.67	11.22	CHF 43,480
Females 90% **BMI 45 kg/m²**	(CHF 19,360)		(0.45)		
Hoerger 2010 - *Gastric bypass versus conservative therapy*	CHF 122,442	CHF 100,494	11.76	9.55	CHF 9,931
Newly diagnosed diabetes	(CHF 21,948)		(11.76)		
Established diabetes	CHF 141,203	CHF 112,486	9.38	7.68	CHF 16,892
	(CHF 28,717)		(1.7)		
Ikramuddin 2009 - *Laparoscopic Rouxen-Y gastric bypass versus conservative therapy*	CHF 104,819	CHF 80,008	6.782	5.833	CHF 27,598
Females 77.9% (Age 50.1) **BMI 48.4 kg/m²**	(CHF 24,810)		(0.899)		
Mäklin 2011 - *Gastric bypass versus conservative therapy*	CHF 102,712	CHF 155,910	7.67	7.04	Bariatric surgery dominant
Females 65% (Age 43) **BMI 47 kg/m²**	(CHF -53,198)		(0.63)		

Michaud 2012 - ***Roux-en-Y gastric bypass versus conservative therapy***	CHF 410,672	CHF 393,614	30.35	28.8	CHF11,005
BMI > 40 or BMI > 35 kg/m² with high risk comorbidities	(CHF 17,058)		(1.55)		
BMI > 35 or BMI > 30 kg/m² with qualifying comorbidities	CHF 391,403	CHF 375,803	29.87	28.87	CHF 14,444
	(CHF 15,600)		(1.08)		
Picot 2009 – *Gastric bypass versus non-surgical intervention*	CHF 81,860	CHF 55,998	12.32	10.8	CHF 17,015
20 years time horizon (age 40) **BMI ≥ 40 kg/m²**	(CHF 25,862)		(1.52)		
Wang 2014 - *Laparoscopic gastric bypass versus conservative therapy*	CHF 187,870	CHF 167,713	13.4	10.6	CHF 7,199
Females 77.9% (Age 50.1) **BMI 48.4 kg/m²**	(CHF 20,157)		(2.8)		

I=intervention, C-comparator, QALY-quality adjusted life years, BMI-body mass index, Bariatric surgery dominant= cost saving

Table 19 Other bariatric surgeries

Study	**Total costs of surgical strategy**	**Total costs of comparator strategy**	**Outcome of surgical strategy (QALY)**	**Outcome of surgical strategy (QALY)**	**ICER per QALY**
	(Difference)		**(Difference)**		
Ackroyd 2006 – *Adjustable gastric banding versus conservative therapy*	CHF 43,441	CHF 54,891	3.03	2.0	Bariatric surgery dominant
Germany	(CHF -11,449)		(1.03)		
France	CHF 47,180	CHF 61,465	3.03	2.0	Bariatric surgery dominant
	(CHF -14,285)		(1.03		
United Kingdom	CHF 44,238	CHF 34,539	3.03	2.0	CHF 9417
	(CHF 9,699)		(1.03)		
Craig 2002 - Open *Gastric bypass versus conservative therapy* Males (age 35)	CHF 131,201	CHF 73,633	19.56	18.51	CHF 54,826
	(CHF 57,568)		(1.05)		

BMI 40 kg/m²					
Males (age 35) **BMI 50 kg/m²**	CHF 143,441	CHF 101,747	18.87	16.83	CHF 20,438
	(CHF 41,694)		(2.04)		
Males (age 55) **BMI 40 kg/m²**	CHF 148,413	CHF 91,611	13.32	12.48	CHF 67,622
	(CHF 56,802)		(0.84)		
Males (age 55) **BMI 50 kg/m²**	CHF 163,140	CHF 121,447	12.81	11.17	CHF 25,423
	(CHF 41,693)		(1.64)		
Females (age 35) **BMI 40 kg/m²**	CHF 112,840	CHF 67,513	19.82	18.21	CHF 28,154
	(CHF 45,327)		(1.61)		
Females (age 35) **BMI 50 kg/m²**	CHF 123,933	CHF 92,758	18.88	16.03	CHF 10,938
	(CHF 31,174)		(2.85)		
Females (age 55) **BMI 40 kg/m²**	CHF 133,113	CHF 92,185	13.94	12.62	CHF 31,006
	(CHF 40,928)		(1.32)		
Females (age 55) **BMI 50 kg/m²**	CHF 147,266	CHF 122,594	13.23	10.88	CHF 10,499
	(CHF 24,672)		(2.35)		
Hoerger 2010 - *Banding surgery versus conservative therapy*	CHF 125,782	CHF 100,494	11.12 9.55		CHF 16,107
Newly diagnosed diabetes	(CHF 25,288)		(1.57)		
Established diabetes	CHF 136,932	CHF 112,486	9.02	7.68	CHF 18,243
	(CHF 24,446)		(1.34)		
Keating 2009 – *Laparoscopic adjustable gastric banding versus conservative therapy*	CHF 230,703	CHF 236,405	15.7	14.5	Bariatric surgery dominant
Females 55% (Age 49) **BMI 30-39.9 kg/m²**	(CHF -5,702)		(1.2)		
Mäklin 2011 – *Gastric banding versus conservative therapy*	CHF 106,451	CHF 129,455	7.39	7.19	Bariatric surgery dominant
Females 65% (Age 43) **BMI 47 kg/m²**	(CHF -23,005)		(0.2)		
Picot 2009 - *Adjustable gastric banding versus non-surgical intervention*	CHF 70,719	CHF 55,998	12.32	10.8	CHF 9,685
20 years time horizon (age 40) **BMI ≥ 40**					

kg/m²					
	(CHF 14,721)		(1.52)		
Picot 2012 – ***Laparoscopic adjustable gastric banding versus conservative therapy***	CHF 50,406	CHF 39,623	4.09	3.48	CHF 17,678
5 years time horizon **BMI ≥30 and <40 kg/m²), with T2D**	(CHF 10,784)		(0.61)		
20 years time horizon **BMI ≥30 and <40 kg/m²), with T2D**	CHF 124,594	CHF 118,222	11.49	10.39	CHF 5,793
	(CHF 6,373)		(1.1)		
5 years time horizon **BMI ≥30 and <35 kg/m²)**	CHF 35,269	CHF 17,064	4.03	3.74	CHF 62,775
	(CHF 18,205)		(0.29)		
20 years time horizon **BMI ≥30 and <35 kg/m²)**	CHF 54,064	CHF 34,654	11.52	11.12	CHF 48,524
	(CHF 19,410)		(0.4)		
Pollock 2013 – ***Laparoscopic adjustable gastric banding versus conservative therapy***	CHF 44,727	CHF 27,816	5.63	5.35	CHF 60,397
10 years time horizon Females 53.5% (Age 46.9) **BMI 42.4 kg/m²**	(CHF 16,911)		(0.28)		
20 years time horizon Females 53.5% (Age 46.9) **BMI 42.4 kg/m²**	CHF 64,293	CHF 52,009	8.63	8.05	CHF 21,178
	(CHF 12,283)		(0.58)		
30 years time horizon Females 53.5% (Age 46.9) **BMI 42.4 kg/m²**	CHF 78,915	CHF 67,698	9.85	8.99	CHF 13,043
	(CHF 14,473)		(0.86)		
Lifetime horizon Females 53.5% (Age 46.9) **BMI 42.4 kg/m²**	CHF 83,745	CHF 72,020	10.05	9.14	CHF 12,885
	(CHF 11,725)		(0.91)		
Wang 2014 ***Laparoscopic adjustable gastric banding versus conservative therapy***	CHF 182,580	CHF 167,713	12.8	10.6	CHF 6,757
Females 77.9% (Age					

50.1) **BMI 48.4 kg/m²**	(CHF 14,866)		(2.2)		
Wang 2014 – conventional open *Roux-en-Y gastric bypass versus conservative therapy*	CHF 216,520	CHF 167,713	13.2	10.6	CHF 18,772
Females 77.9% (Age 50.1) **BMI 48.4 kg/m²**	(CHF 48,807)		(2.6)		

4.2.4 Sensitivity Analysis of adapted costs for Switzerland

The aim of the sensitivity analysis was to assess the impact on ICERs adapted for Switzerland when the adjustment of costs over time was based on inflation rates instead of the increase in health care costs (see section 4.1.6). Overall, this alternative approach to calculating adapted costs and ICERs had no large effect on the ICERs, for the studies reviewed. No differences were generated for the studies by Borg et al. and Castilla et al., which were quite recent (cost year 2012). The largest impact on ICERs was seen for the older studies by Clegg et al. (-31.70 % of difference) and Craig et al. (-28.07 % of difference) when we compared the adapted inflation costs with health care costs growth costs (Table 20). For the other 11 studies, the change remained below 20%.

Table 20 Alternative approach to adaptation of costs, using inflation rates instead of the increase in health care costs in step 3.

Study	Difference (%) in –resulting ICERs	Study	Difference (%) in –resulting ICERS
1. Ackroyd 2006	- 20.4	**8. Mäklin 2011**	-8.5
2. Cambell 2010	- 20.2	**9. Michaud 2012**	-8.5
3. Clegg 2003	-31.7	**10. Picot 2009**	-9.9
4. Craig 2002	-28.1	**11. Picot 2012**	-8.5
5. Hoerger 2010	-20.4	**12. Pollock 2013**	-8.5
6. Ikramuddin 2009	-17.1	**13. Wang 2014**	-8.5
7. Keating 2009	-20.2		

4.2.5 Budget impact analysis

4.2.5.1 Surgery costs

As described in section 4.1.8, the potential, immediate surgery costs for treating all eligible patients in one single year was estimated. In 2012, almost 25,000 persons with a BMI higher than 40 kg/m² lived in Switzerland, whereas the number of persons with BMI 35-40 kg/m² was approximately 100,000. The surgery costs for all persons with a BMI higher than 40 kg/m², if treated in one single year, would reach CHF 353 Mio. Assuming that 10%-to-30% of the persons with BMI 35-40 kg/m² might also be eligible for surgery, the total surgery costs would be CHF

496-781 Mio. In 2014, the estimated total surgery costs for the 4,153 operated persons were CHF 61.1 Mio.

4.2.5.2 Budget impact of bariatric surgery over 10 years

As described in section 4.2.4, cost estimates resulting from the adaptation of qualitatively transferable cost-effectiveness studies to Switzerland were combined with the current use of bariatric surgery in Switzerland. Surgery costs for laparoscopic gastric bypass and follow-up costs per patient were calculated for a 5-year time horizon, using costs estimates by Ackroyd et al., adapted to Switzerland.[80] This choice was to a certain extent arbitrary, driven by the availability of sufficient numerical detail. Ackroyd et al. primarily based their assessment of the effectiveness of the intervention on the remission from diabetes. We calculated two scenarios. In the first scenario, we assumed that all patients undergoing surgery or conservative treatment would have diabetes at baseline (i.e. diabetes prevalence = 100%). In contrast, the second scenario assumed that no patients would have diabetes at baseline (i.e. diabetes prevalence = 0%). Combining the two scenarios allowed us to estimate the budget impact of bariatric surgery assuming different diabetes prevalence rates.

As mentioned in the methods, all presented results depend very strongly on the cost differences reported by Ackroyd et al. (Ackroyd 2006). It is important to emphasize that the use of other cost-effectiveness studies as a basis might have led to very different results. However, the other included cost-effectiveness studies provided less detailed numerical data.

Table 21 Cumulative costs of bariatric surgery versus conservative treatment, depending on the presence or absence of diabetes (Mio. CHF).

Year	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
No. of surgeries	2,300	3,206	3,398	4,153	5,000	5,000	5,000	5,000	5,000	5,000
GBP - diabetes	33.9	48.5	58.3	99.6	151.4	167.2	180.9	200.1	210.2	210.2
Only surgery costs	*33.9*	*47.2*	*50.0*	*61.2*	*73.6*	*73.6*	*73.6*	*73.6*	*73.6*	*73.6*
GBP - no diabetes	33.9	48.5	53.1	66.1	80.8	82.3	83.3	84.2	84.6	84.6
Only surgery costs	*33.9*	*47.2*	*50.0*	*61.2*	*73.6*	*73.6*	*73.6*	*73.6*	*73.6*	*73.6*
Usual care - diabetes	29.3	68.4	109.3	159.7	220.3	252.7	274.3	293.5	303.7	303.7
Usual care - no diabetes	8.3	12.9	15.4	20.2	25.6	27.2	28.3	29.2	29.7	29.7

Table 21 shows the estimated cumulative costs of bariatric surgery versus conservative treatment over 10 years (from year 2011 to 2020). The cost changes from year 2011 to 2015 are mainly due to the increasing number of performed surgeries (from 2,300 in 2011 to 5,000 in 2015) and due to the increasing number of follow-up patients. The presence of diabetes seems to have a higher impact on the total yearly costs of conservative treatment, if compared to surgical intervention.

The reported cost estimates for patients with and without diabetes were combined into two hypothetical scenarios. In the first, hypothetical scenario in which all patients undergoing surgery or conservative treatment had diabetes (i.e. diabetes prevalence = 100%), the mid-term costs would be lower with bariatric surgery. For example, the 5-years cumulative costs for GBP in patients with initial diabetes in 2018 were estimated at CHF 200.1 Mio., compared to CHF 293.5 Mio. for conservative treatment. In contrast, in the second hypothetical scenario in which all patients undergoing surgery or conservative treatment had no diabetes (i.e. diabetes prevalence = 0%), the mid-term costs would be lower with usual care. For example, the 5-years cumulative costs for GBP in patients with initial diabetes in 2018 were estimated at CHF 84.2

Mio., compared to CHF 29.2 Mio. for conservative treatment. Figure 21 and Figure 22 provide a graphical overview of the costs of surgery and usual care for the two scenarios.

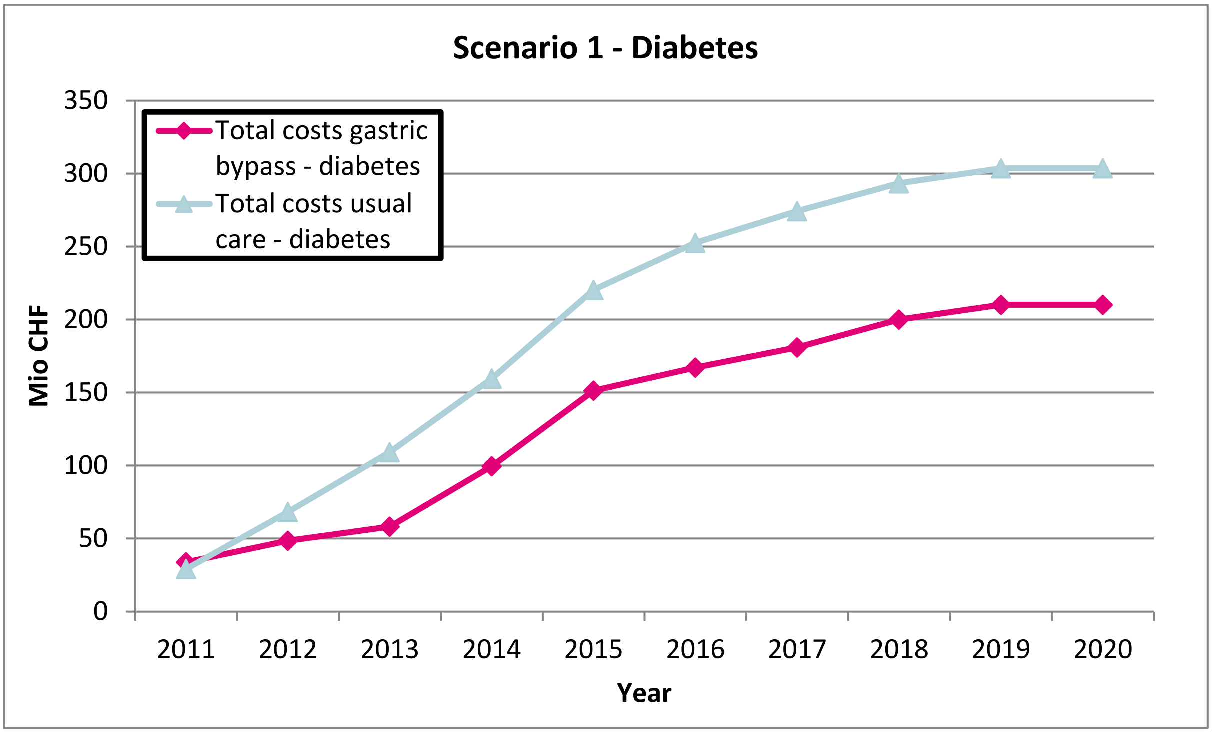


Figure 21 Scenario 1- cumulative total costs for patients with initial diabetes.

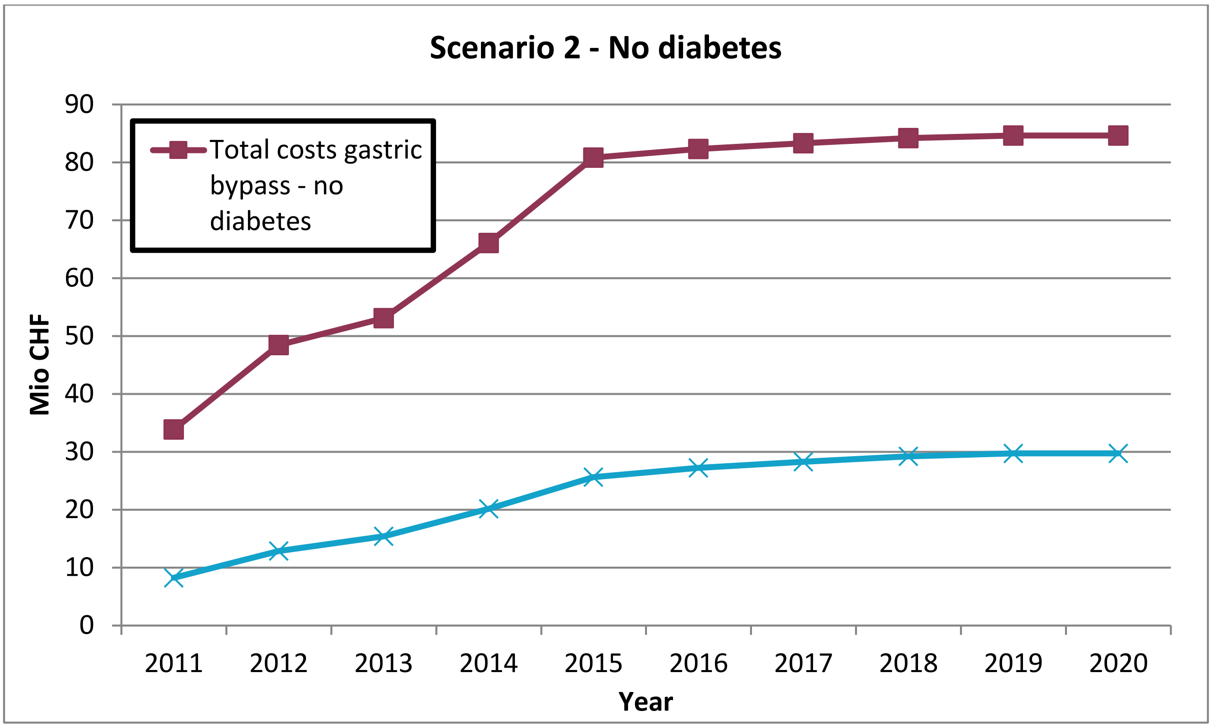


Figure 22 Scenario 2 - cumulative total costs for patients without initial diabetes.

Obviously, neither the first nor the second hypothetical scenario represents reality. For obese patients it may be plausible to assume a diabetes prevalence ranging from 10% to 40%. In Table 22 cost differences between surgery and usual care are reported, assuming different levels of diabetes prevalence. Assuming a diabetes prevalence in obese patients of 10%, the incremental costs of bariatric surgery in 2016 might reach CHF 55.1 Mio. In contrast, assuming a diabetes prevalence in obese patients of 40%, surgery might be cost-saving in 2016 (CHF -1.2 Mio.).

Table 22 Cumulative budget impact (cost differences between bariatric surgery and usual care) depending on diabetes prevalence (in Mio. CHF).

		2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Diabetes prevalence (%)	0	25.6	35.6	37.6	45.9	55.2	55.1	55.0	54.9	54.9	54.9
	10	23.5	30.1	28.8	35.3	42.8	41.0	40.2	40.1	40.1	40.1
	20	21.4	24.5	19.9	24.7	30.4	26.9	25.3	25.3	25.2	25.2
	30	19.3	19.0	11.0	14.1	17.9	12.9	10.5	10.4	10.4	10.4
	40	17.2	13.4	2.2	3.5	5.5	-1.2	-4.4	-4.4	-4.5	-4.5
	50	15.1	7.8	-6.7	-7.1	-6.9	-15.3	-19.2	-19.3	-19.3	-19.3
	60	13.0	2.3	-15.5	-17.7	-19.3	-29.3	-34.0	-34.1	-34.1	-34.1
	70	10.9	-3.3	-24.4	-28.3	-31.7	-43.4	-48.9	-48.9	-49.0	-49.0
	80	8.8	-8.8	-33.3	-38.9	-44.1	-57.5	-63.7	-63.8	-63.8	-63.8
	90	6.7	-14.4	-42.1	-49.5	-56.5	-71.5	-78.5	-78.6	-78.6	-78.6
	100	4.6	-19.9	-51.0	-60.1	-68.9	-85.6	-93.4	-93.4	-93.5	-93.5

Figure 23 shows cost differences between bariatric surgery and conservative treatment, assuming diabetes prevalence rates ranging from 10% to 40%. The variation from year 2011 to year 2015 is mainly due to the changing number of surgeries performed (from 2015 a constant number of 5,000 surgeries per year was assumed) and an increasing number of follow-up patients.

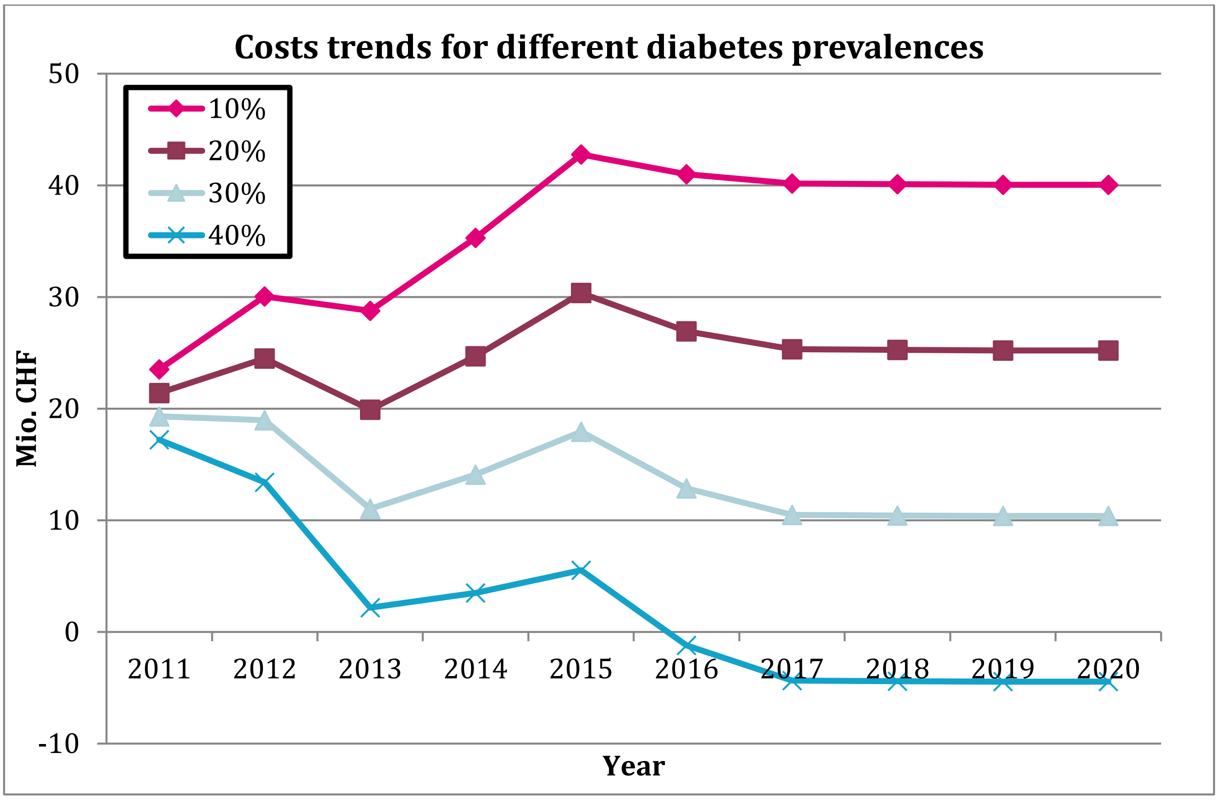


Figure 23 Budget impact (cost differences between bariatric surgery and conservative treatment) estimated for different levels diabetes prevalence (Mio. CHF).

4.3 Discussion

4.3.1 Findings

We investigated the cost-effectiveness of bariatric surgery, using data from one systematic review (Padwal 2011) and 21 cost-effectiveness analyses performed for European countries (Ackroyd 2006, Anselmino 2009, Borg 2014, Castilla 2014, Clegg 2003, Faria 2013, Mäklin 2011, Picot 2009, Picot 2012, Pollock 2013), for the US (Campbell 2010, Craig 2002, Hoerger 2010, Ikramuddin 2009, Jensen 2005, Michaud 2012, Salem 2008, Wang 2013) or for Australia (Keating 2009, Ananthapavan 2010). The available studies mainly examined laparoscopic gastric bypass surgery or laparoscopic gastric banding. Only four studies explicitly included open gastric bypass (Castilla 2014, Craig 2002, McEwen 2010, Wang 2014). Comparator strategies were described as conservative treatment or 'current practice'. They were typically a combination of diet, exercise, and behavioural modification. The vast majority of studies assessed cost-effectiveness as cost per QALY gained, over time horizons ranging from 5 years to lifelong. Results varied between studies, e.g. due to inclusion criteria with regards to age, BMI class / severity of obesity, and comorbidities; the assessment and modelling of costs and effects; and the perspective of cost-assessment taken.

Studies of patients with BMI values >35 kg/m^2 indicated bariatric surgery to be cost-saving or cost-effective based on the criteria set by the authors. This result of our systematic review is consistent with the earlier systematic review published by Padwal et al. in 2011 (Padwal 2011). This review suggested that based on 13 economic evaluations, surgery compared to conservative treatment appeared attractive, with ICERs ranging from $1,000 to $40,000 per QALY gained. However, the authors did not make any definitive conclusions about the cost-effectiveness of alternative surgical procedures.

Based on the evidence we identified, bariatric surgery may be efficient in patients with a BMI of less than 35 kg/m^2 (Borg 2014 and Picot 2012). Borg et al. and Picot et al. concluded that surgery was cost-effective among subgroups of patients with a BMI between 30 kg/m^2 and 35 kg/m^2, whose survival and quality of life were improved with surgery. However, further research may be needed to understand better which subgroups of lower-BMI patients may or may not benefit from bariatric surgery.

A total of 15 studies judged as being qualitatively transferable were included in our adaptation of costs and ICERs for Switzerland. Adapted ICER results for patients with a BMI >35 kg/m^2 indicated a cost saving (dominant) situation or showed ICERs of CHF 8,000 to CHF 44,000 per QALY gained. Differences observed within this range were, amongst others, due to the way how effectiveness was modelled (in terms of duration of BMI changes and dependent effects on morbidity, mortality and costs), time horizon, population studied, and exact type of intervention studied. For patients in the BMI <35 kg/m^2 category, the adapted cost-effectiveness of bariatric surgery compared to conservative treatment was below CHF 3,000 per QALY gained, using a healthcare perspective (Borg 2014) or below CHF 50,000 per QALY gained (Picot 2012).

Procedure-specific differences in terms of benefits and efficiency appear to exist. When laparoscopic gastric bypass and laparoscopic gastric banding were compared to conservative treatment, laparoscopic gastric bypass appeared to be better than gastric banding in terms of benefits (QALYs gained, and life years saved) but also more expensive (Ackroyd 2006, Hoerger

2010, Mäklin 2011, Wang 2014). However, to date there was no direct economic evaluation between gastric bypass and banding. Therefore, further information is required to understand long-term benefits and cost-effectiveness in order to define the advantages and disadvantages of different bariatric procedures.

It also important to highlight that the majority of included studies were from the payers' perspective (Ackroyd 2006, Campbell 2010, Castilla 2014, Clegg 2003, Craig 2002, Ikramuddin 2009, Padwal 2011, Picot 2009, Picot 2012 and Pollock 2013). Three were from a health care perspective (Keating 2009, Mäklin 2011, and Wang 2014) and two were from a societal perspective (Borg 2014, Michaud 2012). Therefore, a certain deviation may be expected to reflect the cost-effectiveness of bariatric surgery from a 'KVG perspective' (taking into account the direct medical costs of all health care services covered by the Swiss statutory health insurance, irrespective of actual payer) and from a social perspective.

Additionally, results of the newer trials (incorporated into the systematic review) with regards to percentage change of body weight (2 years results) were generally consistent with those that were used in the CE analyses such as O'Brien et al, and Dixon et al.[36 111]

4.3.2 Limitations

Our systematic review aimed at gaining an understanding of the to-be-expected cost-effectiveness of bariatric surgery in Switzerland, has several limitations. Firstly, due to a lack of suitable Swiss data, e.g. relating to healthcare resource use, we have not performed a *de novo* health economic modelling. Instead, a systematic review of the international economic literature has been undertaken to understand the impact of bariatric surgery in terms of cost-effectiveness.

Our assessment of qualitative transferability to Switzerland was based on a combination and adaptation of published criteria. For studies judged to be qualitatively transferable, cost estimates were tentatively adapted and ICERs re-calculated on this basis. The adaptation consisted of three distinct steps, namely correction for different levels of resource utilisation, for different prices of healthcare services, and for change in level of resource utilisation and prices over time. Aims were to make international cost-effectiveness results more comparable and to possibly achieve a rough indication of the possible magnitude of ICERs for Switzerland. However, the resulting estimates cannot be directly interpreted as 'ICERS for Switzerland'.

It should be noted that in the identified cost-effectiveness studies, the modelling of effectiveness was based on short-term clinical trials and longer-term observational data such as the Swedish SOS study.[114] However, this is a common aspect of economic evaluation studies with long time horizons, where an extrapolation of effects observed in clinical trials is needed. Although generally affected by higher risks of bias, data from observational studies are crucial to understand the long term impact of interventions in clinical practice. As a result, the available models may overestimate the impact of surgery in some ways, while they may potentially underestimate the true benefit of surgery due to improvement of unmeasured obesity-related comorbidities that were not modelled.

The sources of effectiveness data varied across studies. However, some sources were used frequently. For example, the results of the SOS study[56 114], a prospective, long-term controlled intervention study including 2,010 obese subjects who underwent bariatric surgery and 2,037 matched obese control subjects receiving usual care was used by eleven studies as their main source of effectiveness estimates (Borg 2014, Castilla 2014, Clegg 2003, Craig 2002, Faria 2013, Hoerger 2010, Jensen 2005, Mäklin 2011, Michaud 2012, Padwal 2011, and Picot 2009).

Another limitation of the reviewed studies was that there were no long term data on the impact of changes in body weight on quality of life. Bariatric surgery has been shown to improve quality of life in the short-term, and there were cross-sectional data on the correlation between BMI and quality of life. As such, this information was modelled in the available economic evaluations. Again, this is quite common practice in cost-effectiveness studies with long time horizons. The authors undertook a number of sensitivity analyses to demonstrate the robustness of the data presented in the economic evaluations, but not all followed the rule of varying all parameters with potentially relevant, inherent uncertainty.

A final limitation concerns the outcomes taken into consideration in the available cost-effectiveness analyses. It has been well known for many years that diseases like diabetes, myocardial infarction or stroke are associated with obesity. Not surprisingly, the costs of these diseases are often included in cost assessments of bariatric surgery. However, bariatric surgery may also have a positive or negative impact on the occurrence of additional, potentially costly diseases. For example, the effects of bariatric surgery on bone mineral density have recently been discussed as potentially relevant (Laurent 2014). Bone mineral loss after bariatric surgery could potentially increase the risk of fractures. This would lead to higher costs and reduced cost-effectiveness."

4.3.3 Budget impact analyses

The immediate costs for surgical treatment of all Swiss individuals with a BMI >40 kg/m^2 in a single year would reach CHF 353 Mio. If also 10-30% of the persons with BMI 35-40 kg/m^2 were surgically treated due to the presence of relevant comorbidities, the total costs of surgery would reach CHF 496-781 Mio. This is an obviously unrealistic scenario intended to describe the potential magnitude of economic implications. In 2014, the estimated total surgery costs for the 4,153 surgeries actually performed in Switzerland were CHF 61.1 Mio.

The results of the budget impact estimation over ten years strongly depended on the assumptions made (see sections 4.1.8 and 4.1.8.2). For example, if the diabetes prevalence in obese patients was 10%, the incremental costs of bariatric surgery in 2016 might reach CHF 55.1 Mio. In contrast, a diabetes prevalence in obese patients of 40% might make bariatric surgery cost-saving (CHF -1.2 Mio.).

The estimation over 10 years was performed under numerous assumptions; substantial limitations apply. First, the yearly number of bariatric surgeries was assumed to remain constant after the year 2015, to reflect potential limitations of capacity. Depending on future demographic and epidemiological trends, on the degree of implementation of bariatric surgery and on the availability of surgical centres, the number of persons undergoing surgery may increase or decrease considerably. Second, a single cost-effectiveness study using German data served as the basis for our costs assumptions (Ackroyd 2006). The primary reason for selecting this study was availability of sufficient numerical detail. Specifically, the authors provided treatment costs per year, broken down by bariatric surgery (laparoscopic gastric bypass) vs. conservative treatment; presence vs. absence of diabetes; and, in the applicable strategy, by surgery vs. follow-up. If compared to the other cost-effectiveness studies analysed in this report, the effectiveness impact of bariatric surgery in Ackroyd et al. was in the middle range despite a relatively short follow-up period. This may suggest the use of relatively optimistic effectiveness assumptions. Only short-term costs until 5 years after surgery were included, while several cost-effectiveness analyses suggested that the full benefits of surgery may only emerge in the long-term (see section 4.2.2.9). Therefore, a long-term budget impact analysis might report bigger benefits for patients undergoing surgery versus conservative treatments. A final limitation is that the analysis of Ackroyd et al. was mainly based on the presence or absence of diabetes in the treated population. Effects in terms of mortality reduction or reduction of other comorbidities (e.g. cardiovascular diseases) were not taken into account. There may also be relevant differences concerning diabetes management and conservative treatment of obese patients in Germany vs. Switzerland. Using other cost-effectiveness studies as a basis would have led to different results but due to lack of reporting of suitable details, the range could not be explored.

5 Assessment of the legal domain

5.1 Introduction

5.1.1 Background: the system of compulsory health insurance

An analysis of the legal implications of the results of the comparison between bariatric surgery and conservative treatment for obesity and overweight as undertaken in the present report is to be made on the basis of the applicable legal norms of the Swiss health law, in particular of the pertinent provisions of the Federal Swiss Health Insurance Act (HIA).[1] The HIA establishes a system of compulsory social health insurance: Basic health insurance is mandatory for all persons residing in Switzerland. Every resident must therefore be insured by a recognized health insurance company providing basic sickness coverage in accordance with the HIA.[2] Health insurance providers are required to accept every applicant for basic health insurance (full freedom to choose one's insurer).[3]

Basic health insurance services and treatments to be granted by all health insurance providers are precisely determined by the HIA.[4] The risks covered by the insurance are illness, accidents[5] and maternity.[6] The insured benefits include in particular ambulant treatment by doctors and chiropractors including prescribed medication, psychotherapy, physical therapy and occupational therapy, admission to the general wards of the hospitals officially listed as certified public health service providers by the canton of residence or by another canton. The basic health insurance also takes over certain costs for preventive medicine.[7]

All providers of this basic insurance must cover the full range of benefits as prescribed by law. They are, however, not allowed to reimburse any further so-called "voluntary" benefits. If benefits, which are not part of compulsory insurance, are to be provided or prescribed by a hospital, by a doctor or by a physiotherapist, then there is the legal obligation to inform the patient accordingly and in advance. Such further treatments, which are not part of the benefits of the basic health insurance, may be covered by a voluntary supplementary insurance outside the social health insurance system.

5.1.2 Legal issues

There are the following legal issues pertaining to this report:

---

[1] Bundesgesetz über die Krankenversicherung (KVG) vom 18. März 1994 (SR 832.10).
[2] Art. 3 (1) HIA.
[3] Art. 4 HIA.
[4] Arts 24-31 HIA.
[5] If already insured for accidents by an employer, accident coverage can be excluded.
[6] Art. 1b (2) HIA.
[7] Art. 26 HIA.

- Are obesity and overweight to be qualified as illness covered by the mandatory health insurance under the HIA, and if yes, to which extend?
- If obesity and overweight is an illness: are there any legal standards as to which treatment are to be preferred? Does the applicable law prescribe a certain method, a certain treatment?

5.2 Obesity and Overweight as Illness

5.2.1 Legal definition of illness

The law defines illness as any impairment of physical or mental health not caused by an accident that makes a medical examination or treatment necessary or results in an inability to work.[8]

It is generally recognized that obesity and overweight do not necessarily fall under this definition because they do not necessarily constitute health impairment. However, it is not contested that some extremer forms of obesity may constitute an illness in accordance with the legal definition. An ordinance of the Federal Department of Home Affairs, the HISO,[9] regulates the obesity-related benefits of the federal health insurance.

5.2.2 Benefits covered by the mandatory health insurance

These benefits include:

a) *Advice by a nutritionist*: Coverage includes up to 6consultations with a doctor's prescription or as commissioned by a doctor. If further consultations were necessary, the doctor could renew the prescription. After 12 consultations, the treating doctor must consult one of the insurance company's fiduciary physicians and suggest how his patient's treatment should continue. Health insurance covers nutrition counselling only in case of certain medical conditions or allergies.[10]

b) The treatments covered by the Public Health Insurance are listed in Annex 1 to the HISO:

Bariatric Surgery is covered under the following conditions:

- A BMI higher than 35 kg/m2;

- A therapy, lasting 2 years, to reduce the body weight failed;

- Processes, treatments and surgery in accordance with the Medial Guidelines of the Swiss Society for the Study of Morbid Obesity and Metabolic Disorders (SMOB) of September 25, 2013, regarding Bariatric Surgery;

---

[8] This is the official definition according to Art. 3 (1) of the Federal Act on the General Part of the Social Security Law (ATSG). Art. 1 (2) (a) HIA refers to this definition.

[9] Verordnung des EDI über Leistungen in der obligatorischen Krankenpflegeversicherung vom 29. September 1995 (KLV; SR 832.112.31).

[10] Art. 9b HISO.

- Execution by a recognized centre for bariatric surgery according to the Administrative Guidelines of the SMOB of September 25, 2013, or in non-recognized centres as far as they are able to observe the Medial Guidelines of the SMOB and after consultation of the insurance company's fiduciary physicians.

All this is regulated in detail by the SMOB Medical Guidelines, which are de facto legally binding because also the public health authorities and the insurance companies refer to them. It is a case of non-state standard setting by a professional community to which the practice of state authorities and courts defer.

There is no coverage for treatments with intragastric balloons.

c) *Internal medicine treatments* are covered, i.e. non-surgical, treatments in case of an overweight of more than 20% or generally in case of comorbidity if weight reduction has a positive impact on the principal illness.

d) Special regulations apply to the *treatment of children and teenagers*.

e) There is no coverage for *preventive examinations and consulting services* by doctors or nutritionists.

5.2.3 Summary

To sum up: It is recognized by the HIA and by the HISO that obesity and overweight may constitute an illness according to the definition of Art. 1a (2) (a) HIA. The details are regulated by the ordinance and by the Medical Guidelines of the SMOB. Overweight has to pass a certain threshold to qualify as an illness. To cure that illness, the applicable legal rules provide for both conservative treatments as well as bariatric surgery. The methods and treatments not covered by the mandatory health insurance may be applied nonetheless if they are medically sound and based on an agreement between doctor and patient. Coverage of the costs, however, depends on whether the status of a person may be qualified as an illness and on whether there the patient has concluded a supplementary insurance covering such risks and treatments.

5.3 Does law prescribe certain treatments?

5.3.1 Legal preference for conservative methods

According to the report's findings there is some evidence for a certain superiority of bariatric surgery over conservative treatments. This corresponds to the conclusions of the SMOB in the Medical Guidelines.[11]

[11] The Guideline's statements are crystal clear: "Alle bisher bekannten, konservativen Behandlungsprogramme können die Fettmasse bei der Überzahl Betroffener weder ausreichend noch anhaltend senken….Erst die wenig traumatisierende, laparoskopische Chirurgie zu Beginn des letzten Jahrzehnts des 20. Jahrhunderts machte die bariatrische Chirurgie zur bisher wirksamsten, zweckmässigsten und wirtschaftlichsten Behandlungsform der Adipositas sowie zwei ihrer wichtigsten Folgeerkrankungen Diabetes mellitus Typ II und Dyslipidämie. Die bariatrisch-metabolische Chirurgie ist aktuell die nachhaltigste Behandlungsform der Bariatrie." SMOB Medical Guidelines p. 3.

The applicable legal norms, however, require as a general rule first a two year conservative treatment. Also according to international medical guidelines there usually should be first attempts to reduce the weight by conservative methods. The problem is that in Switzerland this is not just a medical guideline but a legal rule, which severely limits the doctors' professional discretion and which could work against the patients' interest because the public health insurance would not cover bariatric surgery without previous attempts to reduce weight by conservative methods. Against the background of the report's and of others' such as the SMOB's findings it could make sense from a medical as well as from a financial point of view to *modify the Federal Ordinance* and to allow the doctors from the very beginning to take both treatments, conservative and bariatric surgery, seriously into account and to discuss the alternatives with the patients in order to fix an agreed course of action. Such a decision is controlled by general principles, which adequately applied will prevent doctors from an uncritical preference for bariatric surgery.

5.3.2 General principles: efficiency, functionality, cost-effectiveness of a treatment

These legal principles determining, apart from medical considerations, which treatment should be elected are regulated by Art. 32 HIA. According to that provision, treatments and services provided in accordance with the HIA have to be efficient, functional and economical.[12]

Efficiency means that the method used should be generally apt to reach the diagnostic and therapeutic goals,[13] that it is usually successful and that the overall success of the treatment is supported by scientific evidence.[14]

Functionality means that the methods and treatments used in a particular case will with a certain probability reach the goals and that the positive effects of the treatment will prevail over potential adverse effects.[15]

Economical means foremost cost effective.[16] In case there are several comparably beneficial treatments the most cost effective treatments should be chosen by the responsible doctor or health institution. Relevant are only the direct costs, not, e.g. the overall costs for the social insurance system, national income and expenditure, or society at large.[17] To be relevant as a criterion, clear cost differences are required.[18] Otherwise, the criterion does not play a role.

According to the practice of the Swiss Federal Court, there is a general presumption that the medical treatments fulfil the aforementioned criteria. The doctors have, based in their know-how and experience, a considerable margin of appreciation as to which treatments fits in a certain situation for a certain patient.[19] The criteria are not infringed if ex post a chosen

---

[12] In German: the so-called "WZW-Kriterien".
[13] Eugster, G. (2010), Bundesgesetz über die Krankenversicherung [KVG], Kommentar zu Art. 32 KVG, Zürich, Rz. 2.
[14] BGE 133 V 115 E. 3.1.
[15] BGE 130 V 299 E. 6.1. Eugster, G. (2010), Bundesgesetz über die Krankenversicherung [KVG], Kommentar zu Art. 32 KVG, Zürich, Rz. 8.
[16] BGE 130 V 532 E. 2.2. Vgl. Urteil 9C_224/2009 (insb. E. 4.2): Magenbanding und Magenbypass gleichermassen zweckmässig, Magenbanding aber in casu als wirtschaftlicher beurteilt.
[17] See e.g. BAG, Operationalisierung der Begriffe Wirksamkeit, Zweckmässigkeit und Wirtschaftlichkeit, Arbeitspapier 21. Juli 2011, S. 6.
[18] BGE 124 V 196 E. 3.
[19] Decision of the Swiss Federal Court 9C_224/2009 of September 11, 2009, 1.1.

treatment appears to be less apt and adequate than another method. In addition, the application of these criteria always requires a case-by-case assessment.

5.3.3 Significance for obesity and overweight

Therefore, from a legal point of view, the report's findings that bariatric surgery might be more successful than conservative treatment does not imply that there is a legal obligation to generally prefer surgery over conservative treatment. It is up to the patient, together with the doctor, to make his or her decisions.

The results of the present report show that as to effectiveness, bariatric surgery seems to be more effective than conservative treatment for some but not all the critical and important outcomes of this report. As to functionality, both methods may be used but the general outcome of conservative treatment is much less convincing. Given, however, the limitation as to the coverage of bariatric surgery by the mandatory health insurance, in many instances conservative treatment will prevail.

As to the cost-effectiveness, the health economic analysis' focus is on the general costs, not on the specific costs of the treatment. This does, as explained, not correspond to the legal understanding of cost effective treatments according to the last criterion of Art. 32 HIA. Generally, it seems, however, that surgery is less costly than conservative treatment. This is legally relevant only for cases of a BMI higher than 35 kg/m^2, because only in such cases the costs of surgery are borne by the public health insurance.

The criteria of Art. 32 HIA not only apply to decide on the alternative between surgery and conservative treatment but also when choosing between several surgical methods. The Swiss Federal Court held in a decision that in a particular situation gastric banding and gastric bypass were both equally functional but that gastric banding was more cost effective[20] and that therefore the costs of gastric bypass would not be borne by the mandatory health insurance.

5.4 Conclusions

Obesity and overweight are, from a legal point of view, only partly considered as an illness. The consequence thereof is that the treatment costs are often not borne by the public health insurance system. In addition, in case of an illness the applicable legal rules display a certain bias for conservative treatment and against bariatric surgery. This may not be justified as shown by the present report, because for some outcomes bariatric surgery was shown to be more effective e.g. on body weight loss at lower costs than conservative treatment. However, for other critical outcomes e.g. stroke or important outcomes e.g. myocardial infarction, data was too sparse to draw any conclusions.

[20] See the decision referred to in the previous footnote.

It is one thing under which circumstances obesity and overweight has to be considered as an illness, and another thing which treatment is adequate in case of an illness. Illness is a general legal term which applies not only in the public health sector but in other sectors as well. Only medical consideration should determine the pathological character of obesity and overweight, this in accordance with the legal definition of illness. However, the regulations concerning treatment are unduly restrictive. They are like a straitjacket and unjustifiably limit the medical margin of discretion. The report's results show that a preference for conservative treatment may not be justified.

6 Assessment of ethical issues

6.1 Method

The evidence for the ethical assessment was reviewed based on:

1) issues which became apparent during scoping, and in subsequent discussions during assessment;
2) a systematic analysis of possible ethical issues based on three grids based respectively on a principalist approach[129], Norman Daniels' *benchmarks of fairness*[130] – which was developed to assess health system changes but contains elements relevant to "the development of a technology (within the health system)" [131] – and Tavaglione's claims-based vulnerability framework[132];
3) a literature search in Pubmed and EBSCO for ethical issues associated with bariatric surgery using keywords associated with both of these terms, followed by screening of resulting titles, abstracts and papers. As recommended by EUnetHTA, this literature search was complemented by a reflective process of literature consultations on ethical issues associated with other, more studied, situations or technologies that pose similar issues.[133]

6.2 Goals of ethical assessment

The purpose of the ethics component of the assessment phase[131] is to yield a series of questions, issues and comments to be integrated during ethical evaluation in the appraisal phase. Although this list does aim to be comprehensive, it remains likely that further issues will only become apparent during the appraisal phase. This two-step process in the identification of ethical issues thus provides the added advantage of providing a 'four-eyes' principle.

6.3 Ethical issues identified in assessment

The interventions involved in weight loss surgery are, for the most part, not new and have been used in other indications. The main ethical issues are thus not associated with the use of new interventions *per se* but with their increasing use in the treatment of obesity.[133] The main ethical issues are thus associated with approaches to obesity itself and increased use of these interventions. The main issues identified are the following:

6.3.1 Approaches to obesity

6.3.1.1 Obesity as a disease

Collective coverage of surgical approaches to obesity rests on the view that obesity is a disease, or at least a risk factor for disease differing perceptions could lead to disagreement regarding coverage for bariatric surgery. There are, broadly speaking, three types of views on obesity and whether it constitutes a disease:

Obesity is a disease, or at least a risk factor, and should be treated as such;

Obesity is a consequence of individual choices, which may in turn be subjected to varying degrees of influence[134];

Obesity is a personal characteristic, which is negatively viewed on aesthetic or moral grounds in our society and attempting to address it may represent a form of discrimination.[135]

Combinations of these views are of course possible. The most important aspect to note here is that – at the very least – the available data shows some BMI ranges to be clearly associated with surplus mortality and morbidity. Bariatric surgery, in addition, is shown to be effective not only in reducing weight but also in reversing diabetes and improving survival. Whether or not obesity rests on a loss of control by individuals may not be the most relevant factor in deciding whether to 'label' it a disease, since loss of control itself could be viewed as part of a disease process.[133 136] In any case, whatever the views adopted on obesity itself, the comorbidities associated with it are established diseases beyond controversy.[137]

6.3.1.2 Stigma

That obesity is negatively connoted in Switzerland and may expose affected individuals to stigma or even discrimination is clearly an ethically relevant aspect. Treatments what are able to reduce body weight in a short period of time could therefore be considered effective also in reducing the stigma an attached to patients as they become less obese. Some may argue on the other hand that providing surgery for weight loss may confirm negative connotations by underlining the badness of obesity, including for patients who do not wish to undergo surgery, and that this could be problematic. In the presence of clear data showing a marked health benefit, this would clearly not represent sufficient grounds to deny coverage.[133] However, it does point to the need for clarity in communicating the reasons underlying recommendations regarding weight loss surgery, with a view to hindering negative effects on collective representations of obesity.

6.3.1.3 Medicalisation

Viewing obesity as a disease can justify coverage for approaches designed to treat it and decrease negative connotations associated with it. This medicalization can, of course, also present disadvantages for persons 'who do not want to be sick, or who do not try to lose weight' and may increase pressures to accept bariatric surgery.[138] Saarni et al, who discuss this issue at some length, conclude that advantages of medicalization outweigh disadvantages for individuals.[133]

A different way to examine the medicalization of obesity is to recognize that, from a collective standpoint, it can be viewed as 'a surgical solution to the medical implications of a societal problem'.[139] It is important to note that data showing the effectiveness and cost-effectiveness of bariatric surgery all show it to be superior to conservative medical treatment. There are no studies, however, that compare it to alternative policy approaches, such as regulating food content or modifying the lived environment. This is clearly a limitation of the existing data. Policy approaches may work better, they may cost less: it seems we simply do not know. A study cited in one of the ethical commentaries on bariatric surgery suggests for example that restricting food advertisement on television « would be one of the most cost-effective interventions that governments can implement against childhood obesity ».[140 141] There is little evidence of how changes of food availability due to environmental changes and how changes of physical activity would affect the prevalence of obesity.[134 142 143] Currently, there are no studies comparing such approaches to bariatric surgery. Nor is there substantive research into «forms of pleasure that are non-addictive».[144 145] Other commentaries have noted the 'paradoxes associated with the mandate for neoliberal citizens to be economically productive when a form of this production is to consume (…) as good citizens, we must get the balance right. Ideally, our bodies demonstrate that we produce more than we consume.'.[146] Greater medicalization of obesity, which may follow greater uptake of weight loss surgery, should not come at the expense of awareness of its social dimensions.

6.3.1.4 **Personal responsibility for health**

One objection to covering bariatric surgery could be that obese persons are responsible for their own state and should not obtain publicly funded assistance in reversing their obesity. This argument relies on two premises, both of which are questionable.[147] First, it is not clear that obese persons can indeed be held responsible for their weight.[133 141] Second, even if it was accurate to hold them responsible, their responsibility would not clearly constitute grounds to withhold coverage from them.[139 148] Persson argues, on the contrary, that bariatric surgery for obese individuals should have a high priority given the important health risks run by obese individuals, and the obstacle to other life possibilities that obesity represents in our societies.[141]

6.3.2 Increased use of bariatric surgery

6.3.2.1 **Patient autonomy**

Free and informed consent to bariatric surgery can be predicted to be more difficult than for many other interventions. Bariatric surgery is not immediately life-saving, it is irreversible in most cases, and requires patients to change their eating habits permanently.[133] Strong societal norms against obesity may exert undue influence over candidates to consent or demand the intervention.[133 138] Nor can this difficulty be alleviated simply by advocating for the more reversible intervention, since gastric banding is also the least effective intervention according to the available data. Internet information is of variable quality.[139] Hype and aggressive marketing may also represent added difficulties in obtaining valid consent.[149 150] Obesity is associated with a number of psychiatric comorbidities, including anxiety, depression, and impaired self-esteem, leading to predictable situations where self-determination will be more difficult.[139]

It should be noted, however, none of these concerns are entirely specific to bariatric surgery and none are sufficient to throw doubt *a priori* on the ability of patients to make autonomous decisions regarding even difficult interventions.[151] They are, however, reasons to warrant special consideration when designing informed consent processes.

6.3.2.2 **Requirements before undergoing bariatric surgery**

Bariatric surgery is rarely performed on its own and ethical issues can be associated with the work-up before undergoing surgery and with follow-up. The main issues have been raised regarding mandatory weight-loss prior surgery. Weight-loss attempts before undergoing bariatric surgery are required to avoid unnecessary surgical interventions in patients who are capable to achieve significant weight loss with conservative treatment. Another reason for mandatory weight loss prior surgery in some individuals is required to reduce the surgical risk. This requirement, however, is perceived by some patients as a way to silence their request for help, as a futile exercise designed to promote weight consciousness in those already conscious, and sometimes as calling into question the indication for weight-loss surgery after it has, after all, been demonstrated that weight loss could be achieved without it.[152] It should be noted, however, that pre-operative dietary *counselling*, which is mandated by some third party payers and is distinct from mandatory weight loss, has also been defended as a mechanism for ensuring that candidates fully understand the requirements which undergoing surgery will place on them.[137]

6.3.2.3 **Operating children and adolescents**

This is a particularly difficult area.[153] [154] Parents, adolescents, and health care providers differ in their views of obesity and its impact on paediatric patients' lives, and consequently in their assessment of the risk-benefit ratio for bariatric surgery.[144] There is thus a risk for implicit or explicit coercion of minors to assent to bariatric surgery[155], especially as bariatric surgery may itself be viewed as a way of disciplining the behaviour of patients.[153] In short, bariatric surgery in the case of minors raises the same issues as other interventions, which have a long-lasting impact on a child who may or may not have the capacity to appreciate the consequences. In addition, data are scarcer for this population. Again, this is not specific to bariatric surgery.

6.3.2.4 **Unequal access**

Obesity correlates negatively with socioeconomic class. Greater medicalization of obesity as a condition would thus concern primarily disadvantaged populations within society.[138] This means that, again, advantages would outweigh disadvantages for them, but only in the case that access is distributed according to need rather than, say, ability to pay.[133] Even with inclusion in basic coverage, it must be noted that some patients are more likely than others to become candidates for an elective procedure designed to correct obesity. Members of disadvantaged groups are typically more likely to consult a physician only when ill, at which time they are unlikely to prioritize an elective procedure.[156] Increased access to bariatric surgery, if it is used on its own, may thus increase rather than limit the current social stratification of obesity.

6.3.2.5 **Risks associated with a commercialized innovation**

Bariatric surgery is subject to hype and it has been pointed out that there are specific risks associated with an innovative intervention at risk of hype:

As for many novel interventions, initial studies can yield results that are not directly transferable, because scale-up leads to uptake by more diverse patient and physician populations[157], leading to calls for processes enabling the oversight of the quality of surgeons in this field.[158]

These interventions can also be very lucrative. This can lead to over-promotion and there have been calls to limit the commercial promotion of not validated interventions in the field of bariatric surgery.[149]

Conflicts of interests due to strong ties between industry and surgeons in the field of bariatric surgery have also been evoked as representing a risk, since they may for example damage trust within the physician-patient relationship.[139]

Concerns have also been raised regarding the promotion of bariatric surgery abroad and the risks associated with medical tourism in search of interventions not covered in patients' country of residence.[159]

Finally, hype can lead to risk discounting and the risk-benefit ratio for an intervention, which involves operating on a well-functioning organ should be assessed with particular care.[160] This is of particular importance in the case of weight loss surgery, as the procedure is irreversible and may have different impacts at different stages of life, in particular as regards frailty in old age.[161]

In response to all these concerns, clearer evidence-based recommendations would thus represent an improvement, whatever their conclusion might be.

7 References

1. Picot J, Jones J, Colquitt JL, et al. The clinical effectiveness and cost-effectiveness of bariatric (weight loss) surgery for obesity: a systematic review and economic evaluation. Health Technol Assess 2009;**13**(41):1-212.
2. Bundesamt für Gesundheit (BAG). Cost of obesity in Switzerland in 2012, Last accessed October 1, 2015. Secondary Cost of obesity in Switzerland in 2012, Last accessed October 1, 2015. http://www.bag.admin.ch/themen/ernaehrung_bewegung/05207/05218/15444/index.html?lang=de.
3. Swiss Society for the Study of Morbid Obesity and Metabolic Disorders (SMOB). Richtlinien zur operativen Behandlung von Übergewicht (Medizinische Richtlinien). Medizinische Richtlinien zur operativen Behandlung von Übergewicht 2013; (Gültig ab: 1.1.2014). www.smob.ch.
4. Bundesamt für Gesundheit (BAG). Verordnung des EDI über Leistungen in der obligatorischen Krankenpflegeversicherung, Stand 15. Septermber 2015. Secondary Verordnung des EDI über Leistungen in der obligatorischen Krankenpflegeversicherung, Stand 15. Septermber 2015. https://www.admin.ch/opc/de/classified-compilation/19950275/index.html.
5. Bauknecht F. Bariatrische Operationen in der Schweiz 2001-2013. Secondary Bariatrische Operationen in der Schweiz 2001-2013. http://www.smob.ch/pdf/Bariatrie%20CH%202001-2013%20Auswertungen%20Website.pdf.
6. Clegg AJ, Colquitt J, Sidhu MK, et al. The clinical effectiveness and cost-effectiveness of surgery for people with morbid obesity: a systematic review and economic evaluation. Health Technol Assess 2002;**6**(12):1-153.
7. Bundesamt für Gesundheit (BAG). Health Technology Assessment: Effectiveness of preventive and therapeutic interventions in overweight and obesity. Follow-up report published 2007. Accessed October 29th, 2015, http://www.bag.admin.ch/themen/ernaehrung_bewegung/05207/05218/05231/index.html?lang=de. 2007.
8. Swiss Medical Board. Beurteilung medizinischer Verfahren.Methodischer Ansatz, Stand vom 30. April 2009. 2009; (October 1, 2014). http://www.medical-board.ch/fileadmin/docs/public/mb/methodik/Methodischer_Ansatz_2009_05_29.pdf (accessed October 1, 2014).
9. Swiss Medical Board. Medical Board: Literatursuche. Anhang zum methodischen Ansatz, Stand vom 28. Februar 2011. 2011; (1/10/2014). http://www.medical-board.ch/fileadmin/docs/public/mb/methodik/2011-02-28_Methodischer_Ansatz_Literatursuche.pdf (accessed 1/10/2014).
10. Wong SS, Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound treatment studies in EMBASE. Journal of the Medical Library Association : JMLA 2006;**94**(1):41-7.
11. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. www.cochrane.org/training/cochrane-handbook The Cochrane Collaboration last edited 20 March 2011
12. Andrews J, Guyatt G, Oxman AD, et al. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. J Clin Epidemiol 2013;**66**(7):719-25.
13. Andrews JC, Schunemann HJ, Oxman AD, et al. GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. J Clin Epidemiol 2013;**66**(7):726-35.
14. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011.
15. Brunetti M, Shemilt I, Pregno S, et al. Grade guidelines: 10. Considering resource use and rating the quality of economic evidence. J Clin Epidemiol 2012.

16. Guyatt G, Oxman AD, Akl E, et al. GRADE guidelines 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2010.
17. Guyatt G, Oxman AD, Kunz R, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. J Clin Epidemiol 2010.
18. Guyatt G, Oxman AD, Sultan S, et al. GRADE guidelines 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. J Clin Epidemiol 2012.
19. Guyatt G, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence-study limitations (risk of bias) and publication bias. J Clin Epidemiol 2011.
20. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. J Clin Epidemiol 2011;**64**(12):1283-93.
21. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 8. Rating the quality of evidence--indirectness. J Clin Epidemiol 2011;**64**(12):1303-10.
22. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 7. Rating the quality of evidence--inconsistency. J Clin Epidemiol 2011;**64**(12):1294-302.
23. Guyatt GH, Oxman AD, Montori V, et al. GRADE guidelines: 5. Rating the quality of evidence--publication bias. J Clin Epidemiol 2011;**64**(12):1277-82.
24. Guyatt GH, Oxman AD, Santesso N, et al. GRADE guidelines 12. Preparing Summary of Findings tables-binary outcomes. J Clin Epidemiol 2012.
25. Guyatt GH, Oxman AD, Schünemann HJ, et al. GRADE guidelines: A new series of articles in the Journal of Clinical Epidemiology. Journal of Clinical Epidemiology 2011;**64**(4):380-82.
26. Guyatt GH, Oxman AD, Sultan S, et al. GRADE guidelines: 9. Rating up the quality of evidence. JClinEpidemiol 2011;**64**(12):1311-16.
27. Guyatt GH, Thorlund K, Oxman AD, et al. GRADE guidelines: 13. Preparing Summary of Findings tables and evidence profiles - Continuous outcomes. Journal of Clinical Epidemiology 2013;**66**(2):173-83.
28. DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled clinical trials 1986;**7**(3):177-88.
29. Juni P, Reichenbach S, Dieppe P. Osteoarthritis: rational approach to treating the individual. Best practice & research Clinical rheumatology 2006;**20**(4):721-40.
30. Cohen J. *Statistical power analysis for the behavioral sciences. 2nd edition. Hillsdale, NJ: Lawrence Earlbaum Associates*, 1988.
31. Warkentin LM, Majumdar SR, Johnson JA, et al. Weight loss required by the severely obese to achieve clinically important differences in health-related quality of life: two-year prospective cohort study. BMC Med 2014;**12**:175.
32. Rücker G, Schwarzer G, Carpenter JR, et al. Undue reliance on I(2) in assessing heterogeneity may mislead. BMC Med Res Methodol 2008;**8**:79.
33. Courcoulas AP, Belle SH, Neiberg RH, et al. Three-Year Outcomes of Bariatric Surgery vs Lifestyle Intervention for Type 2 Diabetes Mellitus Treatment: A Randomized Clinical Trial. JAMA Surg 2015.
34. Courcoulas AP, Goodpaster BH, Eagleton JK, et al. Surgical vs medical treatments for type 2 diabetes mellitus: a randomized clinical trial. JAMA surgery 2014;**149**(7):707-15.
35. Courcoulas AP, Goodpaster BH, Eagleton JK, et al. Supplement: Surgical vs medical treatments for type 2 diabetes mellitus: a randomized clinical trial. JAMA Surg 2014. http://archsurg.jamanetwork.com/data/Journals/SURG/930571/soi140029.pdf.
36. Dixon JB, O'Brien PE, Playfair J, et al. Adjustable gastric banding and conventional therapy for type 2 diabetes: a randomized controlled trial. Jama 2008;**299**(3):316-23.
37. Dixon JB, Schachter LM, O'Brien PE, et al. Surgical vs conventional therapy for weight loss treatment of obstructive sleep apnea: a randomized controlled trial. Jama 2012;**308**(11):1142-49.
38. Feigel-Guiller B, Drui D, Dimet J, et al. Laparoscopic Gastric Banding in Obese Patients with Sleep Apnea: A 3-Year Controlled Study and Follow-up After 10 Years. Obes Surg 2015;**22**:22.
39. Ikramuddin S, Korner J, Lee W-J, et al. Roux-en-Y gastric bypass vs intensive medical management for the control of type 2 diabetes, hypertension, and hyperlipidemia: the Diabetes Surgery Study randomized clinical trial. Jama 2013;**309**(21):2240-49.

40. Ikramuddin S, Billington CJ, Lee W-J, et al. Roux-en-Y gastric bypass for diabetes (the Diabetes Surgery Study): 2-year outcomes of a 5-year, randomised, controlled trial. The Lancet Diabetes & Endocrinology 2015;**3**(6):413-22.
41. Mingrone G, Panunzi S, De Gaetano A, et al. Bariatric surgery versus conventional medical therapy for type 2 diabetes. New England Journal of Medicine 2012;**366**(17):1577-85.
42. Pokala S. ACP Journal Club. Gastric bypass or biliopancreatic diversion increases remission from type 2 diabetes in obese adults. Ann Intern Med 2012;**157**(2):Jc2-12.
43. O'Brien PE, Dixon JB, Laurie C, et al. Treatment of mild to moderate obesity with laparoscopic adjustable gastric banding or an intensive medical program: a randomized trial. Annals of internal medicine 2006;**144**(9):625-33.
44. Dixon JB, Strauss BJ, Laurie C, et al. Changes in body composition with weight loss: obese subjects randomized to surgical and medical programs. Obesity 2007;**15**(5):1187-98.
45. Reis L, Favaro W, Barreiro G, et al. Erectile dysfunction and hormonal imbalance in morbidly obese male is reversed after gastric bypass surgery: a prospective randomized controlled trial. International journal of andrology 2010;**33**(5):736-44.
46. Oliveira LV, Aguiar IC, Hirata RP, et al. Sleep study, respiratory mechanics, chemosensitive response and quality of life in morbidly obese patients undergoing bariatric surgery: a prospective, randomized, controlled trial. BMC surgery 2011;**11**:28.
47. Kashyap SR, Bhatt DL, Schauer PR. Bariatric surgery vs. advanced practice medical management in the treatment of type 2 diabetes mellitus: rationale and design of the Surgical Therapy And Medications Potentially Eradicate Diabetes Efficiently trial (STAMPEDE). Diabetes, obesity & metabolism 2010;**12**(5):452-4.
48. Kashyap SR, Bhatt DL, Wolski K, et al. Metabolic Effects of Bariatric Surgery in Patients With Moderate Obesity and Type 2 Diabetes Analysis of a randomized control trial comparing surgery with intensive medical treatment. Diabetes care 2013;**36**(8):2175-82.
49. Schauer PR, Bhatt DL, Kirwan JP, et al. Bariatric surgery versus intensive medical therapy for diabetes—3-year outcomes. New England Journal of Medicine 2014;**370**(21):2002-13.
50. Schauer PR, Kashyap SR, Wolski K, et al. Bariatric surgery versus intensive medical therapy in obese patients with diabetes. New England Journal of Medicine 2012;**366**(17):1567-76.
51. Malin SK, Samat A, Wolski K, et al. Improved acylated ghrelin suppression at 2 years in obese patients with type 2 diabetes: Effects of bariatric surgery vs standard medical therapy. International Journal of Obesity 2014;**38**(3):364-70.
52. Wentworth JM, Playfair J, Laurie C, et al. Gastric Band Surgery Leads to Improved Insulin Secretion in Overweight People with Type 2 Diabetes. Obes Surg 2015;**21**:21.
53. Wentworth JM, Playfair J, Laurie C, et al. Multidisciplinary diabetes care with and without bariatric surgery in overweight people: a randomised controlled trial. The Lancet Diabetes & Endocrinology 2014;**2**(7):545-52.
54. Benecke AV, Heiner. Übergewicht und Adipositas. Gesundheitsberichterstattung des Bundes Heft 16 2005. https://www.gbe-bund.de/pdf/Heft16.pdf.
55. WHO World Health Organization. The Asia-Pacific perspective: Redefining obesity and its treatment. February 2000. http://www.wpro.who.int/nutrition/documents/Redefining_obesity/en/.
56. Sjostrom L. Review of the key results from the Swedish Obese Subjects (SOS) trial - a prospective controlled intervention study of bariatric surgery. J Intern Med 2013;**273**(3):219-34.
57. National Institute for Health and Care Excellence. Section 5: Management of obesity 2 (continued). https://www.nice.org.uk/guidance/cg43/history, accessed January 26, 2016. Secondary Section 5: Management of obesity 2 (continued). https://www.nice.org.uk/guidance/cg43/history, accessed January 26, 2016 2006.
58. Bowne WB, Julliard K, Castro AE, et al. Laparoscopic gastric bypass is superior to adjustable gastric band in super morbidly obese patients: A prospective, comparative analysis. Arch Surg 2006;**141**(7):683-9.

59. Jan JC, Hong D, Pereira N, et al. Laparoscopic adjustable gastric banding versus laparoscopic gastric bypass for morbid obesity: a single-institution comparison study of early results. J Gastrointest Surg 2005;**9**(1):30-9; discussion 40-1.
60. Aarts EO, Dogan K, Koehestanie P, et al. Long-term results after laparoscopic adjustable gastric banding: a mean fourteen year follow-up study. Surg Obes Relat Dis 2014;**10**(4):633-40.
61. Peterli R, Donadini A, Peters T, et al. Re-operations following laparoscopic adjustable gastric banding. Obes Surg 2002;**12**(6):851-6.
62. Puzziferri N, Roshek TB, 3rd, Mayo HG, et al. Long-term follow-up after bariatric surgery: a systematic review. Jama 2014;**312**(9):934-42.
63. Chakravarty PD, McLaughlin E, Whittaker D, et al. Comparison of laparoscopic adjustable gastric banding (LAGB) with other bariatric procedures; a systematic review of the randomised controlled trials. Surgeon 2012;**10**(3):172-82.
64. Colquitt JL, Pickett K, Loveman E, et al. Surgery for weight loss in adults. Cochrane Database Syst Rev 2014;**8**(8).
65. Gloy VL, Briel M, Bhatt DL, et al. Bariatric surgery versus non-surgical treatment for obesity: a systematic review and meta-analysis of randomised controlled trials. BMJ 2013;**347**:f5934.
66. National Institute for Health and Care Excellence. Obesity: the prevention, identification, assessment and management of overweight and obesity in adults and children. https://www.nice.org.uk/guidance/cg43/history, accessed January 26, 2016.
67. National Institute for Health and Care Excellence. Obesity: Identification, assessment and management of overweight and obesity in children, young people and adults. https://www.nice.org.uk/guidance/cg189, accessed January 26, 2016.
68. Royle P, Waugh N. Literature searching for clinical and cost-effectiveness studies used in health technology assessment reports carried out for the National Institute for Clinical Excellence appraisal system. Health Technol Assess 2003;**7**(34):iii, ix-x, 1-51.
69. Clegg A, Colquitt J, Sidhu M, et al. Clinical and cost effectiveness of surgery for morbid obesity: a systematic review and economic evaluation. Int J Obes Relat Metab Disord 2003;**27**(10):1167-77.
70. Goeree R, He J, O'Reilly D, et al. Transferability of health technology assessments and economic evaluations: a systematic review of approaches for assessment and application. ClinicoEconomics and outcomes research: CEOR 2011;**3**:89.
71. O'Brien BJ. A tale of two (or more) cities: geographic transferability of pharmacoeconomic data. Am J Manag Care 1997;**3 Suppl**:S33-9.
72. Welte R, Feenstra T, Jager H, et al. A decision chart for assessing and improving the transferability of economic evaluation results between countries. Pharmacoeconomics 2004;**22**(13):857-76.
73. Drummond M, Barbieri M, Cook J, et al. Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force report. Value in health 2009;**12**(4):409-18.
74. OECD. Health Statistics 2015 - Frequently Requested Data - Total health expenditure per capita, US$ PPP. Secondary Health Statistics 2015 - Frequently Requested Data - Total health expenditure per capita, US$ PPP. http://www.oecd.org/els/health-systems/oecd-health-statistics-2014-frequently-requested-data.htm.
75. OECD. Health Statistics 2015 - Purchasing Power Parities for GDP and related indicators. Secondary Health Statistics 2015 - Purchasing Power Parities for GDP and related indicators. http://stats.oecd.org/.
76. Office SFS. Kosten der Gesundheitswesen seit 1960. Secondary Kosten der Gesundheitswesen seit 1960 2015 http://www.bfs.admin.ch/bfs/portal/de/index/themen/14/01/new/nip_detail.html?gnpID=2014-095.
77. Office SFS. Schweizerischer Lohnindex aufgrund der Daten der Sammelstelle für die Statistik der Unfallversicherung (SSUV) 2014. Secondary Schweizerischer Lohnindex aufgrund der Daten der Sammelstelle für die Statistik der Unfallversicherung (SSUV) 2014 2015. http://www.bfs.admin.ch/bfs/portal/de/index/themen/03/04.html.

78. FSO. Swiss Health Survey - Factors influencing health – Data, indicators: Body weight. Secondary Swiss Health Survey - Factors influencing health – Data, indicators: Body weight. http://www.bfs.admin.ch/bfs/portal/en/index/themen/14/02/02/key/02.html.
79. Buri M. STA490: Health risks and social significance of BMI ≥ 35 kg/m2 in Switzerland. Master program in Biostatistics, University of Zurich. June 2015.
80. Ackroyd R, Mouiel J, Chevallier JM, et al. Cost-effectiveness and budget impact of obesity surgery in patients with type-2 diabetes in three European countries. Obes Surg 2006;**16**(11):1488-503.
81. Briggs A, Claxton K, Sculpher M. Decision Modelling for Health Economic Evaluation. 2006.
82. Ananthapavan J, Moodie M, Haby M, et al. Assessing cost-effectiveness in obesity: laparoscopic adjustable gastric banding for severely obese adolescents. Surg Obes Relat Dis 2010;**6**(4):377-85.
83. Anselmino M, Bammer T, Fernandez Cebrian JM, et al. Cost-effectiveness and budget impact of obesity surgery in patients with type 2 diabetes in three European countries(II). Obes Surg 2009;**19**(11):1542-9.
84. Borg S, Näslund I, Persson U, et al. Obesity and Surgical Treatment – A Cost-Effectiveness Assessment for Sweden. Nordic Journal of Health Economics 2014;**2**(1).
85. Campbell J, McGarry LA, Shikora SA, et al. Cost-effectiveness of laparoscopic gastric banding and bypass for morbid obesity. Am J Manag Care 2010;**16**(7):e174-87.
86. Castilla I, Mar J, Valcarcel-Nazco C, et al. Cost-utility analysis of gastric bypass for severely obese patients in Spain. Obes Surg 2014;**24**(12):2061-8.
87. Craig BM, Tseng DS. Cost-effectiveness of gastric bypass for severe obesity. Am J Med 2002;**113**(6):491-8.
88. Faria GR, Preto JR, Costa-Maia J. Gastric bypass is a cost-saving procedure: results from a comprehensive Markov model. Obes Surg 2013;**23**(4):460-6.
89. Hoerger TJ, Zhang P, Segel JE, et al. Cost-effectiveness of bariatric surgery for severely obese adults with diabetes. Diabetes Care 2010;**33**(9):1933-9.
90. Ikramuddin S, Klingman D, Swan T, et al. Cost-Effectiveness of Roux-en-Y Gastric Bypass in Type 2 Diabetes Patients. American Journal of Managed Care 2009;**15**(9):607-15.
91. Jensen C, Flum DR, Conference ABSC. The costs of nonsurgical and surgical weight loss interventions: is an ounce of prevention really worth a pound of cure? Surg Obes Relat Dis 2005;**1**(3):353-7.
92. Keating CL, Dixon JB, Moodie ML, et al. Cost-efficacy of surgically induced weight loss for the management of type 2 diabetes: a randomized controlled trial. Diabetes Care 2009;**32**(4):580-4.
93. Maklin S, Malmivaara A, Linna M, et al. Cost-utility of bariatric surgery for morbid obesity in Finland. Br J Surg 2011;**98**(10):1422-9.
94. McEwen LN, Coelho RB, Baumann LM, et al. The cost, quality of life impact, and cost-utility of bariatric surgery in a managed care population. Obes Surg 2010;**20**(7):919-28.
95. Michaud PC, Goldman DP, Lakdawalla DN, et al. The value of medical and pharmaceutical interventions for reducing obesity. J Health Econ 2012;**31**(4):630-43.
96. Padwal R, Klarenbach S, Wiebe N, et al. Bariatric surgery: a systematic review of the clinical and economic evidence. J Gen Intern Med 2011;**26**(10):1183-94.
97. Picot J, Jones J, Colquitt JL, et al. Weight loss surgery for mild to moderate obesity: a systematic review and economic evaluation. Obes Surg 2012;**22**(9):1496-506.
98. Pollock RF, Muduma G, Valentine WJ. Evaluating the cost-effectiveness of laparoscopic adjustable gastric banding versus standard medical management in obese patients with type 2 diabetes in the UK. Diabetes, obesity & metabolism 2013;**15**(2):121-9.
99. Salem L, Devlin A, Sullivan SD, et al. Cost-effectiveness analysis of laparoscopic gastric bypass, adjustable gastric banding, and nonoperative weight loss interventions. Surgery for Obesity and Related Diseases 2008;**4**(1):26-32.
100. Wang BC, Wong ES, Alfonso-Cristancho R, et al. Cost-effectiveness of bariatric surgical procedures for the treatment of severe obesity. Eur J Health Econ 2014;**15**(3):253-63.
101. Borg S, Näslund I, Persson U, et al. Budget impact analysis of surgical treatment for obesity in Sweden. Scandinavian Journal of Surgery 2012;**101**(3):190-97.
102. Neovius K, Neovius M, Kark M, et al. Association between obesity status and sick-leave in Swedish men: nationwide cohort study. The European Journal of Public Health 2012;**22**(1):112-16.

103. Arterburn DE, Maciejewski ML, Tsevat J. Impact of morbid obesity on medical expenditures in adults. International journal of obesity 2005;**29**(3):334-39.
104. Thompson D, Edelsberg J, Colditz GA, et al. Lifetime health and economic consequences of obesity. Archives of Internal Medicine 1999;**159**(18):2177-83.
105. Parikh MS, Laker S, Weiner M, et al. Objective comparison of complications resulting from laparoscopic bariatric procedures. Journal of the American College of Surgeons 2006;**202**(2):252-61.
106. Nguyen NT, Goldman C, Rosenquist CJ, et al. Laparoscopic versus open gastric bypass: a randomized study of outcomes, quality of life, and costs. Annals of surgery 2001;**234**(3):279.
107. Heshka S, Anderson JW, Atkinson RL, et al. Weight loss with self-help compared with a structured commercial program: a randomized trial. Jama 2003;**289**(14):1792-98.
108. Crémieux P-Y, Buchwald H, Shikora SA, et al. A study on the economic impact of bariatric surgery. Am J Manag Care 2008;**14**(9):589-96.
109. Shekelle P, et al. Pharmacological and Surgical Treatment of Obesity.Evidence Report/Technology Assessment No. 103. AHRQ Publication No 04- E028-2. Rockville, MD: Agency for Healthcare Research and Quality.
110. Gallagher SF, Banasiak M, Gonzalvo JP, et al. The impact of bariatric surgery on the Veterans Administration healthcare system: a cost analysis. Obesity surgery 2003;**13**(2):245-48.
111. O'Brien PE, McPhail T, Chaston TB, et al. Systematic review of medium-term weight loss after bariatric operations. Obesity Surgery 2006;**16**(8):1032-40.
112. Sjöström CD, Peltonen M, Wedel H, et al. Differentiated long-term effects of intentional weight loss on diabetes and hypertension. Hypertension 2000;**36**(1):20-25.
113. Sjöström L, Lindroos A-K, Peltonen M, et al. Lifestyle, diabetes, and cardiovascular risk factors 10 years after bariatric surgery. New England Journal of Medicine 2004;**351**(26):2683-93.
114. Sjöström L, Narbro K, Sjöström CD, et al. Effects of bariatric surgery on mortality in Swedish obese subjects. New England journal of medicine 2007;**357**(8):741-52.
115. Dixon S, Currie CJ, McEwan P. Utility values for obesity and preliminary analysis of the Health Outcomes Data Repository. Exp Rev Pharmacoecon Outcomes Res 2004;**4**:657–65.
116. Buchwald H, Estok R, Fahrbach K, et al. Weight and type 2 diabetes after bariatric surgery: systematic review and meta-analysis. The American journal of medicine 2009;**122**(3):248-56. e5.
117. Pories WJ, Swanson MS, MacDonald KG, et al. Who would have thought it? An operation proves to be the most effective therapy for adult-onset diabetes mellitus. Annals of surgery 1995;**222**(3):339.
118. Haby M, Vos T, Carter R, et al. A new approach to assessing the health benefit from obesity interventions in children and adolescents: the assessing cost-effectiveness in obesity project. International journal of obesity 2006;**30**(10):1463-75.
119. Angrisani L, Lorenzo M, Borrelli V. Laparoscopic adjustable gastric banding versus Roux-en-Y gastric bypass: 5-year results of a prospective randomized trial. Surgery for obesity and related diseases 2007;**3**(2):127-32.
120. McCormack K, Wake B, Perez J, et al. Laparoscopic surgery for inguinal hernia repair: systematic review of effectiveness and economic evaluation. Health Technol Assess 2005;**9**(14):1-203, iii-iv.
121. Hakim Z, Wolf A, Garrison LP. Estimating the effect of changes in body mass index on health state preferences. Pharmacoeconomics 2002;**20**(6):393-404.
122. Bagust A, Beale S. Modelling EuroQol health-related utility values for diabetic complications from CODE-2 data. Health economics 2005;**14**(3):217-30.
123. Colagiuri S, Colagiuri R, Conway BG, D, et al. DiabCo$t Australia: assessing the burden of type 2 diabetes in Australia [article online]. Secondary DiabCo$t Australia: assessing the burden of type 2 diabetes in Australia [article online] 2003. http://www.diabetes.nsw.com.au/PDFs/About_Diabetes_PDFs/diabcost_finalreport.
124. Currie CJ, Poole CD, Woehl A, et al. The health-related utility and health-related quality of life of hospital-treated subjects with type 1 or type 2 diabetes with particular reference to differing severity of peripheral neuropathy. Diabetologia 2006;**49**(10):2272-80.

125. Lee A, Morgan CL, Morrissey M, et al. Evaluation of the association between the EQ-5Dindex (health-related utility) and body mass index (obesity) in hospital-treated people with Type 1 diabetes, Type 2 diabetes and with no diagnosed diabetes. Diabetic Medicine 2005;**22**(11):1482-86.
126. Wong ES, Wang B, Alfonso-Cristancho R, et al. BMI trajectories among the severely obese: results from an electronic medical record population. Obesity 2012;**20**(10):2107-12.
127. Arterburn D, Ichikawa L, Ludman EJ, et al. Validity of clinical body weight measures as substitutes for missing data in a randomized trial. Obesity research & clinical practice 2008;**2**(4):277-81.
128. Franks P, Lubetkin EI, Gold MR, et al. Mapping the SF-12 to preference-based instruments: convergent validity in a low-income, minority population. Medical Care 2003;**41**(11):1277-83.
129. Beauchamp TL, Childress JF. *Principles of Biomedical Ethics*. fifth ed. Oxford: Oxford University Press, 2009.
130. Daniels N, Light DW, Caplan RL. *Benchmarks of Fairness for Health Care Reform*. Oxford: Oxford University Press, 1996.
131. INAHTA Ethics Working Group. INAHTA's Working Group on Handling Ethical Issues, 2005.
132. Tavaglione N, Martin AK, Mezger N, et al. Fleshing out Vulnerability. Bioethics 2015;**29**(2):98-107.
133. Saarni SI, Anttila H, Saarni SE, et al. Ethical issues of obesity surgery--a health technology assessment. Obes Surg 2011;**21**(9):1469-76.
134. Garrett JR, McNolty LA. Bariatric surgery and the social character of the obesity epidemic. The American journal of bioethics : AJOB 2010;**10**(12):20-2.
135. Gard M, Wright J. *The obesity epidemic - Science, morality and ideology*. New York: Routledge, 2005.
136. Fisher CE, Devlin MJ, Appelbaum PS. Parsing neurobiological dysfunctions in obesity: nosologic and ethical consequences. The American journal of bioethics : AJOB 2010;**10**(12):14-6.
137. Brandon AR, Puzziferri N, Sadler JZ. Stuck in the middle: what should a good society do? The American journal of bioethics : AJOB 2010;**10**(12):18-20.
138. de Vries J. The obesity epidemic: medical and ethical considerations. Science and engineering ethics 2007;**13**(1):55-67.
139. Hofmann B. Stuck in the middle: the many moral challenges with bariatric surgery. The American journal of bioethics : AJOB 2010;**10**(12):3-11.
140. Magnus A, Haby MM, Carter R, et al. The cost-effectiveness of removing television advertising of high-fat and/or high-sugar food and beverages to Australian children. Int J Obes (Lond) 2009;**33**(10):1094-102.
141. Persson K. Why Bariatric surgery should be given high priority: an argument from law and morality. Health care analysis : HCA : journal of health philosophy and policy 2014;**22**(4):305-24.
142. Huang TTK, Cawley JH, Ashe M, et al. Mobilisation of public support for policy actions to prevent obesity. The Lancet;**385**(9985):2422-31.
143. Caballero B. The global epidemic of obesity: an overview. Epidemiologic reviews 2007;**29**:1-5.
144. van Geelen SM, Bolt IL, van der Baan-Slootweg OH, et al. The controversy over pediatric bariatric surgery: an explorative study on attitudes and normative beliefs of specialists, parents, and adolescents with obesity. Journal of bioethical inquiry 2013;**10**(2):227-37.
145. Mackenzie R. The neuroethics of pleasure and addiction in public health strategies: Moving beyond harm reduction: Funding the creation of non-addictive drugs and taxonomies of pleasure. Neuroethics 2010;**doi: 10.1007/s12152–010-9074–4**.
146. Mackenzie R. Don't let them eat cake! A view from across the pond. The American journal of bioethics : AJOB 2010;**10**(12):16-8.
147. Vartanian LR, Smyth JM. Primum non nocere: obesity stigma and public health. Journal of bioethical inquiry 2013;**10**(1):49-57.
148. Cappelen AW, Norheim OF. Responsibility, fairness and rationing in health care. Health Policy 2005.
149. Shikora SA. A call for maintaining ethical behavior in bariatric surgery. Obes Surg 2012;**22**(6):849-50.
150. Hofmann B, Hjelmesaeth J, Sovik TT. Moral challenges with surgical treatment of type 2 diabetes. Journal of diabetes and its complications 2013;**27**(6):597-603.
151. Zuraw R. Exceptionalism denied: obesity does not negate the ability to give informed consent. The American journal of bioethics : AJOB 2010;**10**(12):27-8.

152. Glenn NM, Raine KD, Spence JC. Mandatory weight loss during the wait for bariatric surgery. Qualitative health research 2015;**25**(1):51-61.
153. Hofmann B. Bariatric surgery for obese children and adolescents: a review of the moral challenges. BMC medical ethics 2013;**14**:18.
154. Caniano DA. Ethical issues in pediatric bariatric surgery. Seminars in pediatric surgery 2009;**18**(3):186-92.
155. Pratt JS, Lenders CM, Dionne EA, et al. Best practice updates for pediatric/adolescent weight loss surgery. Obesity (Silver Spring) 2009;**17**(5):901-10.
156. McGee S. Getting unstuck: rubber bands and public health. The American journal of bioethics : AJOB 2010;**10**(12):1-2.
157. De Ville K. Bariatric surgery, ethical obligation, and the life cycle of medical innovation. The American journal of bioethics : AJOB 2010;**10**(12):22-4.
158. Dixon JB, Logue J, Komesaroff PA. Promises and ethical pitfalls of surgical innovation: the case of bariatric surgery. Obes Surg 2013;**23**(10):1698-702.
159. Snyder J, Crooks VA. Medical tourism and bariatric surgery: more moral challenges. The American journal of bioethics : AJOB 2010;**10**(12):28-30.
160. Hofmann B. Parachutes for diabetes: bariatric surgery beyond evidence? Diabetes Res Clin Pract 2012;**98**(3):406-7.
161. Busetto L, Dixon J, De Luca M, et al. Bariatric surgery in class I obesity : a Position Statement from the International Federation for the Surgery of Obesity and Metabolic Disorders (IFSO). Obes Surg 2014;**24**(4):487-519.

8 Appendix A – Clinical effectiveness and safety

8.1 Search strategy

Appendix 1 Search strategy with number of records found, stratified by databased searched

Central

Search Name: HTA Bariatric Surgery - population AND intervention 2
Last Saved: 30/06/2015 10:27:45.516
Description:

ID	Search
#1	weight reduc*:ti,ab,kw
#2	weight loss:ti,ab,kw
#3	MeSH descriptor: [Weight Loss] explode all trees
#4	(over weight or overweight or overeating or over eating):ti,ab,kw
#5	obesity:ti,ab,kw
#6	MeSH descriptor: [Obesity] explode all trees
#7	obese:ti,ab,kw
#8	MeSH descriptor: [Overweight] explode all trees
#9	MeSH descriptor: [Weight Reduction Programs] explode all trees
#10	{or #1-#9}
#11	MeSH descriptor: [Bariatric Surgery] explode all trees
#12	bariatric surg*:ti,ab,kw
#13	(surg* near/5 bariatric):ti,ab,kw
#14	(metabol* near/5 bariatric):ti,ab,kw
#15	(anti obesity surg*):ti,ab,kw
#16	(antiobesity surg*):ti,ab,kw
#17	(weight near/5 surg*):ti,ab,kw
#18	(anti-obesity surg*):ti,ab,kw
#19	(overweight near/5 surg*):ti,ab,kw
#20	(obesity near/5 surg*):ti,ab,kw
#21	MeSH descriptor: [Gastric Bypass] explode all trees
#22	(gastric near/5 surgery):ti,ab,kw
#23	(gastric near/5 bypass):ti,ab,kw
#24	(restrictive near/5 surgery):ti,ab,kw
#25	MeSH descriptor: [Gastroplasty] explode all trees
#26	(gastroplasty or gastrogastrostom*):ti,ab,kw
#27	MeSH descriptor: [Digestive System Surgical Procedures] explode all trees
#28	(digestive and surg*):ti,ab,kw
#29	(gastrointestinal and surg*):ti,ab,kw
#30	(digestive and bypass):ti,ab,kw

#31 (intestinal and surg*):ti,ab,kw
#32 (gastrointestinal and diversion):ti,ab,kw
#33 (gastrointestinal and bypass):ti,ab,kw
#34 MeSH descriptor: [Biliopancreatic Diversion] explode all trees
#35 (biliopancreatic and (diversion or surg* or bypass)):ti,ab,kw
#36 (bilio pancreatic and (diversion or surg* or bypass)):ti,ab,kw
#37 (gastric and (sleeve* or band*)):ti,ab,kw
#38 (silicon and (sleeve* or band*)):ti,ab,kw
#39 LAGB:ti,ab,kw
#40 MeSH descriptor: [Gastrectomy] explode all trees
#41 gastrectom*:ti,ab,kw
#42 gastroenterostom*:ti,ab,kw
#43 MeSH descriptor: [Gastroenterostomy] explode all trees
#44 (lap* and band*):ti,ab,kw
#45 stomach stapl*:ti,ab,kw
#46 Roux en Y:ti,ab,kw
#47 RYGB:ti,ab,kw
#48 MeSH descriptor: [Anastomosis, Roux-en-Y] explode all trees
#49 malabsorptive surg*:ti,ab,kw
#50 malabsorptive procedure*:ti,ab,kw
#51 duodenal switch:ti,ab,kw
#52 mason* procedure:ti,ab,kw
#53 MeSH descriptor: [Obesity] explode all trees and with qualifier(s): [Surgery - SU]
#54 MeSH descriptor: [Obesity, Morbid] explode all trees and with qualifier(s): [Surgery - SU]
#55 {or #11-#54}
#56 #10 and #55

Hits: 2360; thereof 1937 trials

Medline via OvidSP

Ovid MEDLINE(R) 1946 to June Week 4 2015, Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations July 02, 2015, Ovid MEDLINE(R) Daily Update July 02, 2015

1. obesity.ti,ab.
2. exp obesity/
3. obese.ti,ab.
4. exp Overweight/
5. exp weight reduction programs/
6. Overweight/
7. over weight.ti,ab.
8. overweight.ti,ab.
9. (overeating or over eating).ti,ab.

10. exp weight loss/
11. weight loss.ti,ab.
12. weight reduc$.ti,ab.
13. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12
14. exp Bariatric Surgery/
15. bariatric surgeries.ti,ab.
16. bariatric surgery.ti,ab.
17. (surg$ adj5 bariatric).ti,ab.
18. (metabol$ adj5 surg$).ti,ab.
19. anti?obesity surg$.ti,ab.
20. (weight adj5 surg$).ti,ab.
21. (overweight adj5 surg$).ti,ab.
22. (obesity adj5 surg$).ti,ab.
23. exp Gastric Bypass/
24. (gastric adj5 bypass).ti,ab.
25. (gastric adj5 surgery).ti,ab.
26. (restrictive adj5 surgery).ti,ab.
27. exp Gastroplasty/
28. (gastroplasty or gastro?gastrostom*).ti,ab.
29. exp Digestive System Surgical Procedures/
30. (digestive and surg$).ti,ab.
31. (gastrointestinal and surg$).ti,ab.
32. (gastrointestinal and bypass).ti,ab.
33. (intestinal and surg$).ti,ab.
34. (gastrointestinal and diversion).ti,ab.
35. exp Biliopancreatic Diversion/
36. ((biliopancreatic or bilio?pancreatic or bilio pancreatic) and (diversion or surg$ or bypass)).ti,ab.
37. gastric band$.ti,ab.
38. silicon band$.ti,ab.
39. LAGB.ti,ab.
40. exp Gastrectomy/
41. gastrectom*.ti,ab.
42. (lap$ and band$).ti,ab.
43. stomach stapl$.ti,ab.
44. Obesity/su [Surgery]
45. Obesity, Morbid/su [Surgery]
46. Roux en Y.ti,ab.

47. RYGB.ti,ab.
48. Anastomosis, Roux-en-Y/
49. malabsorptive surg$.ti,ab.
50. duodenal switch.ti,ab.
51. ((gastric or silicon) and sleeve).ti,ab.
52. exp Gastroenterostomy/
53. gastroenterostom*.ti,ab.
54. mason$ procedure.ti,ab.
55. malabsorptive procedure$.ti,ab.
56. 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55
57. randomized controlled trial.pt.
58. controlled clinical trial.pt.
59. randomized.ab.
60. randomised.ab.
61. placebo.ab.
62. clinical trials as topic.sh.
63. randomly.ab.
64. random*.tw.
65. trial.ti.
66. 57 or 58 or 59 or 60 or 61 or 62 or 63 or 64 or 65
67. 13 and 56 and 66

Hits: 1716

Embase

Embase 1974 to 2015 July 02
Hits 2560

1. exp obesity/
2. obese.ti,ab.
3. obesity.ti,ab.
4. overweight.ti,ab.
5. over weight.ti,ab.
6. (overeating or over eating).ti,ab.
7. exp weight reduction/
8. weight loss.ti,ab.
9. weight reduc$.ti,ab.

10. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9
11. exp bariatric surgery/
12. bariatric surgeries.ti,ab.
13. bariatric surgery.ti,ab.
14. (surg$ adj5 bariatric).ti,ab.
15. (surg$ adj5 metabol$).ti,ab.
16. anti?obesity surg$.ti,ab.
17. (weight adj5 surg$).ti,ab.
18. (overweight adj5 surg$).ti,ab.
19. (obesity adj5 surg$).ti,ab.
20. exp stomach bypass/
21. (stomach adj5 bypass).ti,ab.
22. (gastric adj5 bypass).ti,ab.
23. (gastric adj5 surgery).ti,ab.
24. (restrictive adj5 surgery).ti,ab.
25. exp gastroplasty/
26. (gastroplasty or gastro?gastrostom*).ti,ab.
27. (digestive adj5 surg$).ti,ab.
28. (gastrointestinal and surg$).ti,ab.
29. (intestinal and surg$).ti,ab.
30. (gastrointestinal and bypass).ti,ab.
31. (gastrointestinal and diversion).ti,ab.
32. exp biliopancreatic bypass/
33. ((biliopancreatic or bilio?pancreatic or bilio pancreatic) and bypass).ti,ab.
34. ((biliopancreatic or bilio?pancreatic or bilio pancreatic) and diversion).ti,ab.
35. ((biliopancreatic or bilio?pancreatic or bilio pancreatic) and surg$).ti,ab.
36. exp gastric banding/
37. gastric band$.ti,ab.
38. ((gastric or silicon) and band$).ti,ab.
39. silicon band$.ti,ab.
40. LAGB.ti,ab.
41. exp gastrectomy/
42. gastrectom*.ti,ab.
43. (lap$ and band$).ti,ab.
44. stomach stapl$.ti,ab.
45. exp obesity/su [Surgery]
46. exp morbid obesity/su [Surgery]
47. Roux en Y.ti,ab.

48. RYGB.ti,ab.
49. exp Roux Y anastomosis/
50. malabsorptive surg$.ti,ab.
51. duodenal switch.ti,ab.
52. exp gastroenterostomy/
53. gastroenterostom*.ti,ab.
54. mason$ procedure.ti,ab.
55. malabsorptive procedure$.ti,ab.
56. random$.tw. or placebo$.mp. or double-blind$.tw.
57. 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55
58. 10 and 56 and 57

Pubmed

COMPLETE SEARCH

(obesity or morbid obesity or overweight or weight reduction programs or overeating or weight loss or weight reduc*)

AND (bariatric surgery OR antiobesity surgery OR metabolic surgery OR weight surgery OR overweight surgery OR obesity surgery OR gastric bypass OR gastric surgery OR restrictive surgery OR gastroplasty OR gastrogastrostom* OR digestive system surgical procedures OR digestive surgery OR gastrointestinal surgery OR gastrointestinal bypass OR intestinal surgery OR gastrointestinal diversion OR biliopancreatic diversion OR biliopancreatic surgery OR biliopancreatic bypass OR gastric band OR silicon band OR lagb OR gastrectom* OR lap band OR stomach stapling OR stomach staples OR “roux en y” OR rygb OR obesity/surgery OR morbid obesity/surgery OR malabsorptive surgery OR duodenal switch OR gastroenterostom* OR gastric sleeve OR gastric silicon OR mason procedure OR malabsorptive procedure)

AND (randomized controlled trial OR controlled clinical trial OR randomized OR randomised OR placebo OR clinical trials as topic OR random OR trial OR rct)

AND publisher[sb]

HITS: 89

8.2 Study population at baseline – 2 years results

Appendix 2 Study population at baseline

Study identification	Study groups	Female % (n)	Mean age (SD), years	BMI <35 kg/m², % patients	Mean BMI (SD)	Mean body weight (SD), kg	Type 2 diabetes, % patients	Mean duration of type 2 diabetes, years	Mean HbA1c, % (SD)	Metabolic syndrome, % patients	Hypertension, % patients	Dyslipidemia, % patients	Other comorbidities, % patients
Courcoulas 2014	RYGB	79.2% (19/24)	46.3 (7.2)	43% (26/61)*	35.5 (2.6)	99.8 (12.8)	100% (24/24)	7.4 (4.5)	8.7 (2.2)	not reported	50% (12/24), history of hypertension	58% (14/24), history of dyslipidemia/hypercholesterolemia	not reported
	LAGB	81.8% (18/22)	47.3 (7.0)	43% (26/61)*	35.5 (3.4)	99.5 (14.1)	100% (22/22)	6.1 (4.3)	7.9 (2.2)	not reported	59% (13/22), history of hypertension	73% (16/22), history of dyslipidemia/hypercholesterolemia	not reported
	Cons. treatment	82.6% (19/23)	48.3 (4.7)	43% (26/61)*	35.7 (3.3)	102.6 (13.8)	100% (23/23)	5.7 (5.6)	7.0 (0.8)	not reported	70% (16/23), history of hypertension	65% (15/23), history of dyslipidemia/hypercholesterolemia	not reported
Dixon 2008	LAGB	50.0% (15/30)	46.6 (7.4)	20.0% (6/30)	37.0 (2.7)	105.6 (13.8)	100% (30/30)	not reported	7.8 (1.2)	96.7% (29/30)	93.3% (28/30)	not reported	Coronary artery disease: 0% (0/30)
	Cons. treatment	56.7% (17/30)	47.1 (8.7)	23.3% (7/30)	37.2 (2.5)	105.9 (14.2)	100% (30/30)	not reported	7.6 (1.4)	96.7% (29/30)	90.0% (27/30)	not reported	Coronary artery disease: 3.3% (1/30)

Study identification	Study groups	Female % (n)	Mean age (SD), years	BMI <35 kg/m², % patients	Mean BMI (SD)	Mean body weight (SD), kg	Type 2 diabetes, % patients	Mean duration of type 2 diabetes, years	Mean HbA1c, % (SD)	Metabolic syndrome, % patients	Hypertension, % patients	Dyslipidemia, % patients	Other comorbidities, % patients
Dixon 2012	Surgical group 1, LAGB	43.3% (13/30)	47.4 (8.8)	0% (0/30)	46.3 (6.0)	134.9 (22.1)	33.3% (10/30)	not reported	6.25 (1.1)	63.3% (19/30)	50.0% (15/30)	not reported	Obstructive sleep apnoea: 100% (30/30) Depression: 40% (12/30)
	conservative treatment	40.0% (12/30)	50.0 (8.2)	0% (0/30)	43.8 (4.9)	126.0 (19.3)	33.3% (10/30)	not reported	6.26 (1.1)	80.0% (24/30)	56.7% (17/30)	not reported	Obstructive sleep apnoea: 100% (30/30) Depression: 29.7% (11/30)
Feigel-Guiller 2015	LAGB	46.7% (14/30)	46.9 (8.6)	0% (0/30)	48.8 (9.9)	135.0 (25.3)	33.3% (10/30), history of diabetes mellitus	not reported	not reported	not reported	50.0% (15/30), history of hypertension	not reported	Obstructive sleep apnoea (OSA): 63% (19/33) Obesity-hypoventilation syndrome OHS: 10% (3/33) OSA and OHS: 27% (8/33)

Study identification	Study groups	Female % (n)	Mean age (SD), years	BMI <35 kg/m², % patients	Mean BMI (SD)	Mean body weight (SD), kg	Type 2 diabetes, % patients	Mean duration of type 2 diabetes, years	Mean HbA1c, % (SD)	Metabolic syndrome, % patients	Hypertension, % patients	Dyslipidemia, % patients	Other comorbidities, % patients
	conservative treatment	39.4% (13/33)	50.1 (7.4)	0% (0/33)	44.4 (9.0)	123.0 (25.1)	36.4% (12/33), history of diabetes mellitus	not reported	not reported	not reported	69.7% (23/33), history of hypertension	not reported	Obstructive sleep apnoea (OSA): 58% (19/33) Obesity-hypoventilation syndrome OHS: 3% (1/33) OSA and OHS: 39% (13/33)
Ikramuddin 2015	RYGB	63.3% (38/60)	49 (9)	60.0% (36/90)	34.9 (3.0)	98.8 (14.0)	98.3% (after randomisation one patient was determined to have type 1 diabetes)	8.9 (6.1)	9.6 (1.0)	not reported	not reported	not reported	not reported
	conservative treatment	56.7% (34/60)	49 (8)	58.3% (35/60)	34.3 (3.1)	97.9 (17.0)	100%	9.1 (5.6)	9.6 (1.2)	not reported	not reported	not reported	not reported
Mingrone 2012	RYGB	60.0% (12/20)	43.9 (7.6)	0% (0/20)	44.9 (5.2)	129.8 (22.6)	100% (20/20)	6.0 (1.2)	8.6 (1.4)	not reported	not reported	not reported	not reported
	BPD	50.0% (10/20)	42.8 (8.1)	0% (0/20)	45.1 (7.8)	137.9 (30.4)	100% (20/20)	6.0 (1.3)	8.9 (1.7)	not reported	not reported	not reported	not reported

Study identification	Study groups	Female % (n)	Mean age (SD), years	BMI <35 kg/m², % patients	Mean BMI (SD)	Mean body weight (SD), kg	Type 2 diabetes, % patients	Mean duration of type 2 diabetes, years	Mean HbA1c, % (SD)	Metabolic syndrome, % patients	Hypertension, % patients	Dyslipidemia, % patients	Other comorbidities, % patients
	conservative treatment	50.0% (10/20)	43.5 (7.3)	0% (0/20)	45.6 (6.2)	136.4 (21.9)	100% (20/20)	6.1 (1.2)	8.5 (1.2)	not reported	not reported	not reported	not reported
O'Brien 2006	LAGB	75.0% (30/40)	41.8 (6.4)	100% (40/40)	33.7 (1.8)	96.1 (11.9)	not reported	not reported	not reported	37.5% (15/40)	22.5% (9/40)	not reported	coronary artery disease 0% (0/40)
	conservative treatment	77.5% (31/40)	40.7 (7.0)	100% (40/40)	33.5 (1.4)	93.6 (11.9)	not reported	not reported	not reported	37.5% (15/40)	17.5% (7/40)	not reported	coronary artery disease 0% (0/40)
Reis 2010	RYGB	0% (0/10)	36.7 (11.5)	0% (0/10)	55.7 (7.8), 43.1 (4.7)**	168.6 (28.2), 130.6 (18.3)**	not reported	not reported	not reported	not reported	not reported	not reported	Erectile dysfunction 100% (10/10)
	conservative treatment	0% (0/10)	42.2 (11.0)	0% (0/10)	54.0 (6.1), 51.9 (5.7)**	160.4 (20.1), 154.2 (19.3)**	not reported	not reported	not reported	not reported	not reported	not reported	Erectile dysfunction 100% (10/10)
Schauer 2014	RYGB	58% (29/50)	48.3 (8.4)	28% (14/50)	37.0 (3.3)	106.7 (14.8)	100% (50/50)	8.2 (5.5)	9.3 (1.4) (n=50)	90% (45/50)	70% (35/50), history of hypertension	88% (44/50), history of dyslipidemia	not reported

Study identification	Study groups	Female % (n)	Mean age (SD), years	BMI <35 kg/m², % patients	Mean BMI (SD)	Mean body weight (SD), kg	Type 2 diabetes, % patients	Mean duration of type 2 diabetes, years	Mean HbA1c, % (SD)	Metabolic syndrome, % patients	Hypertension, % patients	Dyslipidemia, % patients	Other comorbidities, % patients
	SG	78% (39/50)	47.9 (8.0)	36% (18/50)	36.2 (3.9)	100.8 (16.4)	100% (50/50)	8.5 (4.8)	9.5 (1.7) (n=49)	94% (47/50)	60% (30/50), history of hypertension	80% (40/50), history of dyslipidemia	not reported
	conservative treatment	62% (31/50)	49.7 (7.4)	38% (19/50)	36.8 (3.0)	106.5 (14.7)	100% (50/50)	8.9 (5.8)	8.9 (1.4) (n=41)	92% (46/50)	60% (26/43), history of hypertension	84% (36/43), history of dyslipidemia	not reported
Wentworth 2014	LAGB	76% (19/25)	53 (6)	100% (25/25)	29 (1)	81 (10)	100% (25/25)	2.2 (1.7)	6.9% (1.2)	not reported	not reported	not reported	not reported
	conservative treatment	65.4% (17/26)	53 (7)	100% (26/26)	29 (1)	83 (12)	100% (26/26)	2.8 (1.8)	7.2% (1.1)	not reported	not reported	not reported	not reported

Study identification	Study groups	Female % (n)	Mean age (SD), years	BMI <35 kg/m², % patients	Mean BMI (SD)	Mean body weight (SD), kg	Type 2 diabetes, % patients	Mean duration of type 2 diabetes, years	Mean HbA1c, % (SD)	Metabolic syndrome, % patients	Hypertension, % patients	Dyslipidemia, % patients	Other comorbidities, % patients
Abbreviation BMI: body mass index BPD: Biliopancreatic diversion LAGB: Laparoscopic adjustable gastric banding RYGB: Roux-en-Y gastric bypass SD: standard deviation SG: Sleeve gastrectomy * given only for the total study population who underwent treatment **This is the time point before RYGB and 4 months after extensive weight-loss program in the bariatric surgery group.													

8.3 Eligibility criteria in the included RCTs

Appendix 3 Eligibility criteria of the included RCTs

RCT	Inclusion criteria	Exclusion criteria
Courcoulas 2014	"Adults were eligible for enrollment if they were 25 to 55 years of age and had a BMI of 30 to 40 because these characteristics represent a high-priority subgroup for comparative effectiveness studies. Diagnosis of T2DM was confirmed by a documented fasting plasma glucose (FPG) level of 126 mg/dL or greater (to convert to millimoles per liter, multiply by 0.0555) and/or treatment with antidiabetics to include a broad spectrum of T2DM severity. For participants with grade I obesity, treatment with antidiabetics and permission from their treating physician were required to participate."	"Adults were excluded for prior weight loss surgery, impaired mental status, alcohol or other drug addiction, current smoking, pregnancy or planned pregnancy, inability to tolerate general anesthesia owing to poor health, type 1 diabetes mellitus, failed nutritional or psychological assessment, unwillingness to be randomized, inability to provide informed consent, or being deemed unlikely to comply with study visits or procedures."
Ding 2015	"Eligible participants were aged 21–65 years with at least 1-year of T2D, BMI 30–45 kg/m2, a strong desire for substantial weight loss and commitment to life-long medical followup, and free from active cardiovascular or eye diseases prohibiting them from exercising safely or undergoing a bariatric surgical procedure. Additionally, potential participants had HbA1c above7%regardless of ongoing treatment, or 6.5% on two oral antihyperglycemic medications or insulin, and with stable-dose treatment(s) for more than eight weeks."	"Participants were excluded for detectable glutamic acid decarboxylase antibody; history of diabetic ketoacidosis; HbA1c 12%; gastrointestinal (GI) diseases; malignancy within 5-years; significant cardiopulmonary or renal diseases; active eating disorders; drug and/or alcohol abuse; impaired mental status; weight loss 3% within the previous 3-months; participation in an alternate weight-reduction program or using weight-reduction medications. Participants had to be nonsmoking for over 2-months. Supplemental Material - Additional Information on Exclusion Criteria: Uncontrolled T2D was defined as consistent fasting blood glucose >200 mg/dl or HbA1c above twice normal. Gastrointestinal disease exclusions included previous major gastrointestinal surgery, inflammatory bowel disease, esophageal diseases including severe intractable esophagitis, Barrett's Disease, esophageal dysmotility or other impaired gastric motility (gastroparesis), or hiatal hernia >3

		cm in size, chronic or acute bleeding conditions including peptic ulcer disease, portal hypertension (gastric or esophageal varices), chronic pancreatitis, or cirrhosis of the liver; malignancy (including personal or family history of Medullary Thyroid Carcinoma (MTC) or Multiple Endocrine Neoplasia syndrome type 2 (MEN 2)), except participants who have been disease free for greater than 5 years, or whose only malignancy has been basal or squamous cell skin carcinoma, or debilitating medical conditions, severe cardiopulmonary disease including uncontrolled hypertension (repeated systolic measures >160 or diastolic > 95 mm Hg on more than one day), unstable angina pectoris; recent myocardial infarction, history of coronary artery bypass surgery or angioplasty within 6 months; congestive heart failure, arrhythmia, stroke or transient ischemic attacks, sustained urinary albumin excretion >1000 mcg/mg creatinine, or serum creatinine ≥132.6 μmol/L (1.5mg/dL) with creatinine clearance <60 ml/min by Cockcroft-Gault equation and eGFR <60 ml/1.73 m2/min, by Modification of Diet in Renal Disease (MDRD) study equation, any endocrine disorder other than T2D or thyroid disease which is stable on replacement therapy, including Cushing's syndrome; any history of eating disorder unless it was in stable remission for over 3 years although binge eating symptoms were not exclusionary, history of drug and/or alcohol abuse within 2 years of the screening visit, history of impaired mental status by DSM4 (Diagnostic and Statistical Manual, 4th Edition) criteria and including, but not limited to active substance abuse, a history of schizophrenia, borderline personality disorder, uncontrolled depression, suicidal attempts within the past two years or current suicidal tendencies or ideations. Subjects were excluded for significant weight loss (>3%) within the previous 3 months, or participation in alternate medically supervised exercise or weight reduction program within the previous 3 months, or use of prescription or over the counter weight reduction medications or supplements within one month of the Screening Visit and for the duration of study participation. Subjects needed to be nonsmoking for ≥ 2 months. Women who were lactating, planning pregnancy, or unwilling to use contraception during the course of the trial were excluded.
Dixon 2008	"Patients were eligible if they were aged between 20 and 60	Candidates were excluded if they had a history of type 1 diabetes,

	years, had a body mass index of 30 to 40, had been diagnosed with clearly documented type 2 diabetes within the previous 2 years, had no evidence of renal impairment or diabetic retinopathy, and were able to understand and comply with the study process."	diabetes secondary to a specific disease, or previous bariatric surgery; a history of medical problems such as mental impairment, drug or alcohol addiction, recent major vascular event, internal malignancy, or portal hypertension; or a contraindication for either study group. Participants were excluded if they did not attend 2 initial information visits.
Dixon 2012	"The inclusion criteria were patients aged 18 to 60 years, body mass index (calculated as weight in kilograms divided by height in meters squared) of 35 to 55, apnea-hypopnea index (AHI) of 20 events/hour or more diagnosed within the previous 6 months with recommendation to commence CPAP therapy,13 and at least 3 prior significant weight loss attempts."	"The exclusion criteria were previous bariatric surgery, obesity hypoventilation syndrome requiring bilevel positive airway pressure, and contraindications to bariatric surgery including cognitive impairment, drug or alcohol addiction, and significant cardiopulmonary, neurological, vascular, gastrointestinal, or neoplastic disease."
Feigel-Guiller 2015	"Patients eligible for inclusion in the RCT were aged 18– 65 years, had BMIs >35 kg/m2 2 months before study inclusion, and were receiving nocturnal NIV treatment for OSA and/or obesity-hypoventilation syndrome (OHS).OSA requiring NIV was defined by an AHI >30 events/h on polysomnography, and OHS requiring NIV was defined by a partial pressure of carbon dioxide in arterial blood (PaCO2) >6.5 kPa, according to French guidelines. NIV was provided using continuous positive airway pressure, bi-level positive airway pressure, or volumetric ventilation, as recommended by pulmonologists at the onset of the study."	"Patients with contraindications for surgery or severe eating disorders were excluded."
Halperin 2014	"Eligible participants were aged 21 to65 years with at least 1 year of type 2 diabetes, BMI 30 to 42, a strong desire for substantial weight loss, and a commitment to life-long medical and nutritional follow-up. They were free from active cardiovascular or other diseases prohibiting them from exercising safely or undergoing a bariatric surgical procedure.Additionally, potential participants had HbA1c levels above 7%(to convert to a proportion of total Hb, multiply by 0.01), regardless of ongoing treatment, or 6.5% or greater while receiving either 2 oral antihyperglycemic agents at greater than or equal to half-maximal dose or insulin, and with stable-dose treatment formore than 8 weeks."	"Individuals were excluded if they had detectable levels of antiglutamic acid decarboxylase antibody, a history of diabetic ketoacidosis, uncontrolled type 2 diabetes (HbA1c >12%), gastrointestinal disease, malignant disease within 5 years, significant cardiopulmonary or renal disease, active eating disorder, drug and/or alcohol abuse, impaired mental status, weight loss greater than 3% within the previous 3 months, participation in another weight reduction program, or were using weight-reduction medications and/or supplements. Participants had to be nonsmoking for more than 2 months. [...] People with a preference for a bariatric procedure other than RYGB were not enrolled."
Heindorff	BMI >38 kg/m2	not reported

1997		
Ikramuddin 2015	"Age 30 to 67 years at eligibility visit. Diagnosed with T2DM at least 6 months prior to enrollment, under the active care of a doctor for at least the six months prior to enrollment, HbA1c ≥ 8.0 % and HbA1c ≤ 14.0%. Body Mass Index (BMI) ≥ 30.0 kg/m2 and ≤ 39.9 kg/m2 at eligibility visit. Willingness to accept random assignment to either treatment group. Expect to live or work within approximately one hour's traveling time from the study clinic for the duration of the twoyear trial. Willingness to comply with the followup protocol and successful completion of the runin (described below). Written informed consent."	"Cardiovascular event (myocardial infarction, acute coronary syndrome, coronary artery angioplasty or bypass, stroke) in the past six months. Current evidence of congestive heart failure, angina pectoris, or symptomatic peripheral vascular disease. Cardiac stress test indicating that surgery or IMM would not be safe. Pulmonary embolus or thrombophlebitis in the past six months. Cancer of any kind (except basal cell skin cancer or cancer in situ) unless documented to be diseasefree for five years. Significant anemia (hemoglobin 1.0 g/dL or more below normal range) or history of coagulopathy. Serum creatinine greater than 1.5 mg/dl. Serum total bilirubin greater than the upper limit of normal in the absence of Gilbert's syndrome, or alkaline phosphatase or ALT greater than twice the upper limit of normal. History of stomach surgery, bile duct surgery, pancreatic surgery, splenectomy, or colon resection. Gastric or duodenal ulcer in the past six months. History of intra-abdominal sepsis (except for uncomplicated appendicitis or diverticulitis more than six months prior to enrollment). Previous organ transplantation. Self-reported HIVpositive status, active tuberculosis, active malaria, chronic hepatitis B or C, cirrhosis, or inflammatory bowel disease. Currently pregnant or nursing, or planning to become pregnant in the next two years. History of alcohol or drug dependency (excluding caffeine and nicotine) in the past five years. Active psychosocial or psychiatric problem that is likely to interferer with adherence to the protocol. Depression: A score of 17 or higher may be used to disqualify a participant. At the clinic's discretion, participants with a CESD score of 17 or higher may be referred to a licensed psychologist or psychiatrist for a formal psychological evaluation. In that case, the formal psychological evaluation and recommendation will be important factors in the determination of eligibility by the eligibility committee. If this evaluation clearly indicates potential risks associated with depression, the participant may not be randomized. If the evaluation is equivocal or unclear, the eligibility committee must take this into account in rendering their decision. Current participation in a conflicting research protocol. Presence of any chronic or debilitating

		disease that would make adherence to the protocol difficult. 12lead EKG indicating that surgery would not be safe. Serum Cpeptide less than 1.0 ng/ml 90 minutes post-challenge. Exclusions may also be made at the discretion of the attending physician or the eligibility committee."
Liang 2013	"The study recruited individuals with T2DM diagnosed according to WHO criteria. Other inclusion criteria were: (1) obesity (body mass index [BMI] > 28 kg/m2) in accordance with the WHO Asia-Pacific classification for obesity [9]; (2) T2DM with hypertension of 5–10 years with hypertension defined as systolic blood pressure (SBP) 140 mmHg and/or diastolic (DBP) 90 mmHg as per 1999 WHO/ISH criteria; (3) insulin therapy in combination with oral administration of drugs for 12 months; (4) glycated hemoglobin (HbA1c) > 7%; (5) age: 30–60 years; (6) seronegative for antibodies against insulin, islet cells and glutamic acid decarboxylase (GAD); (7) C-peptide level 0.3 mg/L."	"Exclusion criteria included: (1) people without diabetes; (2) type 1 diabetes mellitus, presence of autoimmune diabetes indicated by antibodies to insulin, islet cells, and GAD, and gestational diabetes; (3) patients with heart, liver, or renal function impairment; (4) presence of severe infections or cerebrovascular disease; (5) fasting serum insulin was less than one-third of the normal value; (6) diabetes of more than 10 years duration; (7) age > 60 years or <30 years."
Mingrone 2002	not reported ("morbidly obese"; "They showed a normal electrocardiogram at rest and during an exercise test." and "age: 30–45 years")	"pregnancy; history or diagnosis of diabetes, heart disease, hypertension, or other chronic diseases; hormone replacement therapy; chronic steroid therapy; a history of alcohol or drug abuse; and glucose intolerance, defined as a 2 h glucose level of 140 mg/dl after a 75 g oral glucose load and of stable weight (within 2 kg, 6 months before testing)."
Mingrone 2012	"Inclusion criteria were an age of 30 to 60 years, a body-mass index (BMI, the weight in kilograms divided by the square of the height in meters) of 35 or more, a history of type 2 diabetes of at least 5 years, a glycated hemoglobin level of 7.0% or more (as confirmed by at least three analyses), and an ability to understand and comply with the study protocol."	"Exclusion criteria were a history of type 1 diabetes, diabetes secondary to a specific disease or glucocorticoid therapy, previous bariatric surgery, pregnancy, other medical conditions requiring short-term hospitalization, severe diabetes complications, other severe medical conditions, and geographic inaccessibility."
O'Brien 2006	"We considered patients to be eligible if they were between 20 and 50 years of age; had a body mass index of 30 kg/m2 to 35 kg/m2; had identifiable problems, including an obesity-related comorbid condition (such as hypertension, dyslipidemia, diabetes, obstructive sleep apnea, or gastroesophageal reflux disease), severe physical limitations, or clinically significant psychosocial problems	"We excluded candidates with a history of bariatric surgery or medical problems that contraindicated treatment in either study group, such as impaired mental status, drug or alcohol addiction, or portal hypertension. In addition, we excluded participants if they had undergone an intensive, physician-supervised program that used very-low-calorie diets or pharmacotherapy or if they did not attend the 2 initial patient information visits."

	associated with their obesity; had attempted to reduce weight over at least the previous 5 years; could understand the options offered and the randomization process; and were willing to comply with the requirements of each program."	
Parikh 2014	"[...] patients with T2DM and BMI 30 to 35 who were otherwise eligible for bariatric surgery by NIH criteria, specifically (1) overweight for at least 5 years, (2) failure to lose weight with nonsurgical means, (3) absence of medical or psychological contraindications, (4) patient understanding of the procedure and its risks, and (5) strong motivation to comply with the postsurgical regimen."	"Patients were excluded if they were deemed unable to comply with the study protocol (either self-selected or by indicating during screening that s/he could not complete all requested tasks), participation in other obesity- or diabetes-related clinical trials, or diagnosis of cognitive dysfunction or significant psychiatric comorbidity."
Reis 2010	not reported	"Exclusion criteria were co-morbidities requiring regular drug usage (statin, antihypertensive, oral anti-diabetic), endocrine disease (except mild hypogonadism) or recent hormonal manipulation (thyroid / other hormonall reposition / block in the last 3 months), testicular impairment, previous history of alcohol or tobacco abuse and phosphodiesterase type-5 inhibitor usage. Detailed medical history and physical and laboratory examination concerning hormonal, biochemical and metabolic parameters were routinely performed in every patient prior to the randomization. Oral glucose tolerance test was performed when fasting plasma glucose was inconclusive or borderline."
Schauer 2014	"Male and female subjects between the ages of 20 and 60-years-old with clinical diagnosis of T2DM and BMI between 27 and 43 kg/m2 will be included in this study. Biochemical evidence of T2DM with treated HbA1c of ≥7.0% must be present. The subject must be a candidate for general anesthesia and be able to understand the options to comply with the requirement of each intervention program. Non-pregnant and non-lactating female patients must agree to use a reliable method of contraception for 2 years."	"Patients were excluded if they had undergone previous bariatric surgery or other complex abdominal surgery or had poorly controlled medical or psychiatric disorders."
Wentworth 2014	"Inclusion criteria were age between 18 and 65 years, BMI between 25 and 30 kg/m², diabetes duration less than 5 years, willingness to be randomised to either study group, and ability to comply with the treatment protocol."	"Exclusion criteria were positive glutamic acid decarboxylase autoantibody titre, pancreatic disease, previous bariatric surgery, or contraindication to laparoscopic adjustable gastric banding (including previous upper gastrointestinal surgery, hypothalamic disease, pregnancy, history of psychosis, or myocardial infarct in the preceding 6 months). If HbA1c was less than 6·5% (48 mmol/mol),

		diabetes was confi rmed by oral glucose tolerance test."

8.4 Description of the interventions in the conservative treatment group

Appendix 4 Description of the interventions in the conservative treatment group

Study identification	Bariatric surgery group 1	Bariatric surgery group 2	Non-surgical treatment group:			
			In general	**Medication(s)**	**Diet**	**Physical activity**
Courcoulas 2014	RYGB (laparoscopic not reported)	LAGB	Lifestyle weight loss intervention, a standard behavioural weight control program	not reported	Energy-restricted diet (1200-1800 kcal/d), and provision of meal plans, meal replacements, and calorie-counter books	Prescription of moderate intensity exercise 20-60 min per day at 5 days per week
Dixon 2008	LAGB and lifestyle modification as for non-surgical treatment group	Not applicable	Lifestyle modification programs were individually structured	Medical therapies on an individual basis	Reduction of energy intake, intake of fat (<30%) and saturated fats, and increase of intake of low glycemic index and high-fiber foods	Physical activity of moderate intensity was encouraged: 10000 steps per day and 200 minutes per week of structured activity
Dixon 2012	LAGB and the intensity, and nature of the lifestyle program as for non-surgical treatment group	Not applicable	Best available medical practice for the treatment, education, and follow-up of severely obese patients	not reported	Dietary advice including calories reduction	Encouragement of physical activity: walking and 200 minutes/week of structured activity
Feigel-Guiller 2015	LAGB and low energy intake (1400 kcal/day) physical exercise as for non-surgical treatment group	Not applicable	Not applicable	not reported	low-energy (1400 kcal/day) diet	advised to perform physical exercise

Study identification	Bariatric surgery group 1	Bariatric surgery group 2	Non-surgical treatment group:			
			In general	Medication(s)	Diet	Physical activity
Ikramuddin 2015	Laparoscopic RYGB and increase of physical activity to a total of 325 minutes per week	Not applicable	Lifestyle modification	Orlistat or sibutramine could be added to the lifestyle modification when program was unsuccessful to decrease body weight And medications for glycemic control	Calorie intake targets of 1200, 1500, or 1800 kilocalories per day, depending on body weight	Advise to physical activity (such as walking) to a total of 325 minutes per week.
Mingrone 2012	Laparoscopic RYGB	Biliopancreatic diversion (BPD)	Programs for diet and lifestyle modification	Oral hypoglycemic agents and insulin doses	Reduction of overall energy and fat intake (<30% total fat, <10% saturated fat, and high fiber content)	Increase of physical exercise (≥30 minutes every day, twice aerobic activity a week)
O'Brien 2006	LAGB	Not applicable	Behavioural modification program: very-low-calorie diet, pharmacotherapy with education and professional support on appropriate eating and exercise behaviour.	120 mg of orlistat before non–very-low-calorie diet meals, and then 120 mg of orlistat before all meals until the completion of the intensive phase	Very-low-calorie diet (500 to 550 kcal/d) using Optifast daily for 12 weeks, followed by a transition phase over 4 weeks	Encouraged of exercise for at least 200 minutes per week.
Reis 2010	Laparoscopic RYGB	Not applicable	No specific individualized programme was offered to participants. At the end of the study, the same intervention and surgery	not reported	Provision of oral and written information about healthy food choices	Provision of oral and written general guidance on increasing their level of physical activity

Study identificat ion	Bariatric surgery group 1	Bariatric surgery group 2	Non-surgical treatment group:			
			In general	Medication(s)	Diet	Physical activity
			were offered to all subjects in the control group.			
Schauer 2014	Laparoscopic RYGB and lifestyle counselling, weight management, frequent home glucose monitoring, and the use of newer drug therapies	Sleeve gastrectomy (SG) and lifestyle counselling, weight management, frequent home glucose monitoring, and the use of newer drug therapies	Intensive medical therapy: lifestyle counselling, weight management, frequent home glucose monitoring, and the use of newer drug therapies	not reported	not reported	not reported
Wentworth 2014	LAGB and glycemic control, calorie intake and physical activity advise as for the non-surgical treatment group.	Not applicable	not reported	Medications for glycemic control	Calorie-restricted diet to all participants	Advise of physical activity of 150 min each week.
Abbreviation BPD: Biliopancreatic diversion LAGB: Laparoscopic adjustable gastric banding RYGB: Roux-en-Y gastric bypass SG : Sleeve gastrectomy						

8.5 Results of risk of bias and support of judgment ot the 2 year results

Appendix 5 Results of risk of bias and support of judgment of the 2-year results

Study	**Random sequence generation (selection bias) and support for judgement**		**Allocation concealment (selection bias) and support for judgement**		**Blinding of participants and personnel (performance bias) and support for judgement**		**Blinding of outcome assessment (detection bias) and support for judgement**		**Incomplete continuous outcome data (attrition bias) and support for judgement**		**Incomplete binary data (attrition bias) and support for judgement**		**Selective reporting (reporting bias) and support for judgement**	
Courcoulas 2015	low	"computer-generated"	unclear	randomization by computer but "randomization envelopes" used	unclear	not reported	unclear	not reported	high	missing data >20% in either study arm	high	missing data >20% in either study arm	low	results were given for all outcomes mentioned in method paragraph
Dixon 2008	low	"Randomization was computer derived, with blocking into 3	unclear	not reported	high	"The study was not blinded."	high	"The study was not blinded."	high	missing data 10% - 20 %, adressed using inadeq	high	missing data 10% - 20 %, adressed using adequa	low	results were given for all outcomes mention ed in

Study	Random sequence generation (selection bias) and support for judgement		Allocation concealment (selection bias) and support for judgement		Blinding of participants and personnel (performance bias) and support for judgement		Blinding of outcome assessment (detection bias) and support for judgement		Incomplete continuous outcome data (attrition bias) and support for judgement		Incomplete binary data (attrition bias) and support for judgement		Selective reporting (reporting bias) and support for judgement	
		groups to allow for orderly recruitment into both study groups and to reduce the risk of uneven recruitment late in the series."								uate methods		te methods, but not comparable among study arms		method paragraph
Dixon 2012	low	"prior to computer-derived randomization."	unclear	not reported	unclear	not reported	high	not for our outcomes of interest ("diagnostic polysomnography scored by staff blinded to	low	missing data ≤10% and comparable among study arms	low	missing data ≤10% and comparable among study arms	low	results were given for all outcomes mentioned in method

Study	Random sequence generation (selection bias) and support for judgement		Allocation concealment (selection bias) and support for judgement		Blinding of participants and personnel (performance bias) and support for judgement		Blinding of outcome assessment (detection bias) and support for judgement		Incomplete continuous outcome data (attrition bias) and support for judgement		Incomplete binary data (attrition bias) and support for judgement		Selective reporting (reporting bias) and support for judgement	
								randomization")						paragraph apart from hip circumference
Feigel-Guiller 2015	unclear	not reported	unclear	not reported	unclear	not reported	unclear	not reported	high	missing data 10% - 20 %, adressed using inadequate methods	high	missing data 10% - 20 %, adressed using inadequate methods	low	results were given for all outcomes mentioned in method paragraph
Ikramuddin 2015	low	"We generated all randomisation assignments in advance	unclear	Sequence was generated in advance. Publication from 2013 states that allocation	high	"Randomization assignment was unblinded but allocation between	low	Publication 2013: Investigators, data collectors, and outcome adjudicator	low	missing data ≤10% and comparable among study	low	missing data ≤10% and comparable among study	low	results were given for all outcomes mentioned in

Study	Random sequence generation (selection bias) and support for judgement		Allocation concealment (selection bias) and support for judgement		Blinding of participants and personnel (performance bias) and support for judgement		Blinding of outcome assessment (detection bias) and support for judgement		Incomplete continuous outcome data (attrition bias) and support for judgement		Incomplete binary data (attrition bias) and support for judgement		Selective reporting (reporting bias) and support for judgement	
		using a pseudorandom number generator in SAS (version 9.3)."		was concealed but no information how this was achieved.		treatment groups was concealed to the study staff until after randomization"		s were blinded to aggregate outcomes until the final patient completed the 12-month follow-up.Publication 2015: "Staff from the clinical centres had access to data for their individual patients, but were masked to data for other patients		arms		arms		method paragraph

Study	Random sequence generation (selection bias) and support for judgement		Allocation concealment (selection bias) and support for judgement		Blinding of participants and personnel (performance bias) and support for judgement		Blinding of outcome assessment (detection bias) and support for judgement		Incomplete continuous outcome data (attrition bias) and support for judgement		Incomplete binary data (attrition bias) and support for judgement		Selective reporting (reporting bias) and support for judgement	
								and to aggregated data until all 2-year follow-up data had been obtained. The aggregate data was not disclosed until the 2-year visits were completed"						
Mingrone 2012	low	"by simple randomization in a 1:1:1 ratio with the use of a	unclear	"Patients will be assigned to bariatric surgery or standard care by means of	high	"In this single-center, nonblinded, randomized, controlle	high	"Here we report the results of a single-center, nonblinded, randomized, controlled	low	missing data ≤10% and comparable among study	low	missing data ≤10% and comparable among study	low	results were given for all outcomes mentioned in

Study	Random sequence generation (selection bias) and support for judgement		Allocation concealment (selection bias) and support for judgement		Blinding of participants and personnel (performance bias) and support for judgement		Blinding of outcome assessment (detection bias) and support for judgement		Incomplete continuous outcome data (attrition bias) and support for judgement		Incomplete binary data (attrition bias) and support for judgement		Selective reporting (reporting bias) and support for judgement	
		computerized system for generating random numbers."		computer-based randomization procedure Computer-generated random lists will be used, with a 1:1:1 randomization ratio."		d trial, 60 patients between the ages..."		trial"		arms		arms		method paragraph
O'Brien 2006	low	"A computer-derived random allocation sequence, without blocking or stratification, performed the	low	"After assessment had confirmed a participant's suitability for randomization,the coordinator contacted the trial office by	high	"The study was not blinded."	high	"The study was not blinded."	high	missing data >20% in either study arm	not applicable (no relevant binary outcome identified)	not applicable	low	results were given for all outcomes mentioned in method paragraph

Study	Random sequence generation (selection bias) and support for judgement		Allocation concealment (selection bias) and support for judgement		Blinding of participants and personnel (performance bias) and support for judgement		Blinding of outcome assessment (detection bias) and support for judgement		Incomplete continuous outcome data (attrition bias) and support for judgement		Incomplete binary data (attrition bias) and support for judgement		Selective reporting (reporting bias) and support for judgement	
		randomization."		telephone for allocation."										
Reis 2010	low	"centralized computer-generated randomization"	low	"centralized computer-generated randomization"	unclear	not reported	unclear	not reported (only for outcome not relevant for our scope)	low	missing data ≤10% and comparable among study arms	low	missing data ≤10% and comparable among study arms	low	results were given for all outcomes mentioned in method paragraph
Schauer 2014	low	from protocol: "The randomization scheme will be developed by a statisticia	unclear	from protocol: "Sealed and sequentially numbered envelopes containing the treatment assignment	high	"nonblinded design"	high	"nonblinded design"	high	missing data >20% in either study arm	high	missing data >20% in either study arm	low	results were given for all outcomes mentioned in method paragra

Study	Random sequence generation (selection bias) and support for judgement		Allocation concealment (selection bias) and support for judgement		Blinding of participants and personnel (performance bias) and support for judgement		Blinding of outcome assessment (detection bias) and support for judgement		Incomplete continuous outcome data (attrition bias) and support for judgement		Incomplete binary data (attrition bias) and support for judgement		Selective reporting (reporting bias) and support for judgement	
		n at the Cleveland Clinic Coordinating Center for Clinical Research (C5Research). A blocking sequence will be used and subjects will be stratified according to diabetes severity, defined as insulin-requiring		will be provided to the nurse coordinator responsible for patient enrollment. Once eligibility is confirmed, the nurse will select and open the next sealed envelope in sequence to reveal the treatment assignment. Envelopes must be selected in sequence and not to										ph

Study	Random sequence generation (selection bias) and support for judgement		Allocation concealment (selection bias) and support for judgement		Blinding of participants and personnel (performance bias) and support for judgement		Blinding of outcome assessment (detection bias) and support for judgement		Incomplete continuous outcome data (attrition bias) and support for judgement		Incomplete binary data (attrition bias) and support for judgement		Selective reporting (reporting bias) and support for judgement	
		at the time of screening."		be opened prior randomization."										
Wentworth 2014	low	"via a computer-generated random sequence"	low	"A third party assigned treatment groups with numberd envelopes, which were opened by the participant immediately after providing written consent to join the study"	high	"We did an open-label, parallel-group, randomised controlled trial between Nov 1, 2009, and June 30, 2013, at one centre in Melbourne"	high	Clinical data were collected as previously described by study nurses (JP and CL) who were not blinded to treatment group. HbA1c, lipid, and OGTT glucose levels and urine albumin/creatinine	low	missing data ≤10% and comparable among study arms	low	missing data ≤10% and comparable among study arms	low	results were given for all outcomes mentioned in method paragraph

Study	Random sequence generation (selection bias) and support for judgement		Allocation concealment (selection bias) and support for judgement		Blinding of participants and personnel (performance bias) and support for judgement		Blinding of outcome assessment (detection bias) and support for judgement		Incomplete continuous outcome data (attrition bias) and support for judgement		Incomplete binary data (attrition bias) and support for judgement		Selective reporting (reporting bias) and support for judgement	
								ratio were performed by Melbourne Pathology (Abbotsford, Australia).						

8.6 Studies with a follow-up of 6 months to 1 year

8.6.1 List of studies with a follow-up of 6 months to 1 year

Appendix 6 List of studies with a follow-up of 6 months to 1 year

Studies which were not included in meta-analysis because of follow-up of <2 years	
Study identification	**Reference**
Ding 2015	Ding SA, Simonson DC, Wewalka M, et al. (2015). Adjustable Gastric Band Surgery or Medical Management in Patients with Type 2 Diabetes: a Randomized Clinical Trial. The Journal of clinical endocrinology and metabolism jc20151443.
Halperin 2014	Halperin F, Ding SA, Simonson DC, et al. (2014). Roux-en-Y gastric bypass surgery or lifestyle with intensive medical management in patients with type 2 diabetes: feasibility and 1-year results of a randomized clinical trial. JAMA surgery 149(7):716-26.
Heindorff 1997	Heindorff H, Hougaard K, Larsen PN. (1997). Laparoscopic adjustable gastric band increases weight loss compared to dietary treatment: A randomized study. Obesity surgery 7(4):300-01.
Liang 2013	Liang, Z., Q. Wu, et al. (2013). Effect of laparoscopic Roux-en-Y gastric bypass surgery on type 2 diabetes mellitus with hypertension: A randomized controlled trial. Diabetes Research and Clinical Practice.
Mingrone 2002	Mingrone, G., A. V. Greco, et al. (2002). Sex hormone-binding globulin levels and cardiovascular risk factors in morbidly obese subjects before and after weight reduction induced by diet or malabsorptive surgery.Atherosclerosis 161(2): 455-462.
Parikh 2014	Parikh M, Chung M, Sheth S, et al. (2014). Randomized pilot trial of bariatric surgery versus intensive medical weight management on diabetes remission in type 2 diabetic patients who do NOT meet NIH criteria for surgery and the role of soluble RAGE as a novel biomarker of success. Ann Surg260(4):617-22; discussion 22-4.

8.6.2 Study characteristics of studies with maximal 1 year of follow-up

Appendix 7 Study characteristics of studies with maximal 1 year of follow-up

Study identification	Population	Comparison	Setting	Recruitment period Study period	n individuals assessed for eligibility and randomized	patients who withdrew* or patients who refused intervention**	Follow-up visits Maximal follow-up time reported	Adherence rate at last follow-up Adherence rate at 2 years
Ding 2015	Individuals with a BMI of 30-45 kg/m² and type 2 diabetes; prior body weight loss attempts: n.r.; run in phase : n.r.	Laparoscopic gastric banding vs. conservative treatment	Two centres, Brigham and Women's Hospital and Joslin Diabetes Center Boston, Massachusetts, USA	January 2010 to November 2013; not reported	n= 1596 assessed for eligibility (1320 phone screen failure, 119 did not pursue screening visits, 69 excluded after screening with the following reasons: surgical preference for RYGB, PMH exclusionary, time constraints, BMI out of range, insurance coverage, physician deemed inappropriate, unwilling to have surgery, A1c out of range, moved away; 43 excluded but no reason given); n= 45 randomized	LAGB: n= 4* conservative: n= 0*	LAGB: at 2, 6 weeks after surgery, then every 4-6 weeks Not reported for non-surgical treatment group, last follow-up at 12 months; 12 months	Last follow-up: LAGB: 78.3% (18/23) Conservative: 100% (22/22) At 2 years: not applicable
Halperin 2014	Individuals with a BMI of 30-42 kg/m² and type 2 diabetes; prior body weight loss attempts: n.r.; run in phase : n.r.	Gastric bypass (laparoscopic not reported) vs. conservative treatment	Two centres, Brigham and Women's Hospital and Joslin Diabetes Center Boston, Massachusetts, USA	March 2010 to September 2011; not reported	n= 822 assessed for eligibility (729 were ineligible or uninterested or did not persue sceening visit, 50 were excluded after screening with the following reasons: surgical preference for LAGB, HbA1c out of range, physician deemed inappropriate, renal function, BMI out of range, moved away, GAD positive, desired only surgery, smoker, other); n= 43 randomised	RYGB: n= 1 * conservative: n=2 *	When 10% of initial body weight was lost, or latest at 3 months. Last follow-up visit after 12 months; 12 months	Last follow-up: RYGB: 86.4% (19/22) Conservative: 90.5% (19/21) At 2 years: not applicable

Study identification	Population	Comparison	Setting	Recruitment period Study period	n individuals assessed for eligibility and randomized	patients who withdrew* or patients who refused intervention**	Follow-up visits Maximal follow-up time reported	Adherence rate at last follow-up Adherence rate at 2 years
Heindorff 1997	Individuals with a BMI of >38 kg/m^2; prior body weight loss attempts: n.r.; run in phase : n.r.	Laparoscopic gastric banding vs. conservative treatment	not reported	1995-1996; not reported	not reported	not reported	Every 4 weeks; 40 weeks	Last follow-up: not reported At 2 years: not applicable
Liang 2013	Individuals with a BMI of >28 kg/m^2, type 2 diabetes and hypertension; prior body weight loss attempts: n.r.; run in phase : n.r.	Laparoscopic gastric bypass vs. conservative treatment	Single centre: Southwest Hospital of Third Military Medical University, Chongqing, China	June 2008 to July 2011; not reported	not reported n= 108 randomized	not reported	For all study groups: at 3, 6, 9 and 12 months; 12 months	Last follow-up: 94% (for all study participants; per group not reported) At 2 years: not applicable
Mingrone 2002	morbidly obese subjects; prior body weight loss attempts: n.r.; run in phase : n.r.	Biliopancreatic diversion (laparoscopic not reported) vs. conservative treatment	Single centre: Metabolic Disease Unit, city/institution not reported Italy	not reported; June 1998 to December 2000	Eligiblity assessment not reported; n= 79 randomized	RYGB: n= 1* BPD: n= 1* conservative: n= 2*	not reported; 12 months	Last follow-up: not reported At 2 years: not applicable
Parikh 2014	Individuals with a BMI of 30-35 kg/m^2 and type 2 diabetes; failure to lose weight with nonsurgical means; run in phase : n.r.	Laparoscopic gastric bypass/ laparoscopic gastric banding/sleeve gastrectomy vs. conservative treatment	Single centre; Municipal health care system: New York City Health and Hospitals (HHC), USA	not reported; not reported	not reported	surgery: n= 8* conservative n= 3*	Surgical group: at 2 and 4 weeks, then monthly follow-up visits Conservative group: Weekly follow-up; 6 months	Last follow-up: surgery: 69.0% (20/29) conservative: 85.7% (24/28) At 2 years: not applicable

8.6.3 Results of risk of bias assessment and support for judgment of the 1-year results

Appendix 8 Results of risk of bias assessment and support for judgment of the 1-year results

Study	Random sequence generation (selection bias) and support for judgement		Allocation concealment (selection bias) and support for judgement		Blinding of participants and personnel (performance bias) and		Blinding of outcome assessment (detection bias) and support for judgement		Incomplete continuous outcome data (attrition bias) and support for judgement		Incomplete binary data (attrition bias) and support for judgement		Selective reporting (reporting bias) and support for judgement	
Ding 2015	low	"Randomization was **computer-generated** in centrally-allocated blocks of four"	low	"Randomization was computer-generated in **centrally-allocated** blocks of four"	unclear	not reported	unclear	not reported	high	missing data >20% in either study arm	high	missing data >20% in either study arm	low	results were given for all outcomes mentioned in method paragraph
Halperin 2014	low	"Randomization was **computer-generated** in centrally allocated	unclear	not reported (abstract only)	unclear	not reported	unclear	not reported	high	missing data 10% - 20 %, adressed using	low	missing data 10% - 20 %, comparable	low	results were given for all outcomes

Study	Random sequence generation (selection bias) and support for judgement		Allocation concealment (selection bias) and support for judgement		Blinding of participants and personnel (performance bias) and		Blinding of outcome assessment (detection bias) and support for judgement		Incomplete continuous outcome data (attrition bias) and support for judgement		Incomplete binary data (attrition bias) and support for judgement		Selective reporting (reporting bias) and support for judgement	
		blocks of 4, stratified by BMI above or equal to 35 and below 35."								inadequate methods		among study arms, adressed using adequate methods		mentioned in method paragraph
Heindorf	unclear	not reported (abstract only)	unclear	not reported (abstract only)	unclear	not reported (abstract only)	unclear	not reported (abstract only)	unclear	not applicable	unclear	other	unclear	abstract only
Liang 2013	low	"Patients were randomized into three groups in a 1:1:1 ratio with the use of a computeri	unclear	not reported	unclear	not reported	unclear	not reported	unclear	unclear	unclear	unclear	unclear	body weight not reported end of follow-up

Study	Random sequence generation (selection bias) and support for judgement		Allocation concealment (selection bias) and support for judgement		Blinding of participants and personnel (performance bias) and		Blinding of outcome assessment (detection bias) and support for judgement		Incomplete continuous outcome data (attrition bias) and support for judgement		Incomplete binary data (attrition bias) and support for judgement		Selective reporting (reporting bias) and support for judgement	
		zed system for generating random numbers."												
Mingrone 2002	low	computer random number generator system (upon author request)	unclear	not reported	unclear	not reported	unclear	not reported	unclear	not reported	not applicable (no relevant binary outcome identified)	not applicable	low	body weight not mentioned in methods as relevant outcome. Other outcomes were reported

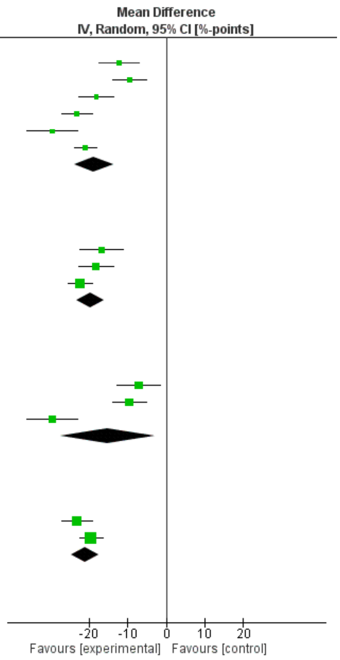
Study	Random sequence generation (selection bias) and support for judgement		Allocation concealment (selection bias) and support for judgement		Blinding of participants and personnel (performance bias) and		Blinding of outcome assessment (detection bias) and support for judgement		Incomplete continuous outcome data (attrition bias) and support for judgement		Incomplete binary data (attrition bias) and support for judgement		Selective reporting (reporting bias) and support for judgement	
Parikh 2014	unclear	method of random sequence generation not reported ("the allocation sequence was generated in advance and stratified to ensure balance of patients with BMI 30 to 35 in each of the arms.")	low	"Patients were randomized by research assistants calling a central phone number to receive allocation"	unclear	not reported	unclear	not reported	high	missing data >20% in either study arm	high	missing data >20% in either study arm	low	results were given for all outcomes mentioned in method paragraph

8.6.4 Results of RCTs with 6 months to 1 year follow-up

Appendix 9 Results of RCTs with 6 months to 1 year follow-up

8.6.4.1 **Percent body weight loss – 6 month to 1 year results**

	Bariatric surgery			Control				Mean Difference
Study or Subgroup	Mean [%-points]	SD [%-points]	Total	Mean [%-points]	SD [%-points]	Total	Weight	IV, Random, 95% CI [%-points]
5.1.1 Bariatric surgery								
Courcoulas 2015	-22.4	9.9	46	-10.2	10.6	23	16.1%	-12.20 [-17.39, -7.01]
Feigel-Guiller 2015 (1)	-15.1	7.3	26	-5.6	9.1	30	17.0%	-9.50 [-13.80, -5.20]
Ikramuddin 2015 (2)	-25.6	13.3	60	-7.5	12.2	60	16.8%	-18.10 [-22.67, -13.53]
Mingrone 2002 (3)	-29.2	12	38	-6.1	4.2	33	17.2%	-23.10 [-27.18, -19.02]
O'Brien 2006 (4)	-19.5	13.7	39	10	14.5	31	14.5%	-29.50 [-36.17, -22.83]
Schauer 2014	-26.1	7.1	97	-5.2	7.7	40	18.4%	-20.90 [-23.67, -18.13]
Subtotal (95% CI)			**306**			**217**	**100.0%**	**-18.71 [-23.75, -13.68]**
Heterogeneity: Tau² = 33.94; Chi² = 40.63, df = 5 (P < 0.00001); I² = 88%								
Test for overall effect: Z = 7.28 (P < 0.00001)								
5.1.2 Gastric bypass								
Courcoulas 2015	-27	8.9	24	-10.2	10.6	23	24.8%	-16.80 [-22.41, -11.19]
Ikramuddin 2015 (5)	-25.6	13.3	60	-7.5	12.2	60	31.5%	-18.10 [-22.67, -13.53]
Schauer 2014	-27.5	7.3	48	-5.2	7.7	40	43.7%	-22.30 [-25.46, -19.14]
Subtotal (95% CI)			**132**			**123**	**100.0%**	**-19.61 [-23.12, -16.11]**
Heterogeneity: Tau² = 4.74; Chi² = 3.93, df = 2 (P = 0.14); I² = 49%								
Test for overall effect: Z = 10.96 (P < 0.00001)								
5.1.3 Adjustable gastric banding								
Courcoulas 2015	-17.3	8.6	22	-10.2	10.6	23	33.3%	-7.10 [-12.73, -1.47]
Feigel-Guiller 2015 (6)	-15.1	7.3	26	-5.6	9.1	30	34.3%	-9.50 [-13.80, -5.20]
O'Brien 2006 (7)	-19.5	13.7	39	10	14.5	31	32.4%	-29.50 [-36.17, -22.83]
Subtotal (95% CI)			**87**			**84**	**100.0%**	**-15.18 [-27.45, -2.91]**
Heterogeneity: Tau² = 109.46; Chi² = 30.25, df = 2 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 2.42 (P = 0.02)								
5.1.4 Other surgical techniques								
Mingrone 2002 (8)	-29.2	12	38	-6.1	4.2	33	42.5%	-23.10 [-27.18, -19.02]
Schauer 2014	-24.7	6.6	49	-5.2	7.7	40	57.5%	-19.50 [-22.52, -16.48]
Subtotal (95% CI)			**87**			**73**	**100.0%**	**-21.03 [-24.52, -17.54]**
Heterogeneity: Tau² = 3.13; Chi² = 1.94, df = 1 (P = 0.16); I² = 48%								
Test for overall effect: Z = 11.82 (P < 0.00001)								



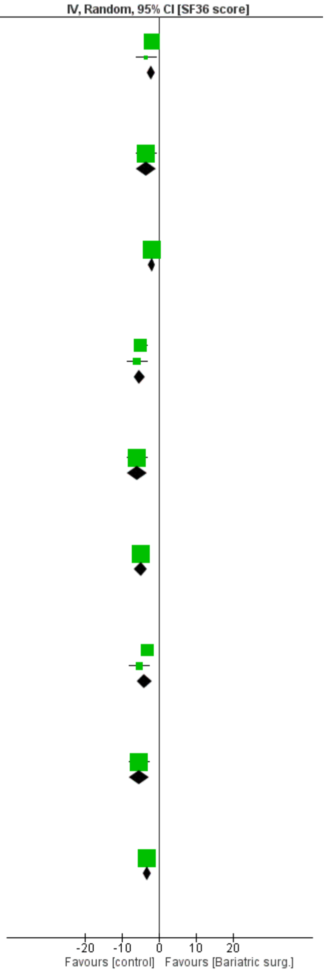
Footnotes
(1) Author inquiry, 07.09.2015
(2) Author inquiry, 08.06.2013
(3) Calculations based on data received from author inquiry 08.06.2013
(4) Extracted from Figure 3, O'Brien 2006
(5) Author inquiry, 08.06.2013
(6) Author inquiry, 07.09.2015
(7) Extracted from Figure 3, O'Brien 2006
(8) Calculations based on data received from author inquiry 08.06.2013

Percent body weight loss – 6 month to 1 year results

8.6.4.2 **Quality of life – 6 month to 1 year results**

Study or Subgroup	Bariatric surgery Mean [SF36 score]	Bariatric surgery SD [SF36 score]	Bariatric surgery Total	Control Mean [SF36 score]	Control SD [SF36 score]	Control Total	Weight	Mean Difference IV, Random, 95% CI [SF36 score]
5.3.1 Bariatric surgery, physical function								
Ding 2015 (1)	4.8	1.5	18	6.8	1.6	22	87.0%	-2.00 [-2.96, -1.04]
Halperin 2014 (2)	3.8	4	19	7.3	4.5	19	13.0%	-3.50 [-6.21, -0.79]
Subtotal (95% CI)			**37**			**41**	**100.0%**	**-2.19 [-3.18, -1.21]**
Heterogeneity: Tau² = 0.05; Chi² = 1.05, df = 1 (P = 0.31); I² = 4%								
Test for overall effect: Z = 4.36 (P < 0.0001)								
5.3.2 Gastric bypass, physical function								
Halperin 2014 (3)	3.8	4	19	7.3	4.5	19	100.0%	-3.50 [-6.21, -0.79]
Subtotal (95% CI)			**19**			**19**	**100.0%**	**-3.50 [-6.21, -0.79]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.53 (P = 0.01)								
5.3.3 Adjustable gastric banding, physical function								
Ding 2015 (4)	4.8	1.5	18	6.8	1.6	22	100.0%	-2.00 [-2.96, -1.04]
Subtotal (95% CI)			**18**			**22**	**100.0%**	**-2.00 [-2.96, -1.04]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 4.07 (P < 0.0001)								
5.3.4 Bariatric surgery, SF-36 mental health								
Ding 2015 (5)	1.9	2.9	18	6.9	3	22	67.4%	-5.00 [-6.83, -3.17]
Halperin 2014 (6)	5	4	19	11	4.3	19	32.6%	-6.00 [-8.64, -3.36]
Subtotal (95% CI)			**37**			**41**	**100.0%**	**-5.33 [-6.83, -3.82]**
Heterogeneity: Tau² = 0.00; Chi² = 0.37, df = 1 (P = 0.54); I² = 0%								
Test for overall effect: Z = 6.93 (P < 0.00001)								
5.3.5 Gastric bypass, SF-36 mental health								
Halperin 2014 (7)	5	4	19	11	4.3	19	100.0%	-6.00 [-8.64, -3.36]
Subtotal (95% CI)			**19**			**19**	**100.0%**	**-6.00 [-8.64, -3.36]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 4.45 (P < 0.00001)								
5.3.6 Adjustable gastric banding, SF-36 mental health								
Ding 2015 (8)	1.9	2.9	18	6.9	3	22	100.0%	-5.00 [-6.83, -3.17]
Subtotal (95% CI)			**18**			**22**	**100.0%**	**-5.00 [-6.83, -3.17]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 5.34 (P < 0.00001)								
5.3.7 Bariatric surgery, overall								
Ding 2015 (9)	2.6	1.8	18	5.8	1.8	22	65.4%	-3.20 [-4.32, -2.08]
Halperin 2014 (10)	2.4	4	19	7.8	4.3	19	34.6%	-5.40 [-8.04, -2.76]
Subtotal (95% CI)			**37**			**41**	**100.0%**	**-3.96 [-6.01, -1.91]**
Heterogeneity: Tau² = 1.35; Chi² = 2.26, df = 1 (P = 0.13); I² = 56%								
Test for overall effect: Z = 3.78 (P = 0.0002)								
5.3.8 Gastric bypass, overall								
Halperin 2014 (11)	2.4	4	19	7.8	4.3	19	100.0%	-5.40 [-8.04, -2.76]
Subtotal (95% CI)			**19**			**19**	**100.0%**	**-5.40 [-8.04, -2.76]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 4.01 (P < 0.0001)								
5.3.9 Adjustable gastric banding, overall								
Ding 2015 (12)	2.6	1.8	18	5.8	1.8	22	100.0%	-3.20 [-4.32, -2.08]
Subtotal (95% CI)			**18**			**22**	**100.0%**	**-3.20 [-4.32, -2.08]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 5.59 (P < 0.00001)								

Mean Difference
IV, Random, 95% CI [SF36 score]

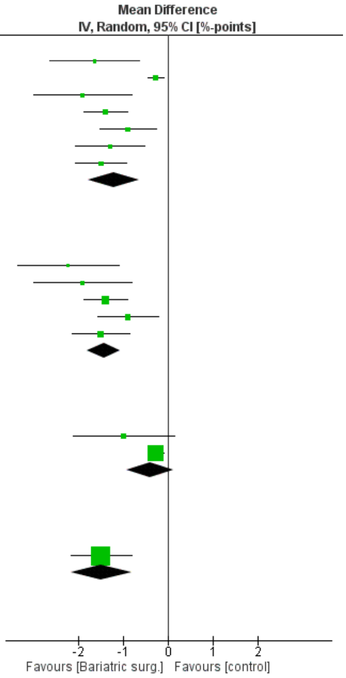


Footnotes
(1) Extracted from Ding 2015 Figure 3
(2) Extracted from Halperin 2015 Figure 3
(3) Extracted from Halperin 2015 Figure 3
(4) Extracted from Ding 2015 Figure 3
(5) Extracted from Ding 2015 Figure 3
(6) Extracted from Halperin 2015 Figure 3
(7) Extracted from Halperin 2015 Figure 3
(8) Extracted from Ding 2015 Figure 3
(9) Extracted from Ding 2015 Figure 3
(10) Extracted from Halperin 2015 Figure 3
(11) Extracted from Halperin 2015 Figure 3
(12) Extracted from Ding 2015 Figure 3

Quality of life – 6 month to 1 year results

8.6.4.3 **HbA1c – 6 month to 1 year results**

	Bariatric surgery			Control				Mean Difference
Study or Subgroup	**Mean [%-points]**	**SD [%-points]**	**Total**	**Mean [%-points]**	**SD [%-points]**	**Total**	**Weight**	**IV, Random, 95% CI [%-points]**
5.6.1 Bariatric surgery								
Courcoulas 2015	-1.7	1.99	46	-0.06	2.01	23	11.6%	-1.64 [-2.64, -0.64]
Ding 2015	-1.23	0.3	18	-0.95	0.28	22	17.7%	-0.28 [-0.46, -0.10]
Halperin 2014 (1)	-1.95	1.74	19	-0.04	1.74	19	10.8%	-1.91 [-3.02, -0.80]
Ikramuddin 2015 (2)	-3.3	1.3	60	-1.9	1.4	60	16.0%	-1.40 [-1.88, -0.92]
Liang 2013 (3)	-4.5	1.5	31	-3.6	1.5	70	14.8%	-0.90 [-1.53, -0.27]
Parikh 2014 (4)	-1.2	1.1	20	0.1	1.5	24	13.6%	-1.30 [-2.07, -0.53]
Schauer 2014	-2.9	1.69	99	-1.4	1.5	41	15.4%	-1.50 [-2.07, -0.93]
Subtotal (95% CI)			**293**			**259**	**100.0%**	**-1.21 [-1.79, -0.64]**
Heterogeneity: Tau² = 0.47; Chi² = 45.68, df = 6 (P < 0.00001); I² = 87%								
Test for overall effect: Z = 4.14 (P < 0.0001)								
5.6.2 Gastric bypass								
Courcoulas 2015	-2.3	1.94	24	-0.06	2.01	23	9.3%	-2.24 [-3.37, -1.11]
Halperin 2014 (5)	-1.95	1.74	19	-0.04	1.74	19	9.7%	-1.91 [-3.02, -0.80]
Ikramuddin 2015 (6)	-3.3	1.3	60	-1.9	1.4	60	35.1%	-1.40 [-1.88, -0.92]
Liang 2013 (7)	-4.5	1.67	31	-3.6	1.45	70	21.9%	-0.90 [-1.58, -0.22]
Schauer 2014	-2.9	1.6	50	-1.4	1.5	41	24.0%	-1.50 [-2.14, -0.86]
Subtotal (95% CI)			**184**			**213**	**100.0%**	**-1.44 [-1.81, -1.08]**
Heterogeneity: Tau² = 0.04; Chi² = 5.10, df = 4 (P = 0.28); I² = 22%								
Test for overall effect: Z = 7.77 (P < 0.00001)								
5.6.3 Adjustable gastric banding								
Courcoulas 2015	-1.05	1.87	22	-0.06	2.01	23	17.7%	-0.99 [-2.12, 0.14]
Ding 2015	-1.23	0.3	18	-0.95	0.28	22	82.3%	-0.28 [-0.46, -0.10]
Subtotal (95% CI)			**40**			**45**	**100.0%**	**-0.41 [-0.94, 0.13]**
Heterogeneity: Tau² = 0.08; Chi² = 1.47, df = 1 (P = 0.23); I² = 32%								
Test for overall effect: Z = 1.50 (P = 0.13)								
5.6.4 Other surgical techniques								
Schauer 2014	-2.9	1.8	49	-1.4	1.5	41	100.0%	-1.50 [-2.18, -0.82]
Subtotal (95% CI)			**49**			**41**	**100.0%**	**-1.50 [-2.18, -0.82]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 4.31 (P < 0.0001)								



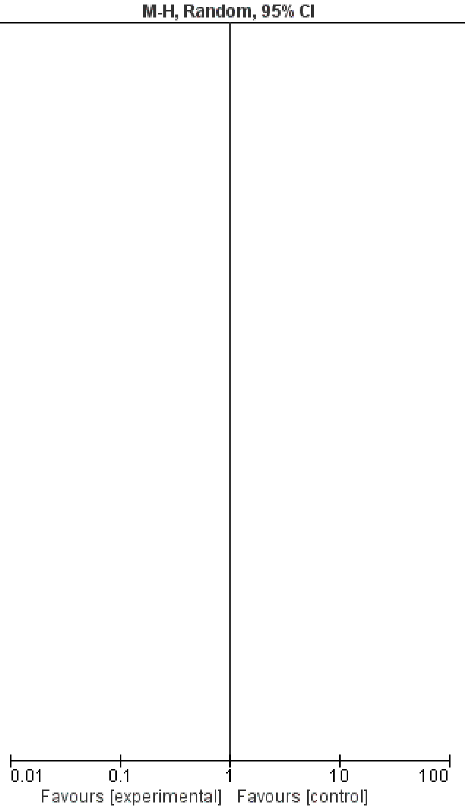
Footnotes
(1) extracted from Halperin Figure 2
(2) Author inquiry, 08.06.2013
(3) SD imputed (median from overall SD for year 1)
(4) results from 6 months follow-up, surgery group consist of mixed techniques: gastric bypass, gastric banding and sleeve gastrectomy
(5) extracted from Halperin Figure 2
(6) Author inquiry, 08.06.2013
(7) SD imputed (median from overall SD for year 1)

HbA1c – 6 month to 1 year results

8.6.4.4 **Mortality – 6 month to 1 year results**

Study or Subgroup	Bariatric surgery Events	Bariatric surgery Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
5.7.1 Bariatric surgery						
Courcoulas 2015	0	46	0	23		Not estimable
Parikh 2014	0	20	0	24		Not estimable
Schauer 2014 (1)	0	99	0	41		Not estimable
Subtotal (95% CI)		**165**		**88**		**Not estimable**
Total events	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
5.7.2 Gastric bypass						
Courcoulas 2015	0	24	0	23		Not estimable
Schauer 2014 (2)	0	50	0	41		Not estimable
Subtotal (95% CI)		**74**		**64**		**Not estimable**
Total events	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
5.7.3 Gastric banding						
Courcoulas 2015	0	22	0	23		Not estimable
Subtotal (95% CI)		**22**		**23**		**Not estimable**
Total events	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
5.7.4 Other surgical techniques						
Schauer 2014 (3)	0	49	0	41		Not estimable
Subtotal (95% CI)		**49**		**41**		**Not estimable**
Total events	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						

Risk Ratio M-H, Random, 95% CI



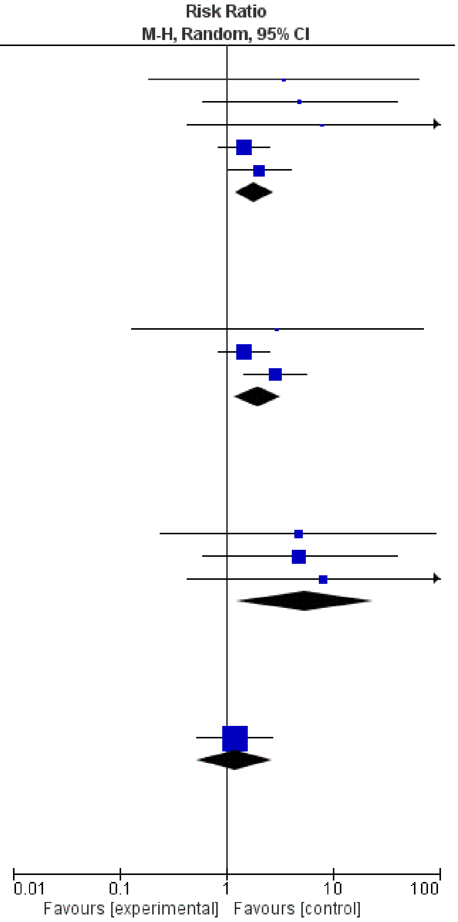
Test for subgroup differences: Not applicable

Footnotes
(1) Data from 3 years follow-up
(2) Data from 3 years follow-up
(3) Data from 3 years follow-up

Mortality – 6 month to 1 year results

8.6.4.5 SAE including reoperations – 6 month to 1 year results

Study or Subgroup	Bariatric surgery Events	Bariatric surgery Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
5.11.5 Bariatric surgery						
Courcoulas 2015 (1)	3	41	0	20	2.0%	3.50 [0.19, 64.67]
Ding 2015 (2)	4	18	1	22	3.8%	4.89 [0.60, 39.96]
Feigel-Guiller 2015 (3)	3	26	0	30	2.0%	8.04 [0.43, 148.71]
Ikramuddin 2015 (4)	22	60	15	60	55.3%	1.47 [0.85, 2.54]
Schauer 2014 (5)	38	99	8	43	37.0%	2.06 [1.05, 4.04]
Subtotal (95% CI)		**244**		**175**	**100.0%**	**1.83 [1.22, 2.76]**
Total events	70		24			
Heterogeneity: Tau² = 0.00; Chi² = 2.88, df = 4 (P = 0.58); I² = 0%						
Test for overall effect: Z = 2.90 (P = 0.004)						
5.11.6 Gastric bypass						
Courcoulas 2015 (6)	1	20	0	20	2.6%	3.00 [0.13, 69.52]
Ikramuddin 2015 (7)	22	60	15	60	55.6%	1.47 [0.85, 2.54]
Schauer 2014 (8)	27	50	8	43	41.8%	2.90 [1.48, 5.70]
Subtotal (95% CI)		**130**		**123**	**100.0%**	**1.99 [1.19, 3.31]**
Total events	50		23			
Heterogeneity: Tau² = 0.04; Chi² = 2.46, df = 2 (P = 0.29); I² = 19%						
Test for overall effect: Z = 2.64 (P = 0.008)						
5.11.7 Gastric banding						
Courcoulas 2015 (9)	2	21	0	20	24.7%	4.77 [0.24, 93.67]
Ding 2015 (10)	4	18	1	22	49.6%	4.89 [0.60, 39.96]
Feigel-Guiller 2015 (11)	3	26	0	30	25.7%	8.04 [0.43, 148.71]
Subtotal (95% CI)		**65**		**72**	**100.0%**	**5.52 [1.26, 24.25]**
Total events	9		1			
Heterogeneity: Tau² = 0.00; Chi² = 0.09, df = 2 (P = 0.96); I² = 0%						
Test for overall effect: Z = 2.26 (P = 0.02)						
5.11.8 Other surgical techniques						
Schauer 2014 (12)	11	49	8	43	100.0%	1.21 [0.53, 2.72]
Subtotal (95% CI)		**49**		**43**	**100.0%**	**1.21 [0.53, 2.72]**
Total events	11		8			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.45 (P = 0.65)						

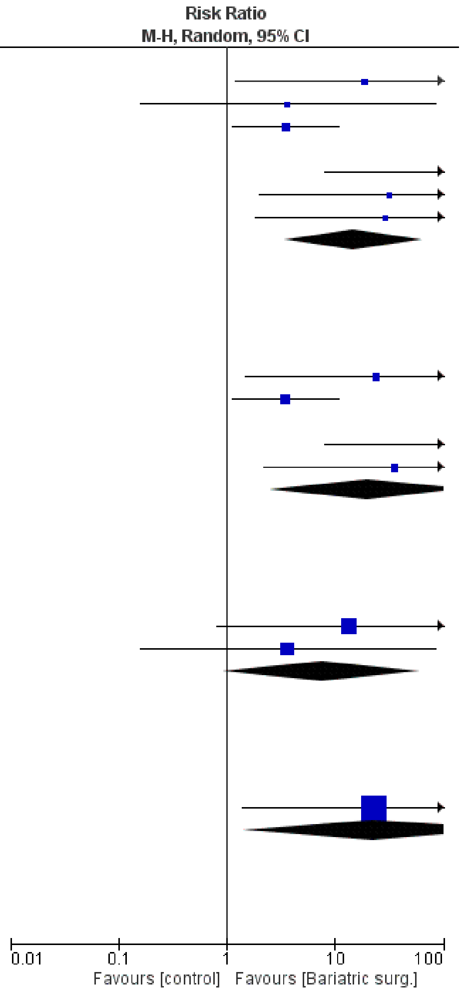


Footnotes
(1) SAE: RYGB 1 ulcer and LAGB 2 dehydration
(2) 1 reoperation (placement of LAGB failed) declared as SAE
(3) No SAE observed, 3 reoperations, Author inquiry 7.9.2015
(4) 2 reoperations (anastomotic leak) declared as SAE
(5) Reoperations declared as SAE: RYGB 3 and SG 1
(6) SAE: RYGB 1 ulcer and LAGB 2 dehydration
(7) 2 reoperations (anastomotic leak) declared as SAE
(8) Reoperations declared as SAE: RYGB 3 and SG 1
(9) SAE: RYGB 1 ulcer and LAGB 2 dehydration
(10) 1 reoperation (placement of LAGB failed) declared as SAE
(11) No SAE observed, 3 reoperations, Author inquiry 7.9.2015
(12) Reoperations declared as SAE: RYGB 3 and SG 1

SAE including reoperations – 6 month to 1 year results

8.6.4.6 Diabetes remission – 6 month to 1 year results

Study or Subgroup	Bariatric surgery Events	Bariatric surgery Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
5.8.5 Bariatric surgery						
Courcoulas 2015 (1)	18	46	0	23	15.0%	18.89 [1.19, 300.20]
Ding 2015 (2)	1	18	0	22	13.0%	3.63 [0.16, 84.11]
Halperin 2014 (3)	11	22	3	21	27.1%	3.50 [1.13, 10.81]
Ikramuddin 2015 (4)	0	60	0	60		Not estimable
Liang 2013 (5)	28	31	0	70	15.0%	126.47 [7.97, 2007.60]
Parikh 2014 (6)	13	20	0	24	15.0%	32.14 [2.03, 509.08]
Schauer 2014 (7)	34	99	0	41	15.0%	28.98 [1.82, 461.77]
Subtotal (95% CI)		**296**		**261**	**100.0%**	**14.84 [3.43, 64.15]**
Total events	105		3			
Heterogeneity: Tau² = 1.73; Chi² = 11.09, df = 5 (P = 0.05); I² = 55%						
Test for overall effect: Z = 3.61 (P = 0.0003)						
5.8.6 Gastric bypass						
Courcoulas 2015 (8)	12	24	0	23	22.3%	24.00 [1.50, 383.26]
Halperin 2014 (9)	11	22	3	21	33.0%	3.50 [1.13, 10.81]
Ikramuddin 2015 (10)	0	60	0	60		Not estimable
Liang 2013 (11)	28	31	0	70	22.4%	126.47 [7.97, 2007.60]
Schauer 2014 (12)	21	50	0	41	22.3%	35.41 [2.21, 567.36]
Subtotal (95% CI)		**187**		**215**	**100.0%**	**20.13 [2.45, 165.72]**
Total events	72		3			
Heterogeneity: Tau² = 3.18; Chi² = 10.44, df = 3 (P = 0.02); I² = 71%						
Test for overall effect: Z = 2.79 (P = 0.005)						
5.8.7 Gastric banding						
Courcoulas 2015 (13)	6	22	0	23	55.4%	13.57 [0.81, 227.36]
Ding 2015 (14)	1	18	0	22	44.6%	3.63 [0.16, 84.11]
Subtotal (95% CI)		**40**		**45**	**100.0%**	**7.54 [0.92, 61.45]**
Total events	7		0			
Heterogeneity: Tau² = 0.00; Chi² = 0.40, df = 1 (P = 0.53); I² = 0%						
Test for overall effect: Z = 1.89 (P = 0.06)						
5.8.8 Other surgical techniques						
Schauer 2014 (15)	13	49	0	41	100.0%	22.68 [1.39, 370.25]
Subtotal (95% CI)		**49**		**41**	**100.0%**	**22.68 [1.39, 370.25]**
Total events	13		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.19 (P = 0.03)						



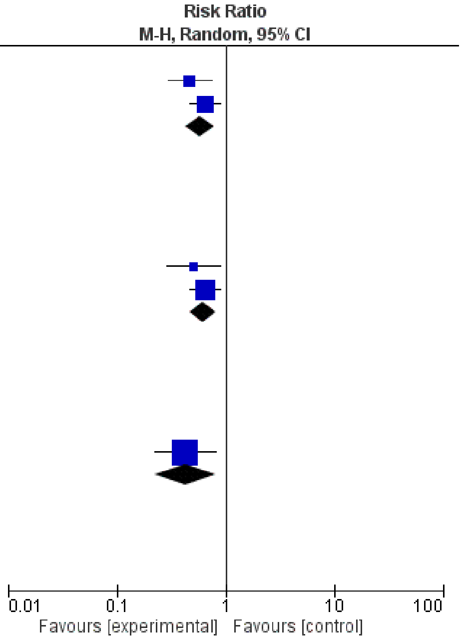
Test for subgroup differences: Chi² = 0.57, df = 3 (P = 0.90), I² = 0%

Footnotes
(1) Defined as no use of antidiabetics at 12 months after surgery with HbA1c levels of less than 6.5% and FPG levels of 125mg/dL or less.
(2) fasting plasma glucose ≤ 7mmol/L (126 mg/dL) and HbA1c ≤ 6.5% without medication
(3) fasting plasma glucose ≤ 7mmol/L (126 mg/dL) and HbA1c ≤ 6.5%, these data were interpreted as diabetes remission, because all individuals in...
(4) HbA1c ≤ 6.5% at every sampling from months 12 to 24 with no medicines for hyperglycaemia
(5) Definition of diabetes remission not reported
(6) fasting plasma glucose ≤ 7mmol/L (126 mg/dL) and HbA1c ≤ 6.5% without medication
(7) HBA1c ≤ 6.5% without diabetes medication
(8) Defined as no use of antidiabetics at 12 months after surgery with HbA1c levels of less than 6.5% and FPG levels of 125mg/dL or less.
(9) fasting plasma glucose ≤ 7mmol/L (126 mg/dL) and HbA1c ≤ 6.5%, these data were interpreted as diabetes remission, because all individuals in...
(10) HbA1c ≤ 6.5% at every sampling from months 12 to 24 with no medicines for hyperglycaemia
(11) Definition of diabetes remission not reported
(12) HBA1c ≤ 6.5% without diabetes medication
(13) Defined as no use of antidiabetics at 12 months after surgery with HbA1c levels of less than 6.5% and FPG levels of 125mg/dL or less.
(14) fasting plasma glucose ≤ 7mmol/L (126 mg/dL) and HbA1c ≤ 6.5% without medication
(15) HBA1c ≤ 6.5% without diabetes medication

Diabetes remission – 6 month to 1 year results

8.6.4.7 **Hypertension – 6 month to 1 year results**

Study or Subgroup	Bariatric surgery Events	Bariatric surgery Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
5.14.5 Bariatric surgery						
Courcoulas 2015	16	46	17	23	36.3%	0.47 [0.30, 0.75]
Ikramuddin 2015 (1)	28	60	42	59	63.7%	0.66 [0.48, 0.90]
Subtotal (95% CI)		**106**		**82**	**100.0%**	**0.58 [0.43, 0.79]**
Total events	44		59			
Heterogeneity: Tau² = 0.01; Chi² = 1.34, df = 1 (P = 0.25); I² = 25%						
Test for overall effect: Z = 3.40 (P = 0.0007)						
5.14.6 Gastric bypass						
Courcoulas 2015	9	24	17	23	23.4%	0.51 [0.29, 0.90]
Ikramuddin 2015	28	60	42	59	76.6%	0.66 [0.48, 0.90]
Subtotal (95% CI)		**84**		**82**	**100.0%**	**0.62 [0.47, 0.81]**
Total events	37		59			
Heterogeneity: Tau² = 0.00; Chi² = 0.60, df = 1 (P = 0.44); I² = 0%						
Test for overall effect: Z = 3.42 (P = 0.0006)						
5.14.7 Gastric banding						
Courcoulas 2015	7	22	17	23	100.0%	0.43 [0.22, 0.83]
Subtotal (95% CI)		**22**		**23**	**100.0%**	**0.43 [0.22, 0.83]**
Total events	7		17			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.51 (P = 0.01)						



Test for subgroup differences: Chi² = 0.98, df = 2 (P = 0.61), I² = 0%

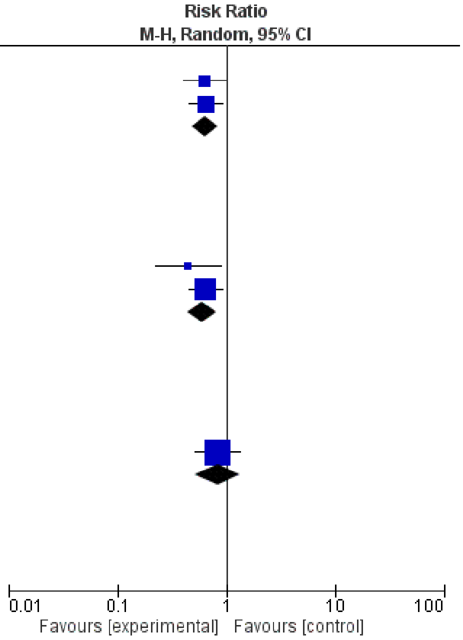
Footnotes

(1) = < 130 mmHg and no antihypertensives

Hypertension – 6 month to 1 year results

8.6.4.8 **Dyslipidemia – 6 month to 1 year results**

Study or Subgroup	Bariatric surgery Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
5.12.5 Bariatric surgery						
Courcoulas 2015	19	46	15	23	35.8%	0.63 [0.40, 1.00]
Ikramuddin 2015 (1)	26	60	40	60	64.2%	0.65 [0.46, 0.91]
Subtotal (95% CI)		**106**		**83**	**100.0%**	**0.64 [0.49, 0.85]**
Total events	45		55			
Heterogeneity: Tau² = 0.00; Chi² = 0.01, df = 1 (P = 0.93); I² = 0%						
Test for overall effect: Z = 3.16 (P = 0.002)						
5.12.6 Gastric bypass						
Courcoulas 2015	7	24	15	23	19.5%	0.45 [0.22, 0.89]
Ikramuddin 2015	26	60	40	60	80.5%	0.65 [0.46, 0.91]
Subtotal (95% CI)		**84**		**83**	**100.0%**	**0.60 [0.45, 0.82]**
Total events	33		55			
Heterogeneity: Tau² = 0.00; Chi² = 0.92, df = 1 (P = 0.34); I² = 0%						
Test for overall effect: Z = 3.23 (P = 0.001)						
5.12.7 Gastric banding						
Courcoulas 2015	12	22	15	23	100.0%	0.84 [0.52, 1.36]
Subtotal (95% CI)		**22**		**23**	**100.0%**	**0.84 [0.52, 1.36]**
Total events	12		15			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.72 (P = 0.47)						



Test for subgroup differences: Chi² = 1.26, df = 2 (P = 0.53), I² = 0%

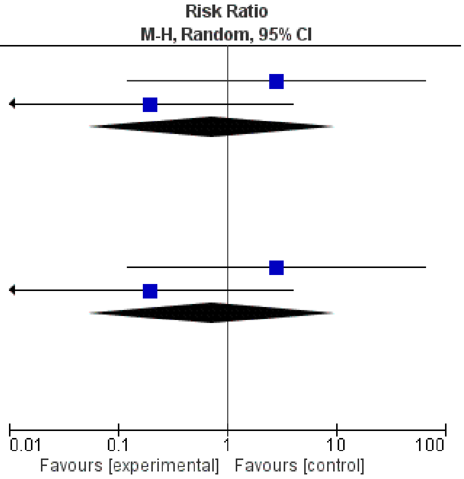
Footnotes

(1) = < 130 mmHg and no antihypertensives

Dyslipidemia – 6 month to 1 year results

8.6.4.9 **Cancer – 6 month to 1 year results**

Study or Subgroup	Bariatric surgery Events	Bariatric surgery Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
5.13.5 Bariatric surgery						
Halperin 2014	1	22	0	21	48.5%	2.87 [0.12, 66.75]
Ikramuddin 2015 (1)	0	60	2	60	51.5%	0.20 [0.01, 4.08]
Subtotal (95% CI)		**82**		**81**	**100.0%**	**0.73 [0.05, 9.96]**
Total events	1		2			
Heterogeneity: Tau² = 1.09; Chi² = 1.44, df = 1 (P = 0.23); I² = 31%						
Test for overall effect: Z = 0.24 (P = 0.81)						
5.13.6 Gastric bypass						
Halperin 2014	1	22	0	21	48.5%	2.87 [0.12, 66.75]
Ikramuddin 2015	0	60	2	60	51.5%	0.20 [0.01, 4.08]
Subtotal (95% CI)		**82**		**81**	**100.0%**	**0.73 [0.05, 9.96]**
Total events	1		2			
Heterogeneity: Tau² = 1.09; Chi² = 1.44, df = 1 (P = 0.23); I² = 31%						
Test for overall effect: Z = 0.24 (P = 0.81)						



Test for subgroup differences: Chi² = 0.00, df = 1 (P = 1.00), I² = 0%

Footnotes

(1) = < 130 mmHg and no antihypertensives

Cancer – 6 month to 1 year results

8.7 Definitions of diabetes remission

Appendix 10 Definitions of diabetes remission

Definition of diabetes remission reported by the RCTs included for the 2-years but as well for the 1-year analysis	
Study	**Definitions**
Courcoulas 2015	Partial remission = fasting plasma glucose ≤ 7 mmol/L (125 mg/dL) and HbA1c ≤ 6.5% without antidiabetic medication (according to the American Diabetes Association 2009) Complete remission = fasting plasma glucose ≤ 5.6 mmol/L (100 mg/dL) and HbA1c ≤ 5.7% without antidiabetic medication (according to the American Diabetes Association 2009)
Ding 2015	= fasting plasma glucose ≤ 7mmol/L (126 mg/dL) and HbA1c ≤ 6.5% without medication (authors reported patients still requiring medications as well)
Dixon 2008	= fasting plasma glucose ≤ 7mmol/L (126 mg/dL) and HbA1c ≤ 6.2% without medication
Ikramuddin 2015	Partial remission = HbA1c ≤ 6.5% at every sampling from months 12 to 24 with no medicines for hyperglycaemia Full remission = HbA1c ≤ 6.0% at every sampling from months 12 to 24 with no medicines for hyperglycaemia
Halperin 2014	= fasting plasma glucose ≤ 7mmol/L (126 mg/dL) and HbA1c ≤ 6.5%, all subjects in the surgery group were free of antidiabetic medication at 12 months, but the medication status for the control group was not reported
Mingrone 2012	= fasting plasma glucose ≤ 5.6mmol/L and HbA1c ≤ 6.5% for one year without medication
Parikh 2014	= according to the American Diabetes Association. Authors give definition for diabetes mellitus, which correspond to cut-off for partial remission defined by the American Diabetes Association: fasting plasma glucose ≤ 7 mmol/L (125 mg/dL) and HbA1c ≤ 6.5% without antidiabetic medication (author request)
Schauer 2014	= HBA1c ≤ 6.5% without diabetes medication (only reported for 3 years of follow-up)

	= HBA1c ≤ 6.0% without diabetes medication (reported for 1 and 3 years of follow-up) = HBA1c ≤ 6.0% (only reported for 2 years of follow-up)
Liang 2013	Not reported
Wentworth 2014	= fasting plasma glucose < 7mmol/L (54mmol/mol) and < 11.1mmol/L 2h after oral glucose at least 2 days after stopping glucose lowering drugs

8.8 Adverse events reported by each study

Appendix 11 Adverse events reported by each study

Study	Courcoulas 2015 RYGB	Courcoulas 2015 LAGB	Courcoulas 2015 control	Ding 2015 LAGB	Ding 2015 control	Halperin 2014 RYGB	Halperin 2014 control	Mingrone 2012 BPD	Mingrone 2012 RYGB	Mingrone 2012 control	Schauer 2014 RYGB	Schauer 2014 SG	Schauer 2014 control	Ikramuddin 2013 RYGB	Ikramuddin 2013 control	Liang 2013 RYGB	Liang 2013 control 1	Liang 2013 control 2	Liang 2013 Control 1+2
Patients (n) which attended allocated treatment	20	21	20	18	22	19	19	19	19	20				60	59	31	36	34	70
Acute abdomen																			
Amputation below the knee														1	0				
Anastomotic ulcer											4	0	1	3	1				
Anoxic brain injury														1	0				
Angina																			
Arrhythmia or palpitations																			
Band removed		0																	
Bleeding											2	0	0	2					
Blood transfusion needed																			
Bowel obstruction											1	1	1	2					
Bronchitis														0	1				
Buttock abscess																			
Cardiac and renal failure																			
Cataract														1	0				
Cellulitis																			
Cholelithiasis														1	0				

Study	Courcoulas 2015 RYGB	Courcoulas 2015 LAGB	Courcoulas 2015 control	Ding 2015 LAGB	Ding 2015 control	Halperin 2014 RYGB	Halperin 2014 control	Mingrone 2012 BPD	Mingrone 2012 RYGB	Mingrone 2012 control	Schauer 2014 RYGB	Schauer 2014 SG	Schauer 2014 control	Ikramuddin 2013 RYGB	Ikramuddin 2013 control	Liang 2013 RYGB	Liang 2013 control 1	Liang 2013 control 2	Liang 2013 Control 1+2
Cholecystectomy																			
Cold, flu and sinusitis/virus, sinusitis and gastro																			
Congestive heart failure														0	1				
Deep venous thrombosis														1	0				
Dehydration	0	2	0																
Depression														0	1				
Diarrhoea														1	0				
Dumping											4	1	0						
Duodenal ulcer																			
Duodenitis														0	1				
Elective arthroplasties																			
Eosinophilic fasciitis																			
Excessive weight gain											0	0	7						
Epilepsy																			
Fall with fracture														5	1				
Fall with other injury														2	2				
Febrile episodes																			
Food intolerance																			
Foot ulcer											2	1	0						
Gallstone diseases																			

Study	Courcoulas 2015 RYGB	Courcoulas 2015 LAGB	Courcoulas 2015 control	Ding 2015 LAGB	Ding 2015 control	Halperin 2014 RYGB	Halperin 2014 control	Mingrone 2012 BPD	Mingrone 2012 RYGB	Mingrone 2012 control	Schauer 2014 RYGB	Schauer 2014 SG	Schauer 2014 control	Ikramuddin 2013 RYGB	Ikramuddin 2013 control	Liang 2013 RYGB	Liang 2013 control 1	Liang 2013 control 2	Liang 2013 Control 1+2
Gastric pouch enlargement																			
Gastritis														1					
Gastrointestinal leak											0	1	0	2					
Headache														1	0				
Haemoglobin decrease > 5g/dL																			
Hernia								1			3	1	1						
Herniated spinal disc with foot drop														1	0				
Hyper-triglyceridaemia														1	5				
Hypertension with hospital admission														1	0				
Hypoalbuminemia								2						4	0				
Hypoglycaemic episode				0*	0*	0*	0*				33	40	39	5	4				
Hypokalaemia																			
Hypotension and light-headedness	0	1	0																
Inguinal hernia repair																			
Intestinal occlusion									1										
Intolerance to metformin																			
intolerance to orlistat																			
Intolerance to very-low-calorie diet																			
Intravenous treatment for											7	4	3	1	0				

Study	Courcoulas 2015 RYGB	Courcoulas 2015 LAGB	Courcoulas 2015 control	Ding 2015 LAGB	Ding 2015 control	Halperin 2014 RYGB	Halperin 2014 control	Mingrone 2012 BPD	Mingrone 2012 RYGB	Mingrone 2012 control	Schauer 2014 RYGB	Schauer 2014 SG	Schauer 2014 control	Ikramuddin 2013 RYGB	Ikramuddin 2013 control	Liang 2013 RYGB	Liang 2013 control 1	Liang 2013 control 2	Liang 2013 Control 1+2
dehydration/ dehydration																			
Iron deficiency anaemia								2	2		8	15	6	13	0				
Ketoacidosis														0	1				
Kidney stone	1	0	0								5	4	6	1	2				
Knee arthroscopy																			
Knee reconstruction																			
Loss of body strength														1	0				
Malabsorption/maldigestion																			
Multiple sclerosis														1	0				
Nausea and emesis requiring IV hydration	1	0	0																
Neuropathy											7	5	4	2	2				
Osteopenia (BMD T score, -2)								1											
Osteoporosis (BMD T score, -2.7)								1											
Pain	0	1	0											7	3				
Pancreatitis														0	3				
Pancreatic cancer														0	2				
Partial third cranial nerve palsy														0	1				
Perforation																			
Pleural effusion																			

Study	Courcoulas 2015 RYGB	Courcoulas 2015 LAGB	Courcoulas 2015 control	Ding 2015 LAGB	Ding 2015 control	Halperin 2014 RYGB	Halperin 2014 control	Mingrone 2012 BPD	Mingrone 2012 RYGB	Mingrone 2012 control	Schauer 2014 RYGB	Schauer 2014 SG	Schauer 2014 control	Ikramuddin 2013 RYGB	Ikramuddin 2013 control	Liang 2013 RYGB	Liang 2013 control 1	Liang 2013 control 2	Liang 2013 Control 1+2
Pneumonia											2	1	0						
Pregnancy																			
Prolonged hospitalization (+1 day)	2	4	0																
Pruritus/erythema at incision site	0	1	0																
Pyelonephritis														0	1				
Reflux oesophagitis														3	2				
Reoperation	0	1	0					1	1										
Retinal photocoagulation																			
Retinopathy											1	2	0						
Rupture of suture or staple line																			
Spinal stimulator placement for pain																			
Suicide attempt														0	1				
Surgery, hand																			
Surgery, heel																			
Surgery, nasal																			
Stricture (gastrointestinal)											1	1	0	2					
Stroke											0	1	0						
Transient renal insufficiency																			
Transurethral resection of prostate																			

Study	Courcoulas 2015 RYGB	Courcoulas 2015 LAGB	Courcoulas 2015 control	Ding 2015 LAGB	Ding 2015 control	Halperin 2014 RYGB	Halperin 2014 control	Mingrone 2012 BPD	Mingrone 2012 RYGB	Mingrone 2012 control	Schauer 2014 RYGB	Schauer 2014 SG	Schauer 2014 control	Ikramuddin 2013 RYGB	Ikramuddin 2013 control	Liang 2013 RYGB	Liang 2013 control 1	Liang 2013 control 2	Liang 2013 Control 1+2
Trocar site abscess																			
Toe amputation														1	0				
Unwanted pregnancy														1	0				
Urinary tract infection														1	0				
Uretal stone														1	1				
Uterine curettage																			
Vasculitic rash																			
Vitamin-deficiencies (B and D)														15	7				
Vomiting/nausea																5	3	13	16
Wound /5 mm port site infection											1	0	0	1		6			
Wound hematoma														1					

9 Appendix B - Health economic analyses

9.1 Search strategy

Appendix 12 Search strategy

A- Initial search strings for the health economic systematic review (dated, August 6th, 2015)

Medline library

Clinical search string

1. obesity.ti,ab.
2. exp obesity/
3. obese.ti,ab.
4. exp Overweight/
5. overweight/
6. over weight.ti,ab.
7. overweight.ti,ab.
8. (overeating or over eating).ti,ab.
9. or 1-8

Intervention search string

10. exp Gastric Bypass/
11. (gastric adj5 bypass).ti,ab.
12. (gastric adj5 surgery).ti,ab.
13. exp Gastroplasty/
14. (gastroplasty or gastro?gastrostom*).ti,ab.
15. (gastrointestinal and bypass).ti,ab.
16. exp Biliopancreatic Diversion/
17. ((biliopancreatic or bilio?pancreatic or bilio pancreatic) and (diversion or surg$ or bypass)).ti,ab.
18. gastric band$.ti,ab.
19. silicon band$.ti,ab.
20. LAGB.ti,ab.
21. exp Gastrectomy/
22. gastrectom*.ti,ab.
23. (lap$ and band$).ti,ab.
24. Roux en Y.ti,ab.
25. RYGB.ti,ab.
26. Anastomosis, Roux-en-Y/
27. duodenal switch.ti,ab.

28. ((gastric or silicon) and sleeve).ti,ab.
29. exp Gastroenterostomy/
30. gastroenterostom*.ti,ab.
31. mason$ procedure.ti,ab.
32. or 10-31

Economics search string

33. costs
34. cost
35. costed
36. costing
37. economic*
38. price*
39. or 33-38

---

40. 9 AND 32 AND 39
41. animal/ not humans/
42. 40 not 41
43. limit 42 to yr="2005 -Current"

44. remove duplicates from 43 **Hits = (300)**

Embase library

Clinical search string

1. exp obesity/
2. obese.ti,ab.
3. obesity.ti,ab.
4. overweight.ti,ab.
5. over weight.ti,ab.
6. (overeating or over eating).ti,ab.
7. or 1-6

Intervention search string

8. exp stomach bypass/
9. (stomach adj5 bypass).ti,ab.
10. (gastric adj5 bypass).ti,ab.
11. exp gastroplasty/
12. (gastroplasty or gastro?gastrostom*).ti,ab.
13. (gastrointestinal and bypass).ti,ab.
14. exp biliopancreatic bypass/
15. ((biliopancreatic or bilio?pancreatic or bilio pancreatic) and bypass).ti,ab.
16. ((biliopancreatic or bilio?pancreatic or bilio pancreatic) and diversion).ti,ab.
17. ((biliopancreatic or bilio?pancreatic or bilio pancreatic) and surg$).ti,ab.

18. exp gastric banding/
19. gastric band$.ti,ab.
20. ((gastric or silicon) and band$).ti,ab.
21. silicon band$.ti,ab.
22. LAGB.ti,ab.
23. exp gastrectomy/
24. gastrectom*.ti,ab.
25. (lap$ and band$).ti,ab.
26. Roux en Y.ti,ab.
27. RYGB.ti,ab.
28. exp Roux Y anastomosis/
29. duodenal switch.ti,ab.
30. exp gastroenterostomy/
31. gastroenterostom*.ti,ab.
32. mason$ procedure.ti,ab.
33. or 8-32

Economics search string

34. costs
35. cost
36. costed
37. costing
38. economic*
39. price*
40. or 34-39

---

41. 7 AND 33 AND 40
42. animal/ not humans/
43. 41 not 42
44. limit 43 to yr="2005 -Current"

45. remove duplicates from 44 **Hits = (715)**

Cochrane library

Clinical search string

1. (over weight or overweight or overeating or over eating):ti,ab,kw
2. obesity:ti,ab,kw
3. MeSH descriptor: [Obesity] explode all trees
4. obese:ti,ab,kw
5. MeSH descriptor: [Overweight] explode all trees
6. or #1-#5

Intervention search string

7. MeSH descriptor: [Gastric Bypass] explode all trees
8. (gastric near/5 bypass):ti,ab,kw
9. MeSH descriptor: [Gastroplasty] explode all trees

10. (gastroplasty or gastrogastrostom*):ti,ab,kw
11. (digestive and bypass):ti,ab,kw
12. (gastrointestinal and bypass):ti,ab,kw
13. MeSH descriptor: [Biliopancreatic Diversion] explode all trees
14. (biliopancreatic and (diversion or surg* or bypass)):ti,ab,kw
15. (bilio pancreatic and (diversion or surg* or bypass)):ti,ab,kw
16. (gastric and (sleeve* or band*)):ti,ab,kw
17. (silicon and (sleeve* or band*)):ti,ab,kw
18. LAGB:ti,ab,kw
19. MeSH descriptor: [Gastrectomy] explode all trees
20. gastrectom*:ti,ab,kw
21. gastroenterostom*:ti,ab,kw
22. MeSH descriptor: [Gastroenterostomy] explode all trees
23. (lap* and band*):ti,ab,kw
24. Roux en Y:ti,ab,kw
25. RYGB:ti,ab,kw
26. MeSH descriptor: [Anastomosis, Roux-en-Y] explode all trees
27. duodenal switch:ti,ab,kw
28. mason* procedure:ti,ab,kw
29. or #7-#28

Economics search string

30. costs
31. cost
32. costed
33. costing
34. economic*
35. price*
36. or #30-#36

---

37. #6 and #29 and #36

38. Publication Year from 2000 to 2015, in Technology Assessments and Economic Evaluation **(Hits = 76)**

Center for review and dessimination

Clinical search string

1. obesity
2. obese
3. overweight
4. overeating
5. or #1-#4

Intervention search string

6. gastric Bypass
7. Gastroplasty
8. gastrostom
9. gastrointestinal and bypass
10. Biliopancreatic Diversion
11. Biliopancreatic surgery
12. Biliopancreatic bypass
13. gastric band
14. silicon band
15. LAGB
16. gastrectomy
17. lap and band
18. Roux en Y
19. RYGB
20. duodenal switch
21. gastric or silicon and sleeve
22. gastroenterostomy
23. gastroenterostom
24. mason procedure
25. or #6-#26

Economics search string

26. costs
27. cost
28. costed
29. costing
30. economic*
31. price*
32. or #28-#34

B- 5 AND 27 AND 34 **(Hits = 92)**

9.2 CHEERS checklist

Appendix 13 B- CHEERS checklist

Section/item	Item No	Recommendation
Title and abstract		
Title	1	Identify the study as an economic evaluation or use more specific terms such as “cost-effectiveness analysis”, and describe the interventions compared.
Abstract	2	Provide a structured summary of objectives, perspective,

Section/item	Item No	Recommendation
		setting, methods (including study design and inputs), results (including base case and uncertainty analyses), and conclusions.
Introduction		
Background and objectives	3	Provide an explicit statement of the broader context for the study. Present the study question and its relevance for health policy or practice decisions.
Methods		
Target population and subgroups	4	Describe characteristics of the base case population and subgroups analysed, including why they were chosen.
Setting and location	5	State relevant aspects of the system(s) in which the decision(s) need(s) to be made.
Study perspective	6	Describe the perspective of the study and relate this to the costs being evaluated.
Comparators	7	Describe the interventions or strategies being compared and state why they were chosen.
Time horizon	8	State the time horizon(s) over which costs and consequences are being evaluated and say why appropriate.
Discount rate	9	Report the choice of discount rate(s) used for costs and outcomes and say why appropriate.
Choice of health outcomes	10	Describe what outcomes were used as the measure(s) of benefit in the evaluation and their relevance for the type of analysis performed.
Measurement of effectiveness	11a	*Single study-based estimates:* Describe fully the design features of the single effectiveness study and why the single study was a sufficient source of clinical effectiveness data.
	11b	*Synthesis-based estimates:* Describe fully the methods used for identification of included studies and synthesis of clinical effectiveness data.
Measurement and valuation of preference based outcomes	12	If applicable, describe the population and methods used to elicit preferences for outcomes.
Estimating resources and costs	13a	*Single study-based economic evaluation:* Describe approaches used to estimate resource use associated with the alternative interventions. Describe primary or secondary research methods for valuing each resource item in terms of its unit cost. Describe any adjustments made to approximate to opportunity costs.
	13b	*Model-based economic evaluation:* Describe approaches and data sources used to estimate resource use associated with

Section/item	Item No	Recommendation
		model health states. Describe primary or secondary research methods for valuing each resource item in terms of its unit cost. Describe any adjustments made to approximate to opportunity costs.
Currency, price date, and conversion	14	Report the dates of the estimated resource quantities and unit costs. Describe methods for adjusting estimated unit costs to the year of reported costs if necessary. Describe methods for converting costs into a common currency base and the exchange rate.
Choice of model	15	Describe and give reasons for the specific type of decision-analytical model used. Providing a figure to show model structure is strongly recommended.
Assumptions	16	Describe all structural or other assumptions underpinning the decision-analytical model.
Analytical methods	17	Describe all analytical methods supporting the evaluation. This could include methods for dealing with skewed, missing, or censored data; extrapolation methods; methods for pooling data; approaches to validate or make adjustments (such as half cycle corrections) to a model; and methods for handling population heterogeneity and uncertainty.
Results		
Study parameters	18	Report the values, ranges, references, and, if used, probability distributions for all parameters. Report reasons or sources for distributions used to represent uncertainty where appropriate. Providing a table to show the input values is strongly recommended.
Incremental costs and outcomes	19	For each intervention, report mean values for the main categories of estimated costs and outcomes of interest, as well as mean differences between the comparator groups. If applicable, report incremental cost-effectiveness ratios.
Characterising uncertainty	20a	*Single study-based economic evaluation:* Describe the effects of sampling uncertainty for the estimated incremental cost and incremental effectiveness parameters, together with the impact of methodological assumptions (such as discount rate, study perspective).
	20b	*Model-based economic evaluation:* Describe the effects on the results of uncertainty for all input parameters, and uncertainty related to the structure of the model and assumptions.

Section/item	Item No	Recommendation
Characterising heterogeneity	21	If applicable, report differences in costs, outcomes, or cost-effectiveness that can be explained by variations between subgroups of patients with different baseline characteristics or other observed variability in effects that are not reducible by more information.
Discussion		
Study findings, limitations, generalisability, and current knowledge	22	Summarise key study findings and describe how they support the conclusions reached. Discuss limitations and the generalisability of the findings and how the findings fit with current knowledge.
Other		
Source of funding	23	Describe how the study was funded and the role of the funder in the identification, design, conduct, and reporting of the analysis. Describe other non-monetary sources of support.
Conflicts of interest	24	Describe any potential for conflict of interest of study contributors in accordance with journal policy. In the absence of a journal policy, we recommend authors comply with International Committee of Medical Journal Editors recommendations.

9.3 Summary of studies

Appendix 14 Summary of studies, information about effectiveness and costs of individual studies

In this section you will find 22 studies that were assessed with the CHEERS checklist, including an overall description and detailed information about how was the effectiveness and cost information derived.

Ackroyd 2006

- Ackroyd et al. utilised a decision model which examined the cost-effectiveness and budget impact of surgery (adjustable gastric banding and gastric bypass) compared to conventional treatment in the United Kingdom, Germany, and France from a payer-perspective (Ackroyd 2006). The conventional treatment was described as watchful waiting or medically guided dieting. Clinical evidence was obtained from several publication quotes and included health technology assessment reports issued by NICE, ANAES, ASERNIP-S, SBU, and the DGA. Quality of life status was measured within a database of 13,547 individuals (HODaR Cardiff Research Consortium) by use of the EuroQol 3-level 5-dimensional health outcome questionnaire. Recourse unit costs were taken from national tariffs, registries, publications and interviews. All patients had a BMI of > 35 kg/m^2 and type 2 diabetes. Costs and effects were discounted by 3.5% with a time horizon of 5 years. Use of gastric bypass and adjustable gastric banding resulted in a QALY gain of 1.34 and 1.03 respectively. Both options were dominant (cost-saving) in Germany and France, in short-term (meaning over one year) and long-term over 5 years time horizon. For the UK results were cost-effective, over 5 years, and GBP vs. conventional treatment showed an ICER of £ 1,521 per QALY gained. The budget impact was calculated with a cohort of 1,000 patients over a time horizon of 5 years. This resulted in a net saving of € 5.03 million and € 3.59 million for Germany, in the case of a gastric bypass and adjustable gastric bypass, respectively. In France, a net saving of € 5.88 million in case of the gastric bypass, and € 4.48 million in the case of the adjustable gastric bypass. In the UK, there was a cost increase of £ 2.03 million and £ 1.98 million for patients treated with gastric bypass and adjustable gastric bypass, respectively. Based on the model, assumptions and structure gastric bypass and adjustable gastric banding is a dominant strategy in the treatment of patients with a BMI of > 35 kg/m^2 and type 2 diabetes in Germany and France and cost-effective in the UK.

- In the study of Ackroyd et al., clinical evidence was obtained from publications quoted in HTA reports issued by the National Institute for Clinical Excellence (NICE) (Clegg 2002), the "Agence Nationale d'Accréditation et d'Evaluation en Santé" (ANAES), the Australian Safety and Efficacy Register of New Interventional Procedures (ASERNIP-S), the Swedish Council on Technology Assessment in Health Care (SBU) and the "Deutsche Adipositas Gesellschaft" (DGA) (Ackroyd 2006). For conventional treatment, a BMI reduction of 3 kg/m^2 and a type 2 diabetes prevalence of 80% were assumed for the first follow-up year with continued intensive medical treatment. For years 2-5 with medical monitoring but without intervention, no changes in BMI and type 2 diabetes prevalence were assumed (i.e. type 2 diabetes prevalence went back to 100% and original weight was assumed). For the GBP group, a BMI reduction ranging from 16.1 to 17.7 kg/m^2 during the 5-years follow-up was assumed. The type 2 diabetes prevalence after five years was 50%. Mortality was not included in the model. Patients' self-rating of their current health-related quality of life according to their BMI and their diabetes status was estimated from a database of 13 547 individuals provided by the Health Outcomes Data Repository (HODaR) Cardiff Research Consortium. The HODaR database societal-perspective utility scores were calculated for each patient using their answer to the EuroQol 3-level 5-dimensional health outcome (EQ-5D) questionnaire. The same utility scores were applied for the Germany, France, and the UK. Five years after GBP, the cumulative utility per patient were 3.34 QALYs. For diabetes patients in conventional treatment, the cumulative utility after five years was 2.00 QALYs.

- A comprehensive list of health-care resources necessary for preoperative assessment, laparoscopic GBP and AGB surgical operations, follow-up, and the treatment of complications up to 5 years after surgery was established for three countries (Germany, the UK, and France) (Ackroyd 2006). Resource items for surgery included the amounts of human resources (surgeons, physicians, nurses, and nutritionists), imaging and laboratory tests, operating-room overhead, post-surgical recovery room, hospital stay, consultations, complications, implants, and other not specified factors. Ranges of unit costs for each resource item were collected using applicable national tariffs, registries, publications and interviews when no other source was available. A similar cost analysis process was conducted for the conventional treatment of obesity after failure of at least one prior year of well-conducted medical treatment. The following resource items were considered: consultations (with physician, district nurse, practice nurse, dietician, nutritionist, and specialist), laboratory tests, food substitutes, and admission in institutions. Since no standard treatment was identified in the three countries, the resource use for conventional treatment was based on HTA reports or authors' opinion. The annual cost of treating diabetes in the three countries was obtained from the CODE-2 survey published results. The total direct costs per patient in the first five years after GBP were estimated at € 12,166 in Germany, € 13,399 in France, and £ 9,121 in the UK. The costs for conventional treatment during the same period were € 17,197, € 19,276, and £ 7,083 for Germany, France and the UK, respectively.

Anathapavan 2010

- Anathapavan et al. assessed the cost-effectiveness of laparoscopic adjustable gastric banding versus current practice (combination of diet, exercise and behaviour modification) in Australia, from a societal perspective (Ananthapavan 2010). Patients with a BMI > 35 kg/m^2 and aged 14–19 years were included. 28 adolescents were included and used to calculate the effectiveness and costs. The calculation were then extrapolated to the Australian adolescent population. Costs and disability-adjusted life year (DALY) were based on the BMI at 3 years after surgery. Benefits were tracked until the patient had reached age of 100 years. Costs and effects were discounted by 3%. Simulation-modelling techniques were used to assess the 95% confidence interval (CI) around the cost-effectiveness ratio. Costs of a patients treated with laparoscopic adjustable gastric banding were AUD 31,000. Costs per BMI unit saved were AUD 2,200 (CI 95% 1,300 – 7,300) and AUD 10,400 (CI 95% 9,000 – 12,200) per DALY saved. Savings from the reduction in obesity-related diseases were AUD 18,200 on average per patient. The authors concluded that whilst the cost-effectiveness analyses is unlikely to be accepted by all stakeholders, nevertheless, adjustable gastric banding is an important intervention to treat severely obese adolescents. There was no information on comparator and intervention effects. Therefore there was no clear ICER.

- In the Australian study of Ananthapavan et al., effectiveness data was obtained by an audit of in-patient and out-patient records of 28 adolescents (2 boys and 26 girls, mean age 18 years, range 14–19), who underwent LAGB at the Centre for Bariatric Surgery, Avenue Hospital (Windsor, VIC, Australia) (Ananthapavan 2010). The mean change in BMI at 3 years was 13.9 kg/m^2, and it was assumed that it was maintained at 4 years and into the future. Mortality was included in the model. The utility was quantified using DALYs model developed by Haby et al. and based on BMI reduction, morbidity, and mortality after surgery (Haby et al. 2006). For a lifetime horizon, LAGB resulted in a DALYs gain range of 2.54 to 3.46. Unfortunately, overall DALYs for intervention and conventional treatment were not reported.

- In the Australian study of Ananthapavan et al., the assessment of costs included physician and surgical consultations, laboratory tests, hospitalisation and parent time and travel (Ananthapavan 2010). Information concerning unit costs were extrapolated from the Live, Eat and Play (LEAP) pilot study, from the Medicare Benefits Schedule, from the Department of Health and Aging, from the Australian Bureau of Statistics and from the Royal Automobile Club of Victoria. The costs for any bariatric procedure in a private hospital (Australian Refined Diagnostic Related Group (AN-DRG) K04Z “major procedure for obesity") were used as the starting point to cost the in-patient services. Costs of comorbidities, informal care and productivity were excluded. The average lifetime cost per adolescent treated was estimated to be around AUD 31'000 (21% of which were out-of pocket expenses to the patients or families). For conventional treatment, resource use and unit costs were not reported.

Anselmino 2009

- The article by Anselmino et al. forms a pair with a previous publication by Ackroyd et al. (Anselmino 2009, Ackroyd 2006). The aim of the study was to establish, using a payer-perspective, the cost-effectiveness and budget impact of adjustable gastric banding (AGB) and gastric bypass (GBP) vs. conventional treatment in patients with a BMI ≥ 35 kg/m^2 and type 2 diabetes mellitus (type 2 diabetes) in Austria, Italy, and Spain. However, it is very difficult to understand how and from where the QALYs were derived, as there was no information given regarding effectiveness. The authors describe results as such: in Spain, adjustable gastric banding and gastric banding yielded a moderate cost increase but were cost-effective, assuming a willingness-to-pay threshold of € 30,000 per QALY gained. Under the worst-case analysis, AGB and GBP remained a cost-saving or around breakeven in Austria and Italy and remained cost-effective in Spain. However, there was no information about costs and effects, and as such this study was not included in the further evaluation and cost adaptation for Switzerland.

- In the paper of Anselmino et al., clinical evidence was obtained from the literature. In particular, the study used the same model and sources of Ackroyd 2006 (Anselmino 2009, Ackroyd 2006). The effects of surgery were assessed using QALYs. Unfortunately, QALY information (including the sources) for GBP and conventional treatment were not provided in the article.

- In the paper of Anselmino et al., strictly related to the article of Ackroyd et al., only healthcare costs in Italy, Austria, and Spain were included (Anselmino 2009, Ackroyd 2006). Inputs for AGB and GBP hospitalisation came from the Lombardy diagnostic-related group (DRG) tariff in Italy, the average service-based hospital tariff (Leistungsorientierte Krankenanstaltenfinanzierung, LKF) in Austria, and microcosting estimates from two hospitals in Madrid (Hospital Clinico "San Carlos" and Fundación Hospital Alcorcón) in Spain. To these costs were added to the pre-operative assessment, the complication, and the 5 years follow-up costs. The conventional treatment was defined as the continuation of medically guided diet during one year in spite of previous failure, followed by four years of watchful waiting. Items for modelling included consultations (physician, nurse, and dietician), laboratory tests, food substitutes, and medication for 5 follow-up years. Depending on the country, difference sources were used (from national to international literature, including in particular the HTA reports published by the National Institute for Clinical Excellence). Costs of diabetes were obtained from the CODE-2 survey published results. Overall, the costs of GBP over 5 years were € 6,361 in Austria, € 7,831 in Italy, and € 8,344 in Spain. During the same period, the costs for conservative treatment were € 135, € 468, and € 831 for Austria, Italy, and Spain, respectively.

Borg 2014

- Borg et al. utilised a Markov micro-simulation model that simulates the outcomes of individual patients in terms of costs, life years, and QALYs over a 50 year time period to evaluate the cost-effectiveness of gastric bypass surgical treatment compared to conventional treatment for obesity in adult Swedish patients, from a societal perspective. Conventional treatment consisted of the prevalent mixture of non-surgical obesity treatments. Data from the Swedish Obesity Study (SOS), a longitudinal intervention trial that evaluated surgical treatment of obesity was used to derive the effectiveness of surgery. Information from the SOS study was combined with other information to measure relative risks of death by age group and BMI according to the NHANES studies (Flegal 2005), and age and gender specific mortality of the Swedish general population in 2012. The model recorded life years and then were weighted into QALY estimates. To estimate QALY, weights from the HODaR study, depending on the patient's BMI value were derived from Ackroyd et al. (Ackroyd 2006). In this study, three types of direct costs were included: cost of surgical procedure (including any plastic surgery required), cost of adverse events, and excess costs of treating obesity related diseases. In addition, this study also incorporated indirect costs (productivity loss due to sick leave). Costs were reported for the year 2012 in Swedish Krona (SEK). In the base-case analyses a stratum age of 45-54 years of age was chosen for assessment, while the scenario analyses assessed the outcomes for different BMIs, age and gender. Results showed that in patients with a BMI of 40-44 kg/m^2, surgery was estimated to be a cost-saving in men and judged cost-effective in women, with an incremental cost per QALY gained of SEK 26,000 (€ 3,000). The incremental costs associated with gastric bypass decreased with BMI, increased with the patient's age, and were higher in women than in men. In the group of patients with BMI 30-34 kg/m^2 the ICER was SEK 75,625 per QALY gained for men, and SEK 62,617 per QALY gained for females. Based on the scenario analyses and assumptions made around the input parameters and design of the model, the study by Borg et al. showed that the ICER was estimated to be at most SEK 160,000 per QALY gained (approximately € 18,000 per QALY). There were a number of scenario analyses performed but there was no information about uncertainty ranges, the distribution of input variables and probabilistic sensitivity analyses.

- In the study by Borg et al., surgery effectiveness was primarily based on the Swedish Obesity Study (SOS), which was a longitudinal intervention trial (Borg 2012, Sjöstrom 2013). Compared to the control group, for GBP an initial average weight loss of 27% was assumed. In the following years, only 75% of the reduction was applied to avoid overestimating the cost and survival benefits of BMI reduction. Surgical and post-surgical mortality were included in the model. Utility of GBP was quantified through QALY weight from the HODaR study, depending on the patient's BMI value (Ackroyd 2006). As for the costs, a great range of QALY gained was shown depending on age, BMI, and the gender of the patients (from 6.78 to 15.53 QALY for GBP and from 4.54 to 11.68 QALY for conventional treatment, with a lifetime horizon).

- The Swedish study of Borg et al. included three types of direct costs: the intervention cost of the surgical procedure including any plastic surgery required, cost of adverse events, and excess healthcare costs of treating obesity related diseases (Borg 2014). In addition, indirect costs related to productivity loss were also included. For surgery and adverse events costs, the prices for one of the major organisations performing GBP in Sweden as well as official hospital prices for the Southern Healthcare Region in Sweden were used. Excess healthcare costs related to obesity from previous publications of the same authors were used (Borg 2014). The indirect costs were based on a study of Neovius et al., indicating the incidence of sick leave by interval of sick-leave duration (Neovius 2012). Average salaries including social fees, primarily extrapolated from Statistics Sweden, and were applied to the incidence rates, to estimate the annual indirect costs for sick-leave. The conventional treatment alternative consisted of the prevalent mixture of non-surgical obesity treatments. The unit costs and sources of these treatments were not reported in the main article. Total lifetime direct costs for GBP and conventional treatment showed a great range of variation depending on age, BMI, and gender (SEK 63,000 to SEK 298,000 for GBP, and SEK 42,000 to SEK 887,000 for conventional treatment).

Campbell 2010

- Campbell et al. utilised a Markov model to simulate weight loss, health consequences, and costs for surgical treatment including laparoscopic Roux-en-Y gastric bypass and laparoscopic Gastric banding among morbidly obese, over a lifetime horizon. Patients aged 18-74 years who satisfied clinical eligibility criteria for bariatric surgery (BMI ≥ 40 kg/m^2 or BMI ≥ 35 kg/m^2 with comorbidities) were included. The study was conducted from the payers' perspective (US), costs were expressed in 2006 US dollars (USD), and QALYs and costs were discounted by 3% per annum. Data in relation to treatment weight loss were estimated from a prospective study by Angrisani et al., in which morbidly obese patients were randomized to LAGB and LRYGB and followed for 5 years post-surgery. In base-case analyses, it was assumed that BMI remained constant after the 5-year follow-up period. Patients receiving no treatment were assumed to maintain a constant BMI for the duration of the model. Additional data with regards to 5-year efficacy were sourced from a systematic review of 36 English-language studies of weight loss outcomes following LAGB or LRYGB by O'Brien et al. BMI-specific mortality rates were estimated based on age- and sex-specific all-cause mortality rates derived from US life tables, to which they applied BMI-specific relative risk ratios from an analysis of the National Health and Nutrition Examination Survey by Flegal et al.. Physician costs for the LAGB and LRYGB procedures were estimated using Current Procedural Terminology codes, and associated inpatient costs were estimated using the Healthcare Cost and Utilization Project (HCUP) Database with International Classification of Diseases, Ninth Revision codes. LRYGB was more beneficial than no treatment, providing an (LRYGB) discounted lifetime QALYs for each treated patient. The ICERs for LRYGB versus no treatment was \$ 14,300 per QALY gained for females and \$ 14,200 per QALY gained for males, among patients with morbid obesity (35-39.9 kg/m^2). Better results were achieved in terms of ICER per QALY gained for individuals with a higher initial body mass index compared to patients who had a lower initial BMI. In addition, the ICER per QALY gained was higher for older individuals. Furthermore, ICERs for men were generally higher compared to women. Sensitivity analyses showed these results to be robust to reasonable variation in model parameters and overall parameter uncertainty. The author's conclusion was that LAGB and LRYGB provided significant weight loss and were cost-effective compared with no treatment at conventionally accepted thresholds for medical interventions.

- In the study by Campbell et al., treatment efficacy was estimated from a prospective study by Angrisani et al., in which morbidly obese patients were randomized to LAGB and LRYGB and followed for 5 years post-surgery (Campbell 2010, Angrisani 2007). The study by Angrisani et al., despite having a small sample of patients (N=51), was judged to be the strongest available based on its randomized head-to-head design and minimal attrition over the 5-year study period. In addition, a study performed by O'Brien et al. was used as an alternative source for 5-year efficacy data (O'Brien 2006). Using the data from Angrisani et al., the cumulative change in BMI after LRYGB was -19.2% and -32.0% after one and five years respectively. In the base case analyses it was assumed that BMI remained constant after 5 years follow-up. Patients receiving no treatment were assumed to maintain a constant BMI for the duration of the model. Surgical and post-surgical mortality were included in the model. Health-related quality of life estimates were derived from EQ-5D data from the 2000 Medical Expenditure Panel Survey (Jia and Lubetkin 2005). Estimated utility decrements for surgery procedures were derived from utilities reported for laparoscopic surgery and for hernia repair in the United Kingdom (McCormack 2005). The lifetime QALY gained ranged from 16.4 to 19.7 and from 12.8 to 17.8, for LRYGB and no treatment, respectively.

- In the US study of Campbell et al., costs for surgery were divided into the initial procedure costs, complication/death costs, other medical expenditures, and follow-up costs (Campbell 2010). Main costs sources were derived from the Healthcare Cost and Utilization Project (HCUP) Database with International Classification of Diseases, Ninth Revision codes (inpatient costs), the guidelines of the American Society for Metabolic and Bariatric Surgery supplemented by input from a clinical expert (follow-up), the Physician Fee Guide (physician costs), the Red Book (drug costs), and a study of the Medical Expenditure Panel Survey ("Other" medical expenditures specific to each BMI level, Arterburn 2005). For the no-treatment group only "other" medical expenditures were taken into account. Lifetime costs for LRYGB ranged from $ 117,000 to $ 141,000 depending on gender and BMI. Costs for patients without treatment ranged from $ 88,000 to $ 132,000.

Castilla 2014

- Castilla et al. utilised a discrete-event simulation model with two groups: patients undergoing gastric bypass and patients following usual care (Castilla 2014). A cost-utility of gastric bypass versus usual care for severely obese patients in Spain from a payer's perspective was implemented. 79 patients (aged 18-55) with an average BMI of 50.7 kg/m^2 (ranging from 36.6 to 76.3 kg/m^2) and with comorbidities, such as stroke, coronary obstructive disease, and diabetes, were modelled within each arm. The main outcome of interest was QALY and the costs. Both costs and effects (year 2012) were discounted by 3%. Results were assessed over a time horizon of 5, 10, 15, and 20 years. Long term effects of gastric bypass were based on the average results of the Swedish Obese Subjects (SOS) study. Patients that were not operated were considered to remain in the same BMI range their whole lifetime, whereas operated patients would see their BMI modified as in the base study. The long-term treatment effect of GBP on BMI was based on the average SOS results from 6 to 20 years, (with a BMI reduction of approximately 25%). Probabilistic sensitivity analysis was used to determine the uncertainty around each of the parameters. Severely obese patients with a gastric bypass showed higher QALYs (18.18 vs. 12.55) and lower costs (€ 17,431 vs. € 31,425) when compared to usual care. Over 10 years, the ICER was € 3,920 per QALY gained. This study observed very low QALY differences in the first five years, which meant GBP was disfavoured for this period, whereas very large QALY differences, thereby favouring GBP, were reported over a life time horizon. The authors concluded that gastric bypass is a dominant strategy (cost-saving) when a lifetime horizon is used.

- In the study by Castilla et al., short-term efficacy was based on a Spanish study of Mar et al., in which 79 patients aged 18–55 with an average BMI of 50.7 kg/m^2 were selected, referred for GBP, and followed for 2 years after surgery (Castilla 2014, Mar 2013). Two years after surgery a 37% BMI reduction was reported. The long-term effect of GB on BMI was based on the average Swedish Obese Subjects (SOS) study results from 6 to 20 years, which revealed a sustained BMI reduction of approximately 25% (Sjöstrom 2013). Patients not operated were considered to remain in the same BMI range for their whole lifetime. Surgical and post-surgical mortality were included in the model. Utilities were estimated from the literature, by paying attention to Spanish studies using EQ-5D utilities (Castilla 2014, Mar 2011). Using a lifetime perspective, the QALYs gained with GBP and conventional treatment were 18.18 and 12.55, respectively.

- In the Spanish study by Castilla et al., only direct costs involving the Spanish National Health System (NHS) were considered (Castilla 2014). The resource use was estimated from the literature, whereas costs of GBP was estimated by assuming 50% of laparoscopic surgeries, and comprised of time of use of operating theatres, materials, and ICU and ward stays. Costs of comorbidities (including diabetes, acute myocardial infarction, stroke, and breast cancer) were extracted from several publications. Conventional treatment including diet, exercise, and behaviour modification therapy were used as a comparator. Using a lifetime horizon, the costs of GBP and conventional treatment were € 17,400 and € 31,400 respectively.

Clegg 2003

- Clegg et al. used a hypothetical cohort of 100 patients with 90 females over a time horizon of 20 years to examine the cost-effectiveness of surgery for morbid obesity. A systematic review of economic evaluations, prospective clinical trials and randomized control trials was performed beforehand (Clegg 2003). A National Health Services (NHS) perspective was used in general, and a personal social services perspective was used for costs and benefits. Sources of costs were from the years 1999-2000. Discount rates of 6% and 1.5% were used for the costs and QALYs, respectively. Patients with a BMI of > 40 kg/m^2 or BMI > 35 kg/m^2, with a serious comorbid diseases were included. The search resulted in three clinically effective surgeries: gastric bypass (Roux-en-Y), vertical banded gastroplasty, and adjustable gastric banding, versus non-surgery. A comparison of surgery with non-surgery resulted in ICERs of £ 6,289 per QALY gained for gastric bypass, £ 8,527 per QALY gained for silicone adjustable gastric banding, and £ 10,237 per QALY for vertical banded gastroplasty. Surgery compared with non-surgery was cost-effective at £ 11'000. This study found small QALY differences, despite the low discount rate of effects. The authors concluded that surgery is a cost-effective treatment in morbidly obese patients.

- In the study by Clegg et al., effectiveness was based on the results of a HTA including 17 published randomised controlled trials and one published cohort study, which was the SOS Study (Clegg 2003, Clegg 2002). For non-surgical treatment, a constant BMI of 45 kg/m^2 was assumed for a period of 20 years. For GBP, the baseline BMI of 45 kg/m^2 was reduced to 29 kg/m^2 for the first five years after surgery. Thereafter, from years 6 to 20, a BMI of 45 kg/m^2 was assumed. Surgical mortality was included in the model. Utility values were extrapolated from an economic evaluation of Orlistat (Clegg 2003, Hakim

2002). For one single unit change in BMI, a QALYs change of 0.077 was assumed. The estimated QALYs gained with 20 years follow-up was 11.67 for GBP, and 11.22 for usual care.

- The article of Clegg et al. is a summary of a more extensive HTA commissioned on behalf of the National Institute for Clinical Excellence (NICE) in the UK (Clegg 2003, Clegg 2002). Sources of costs were restricted to published information for 1999-2000 (National Health Services in Scotland Information and Statistics Division, and Scottish Health Service Costs). For GBP, the costs included pre-operative assessment (visits, specialist consultations), hospitalisation, complications, and 20 years follow-up. Costs for conventional treatment included physician visits and contacts with a dietician and practice/district nurses. The direct costs for GBP and usual care were estimated at £ 9,800 and £ 7,000, respectively, with a time horizon of 20 years.

Craig 2002

- Craig et al. utilized a deterministic decision model to compare lifetime expected costs and outcomes between gastric bypass and no treatment of severely obese patients from the US payers' perspective. The included patients were 35 to 55 old, with a BMI between 40 and 50 kg/m^2. Smokers, patients with cardiovascular risks, and patients that had already tried a conservative bariatric therapy were excluded. The estimates of weight loss and complication rates were based on those from a study by Pories et al., which involved a large sample (608 patients) and 14 years of follow-up with a 96.3% follow-up rate. Additionally, rates of abdominoplasty and reversal surgery were obtained from the Adelaide Study, a randomized clinical trial that compared the outcomes of gastric bypass with those of vertical banded gastroplasty over 5 years. Furthermore, the life expectancy was derived using data from the Framingham Heart Study (Thompson 1999), which estimated life expectancy across age, sex and body mass index. It considered only mortality related to coronary heart disease and stroke, and excluded patients with initial history of those conditions. The analyses incorporated medical costs (in US 2001 dollars) associated with the initial surgery, treatment of complications, follow-up care, and treatment of obesity-related diseases, such as coronary heart disease, stroke, type 2 diabetes, hypercholesterolemia, and hypertension. All cost estimates were adjusted for inflation. The Medical Care Component of the Consumer Price Index for All Urban Consumers was used to adjust prices, where necessary. Expected lifetime medical cost estimates were obtained from the published literature. Base case analysis of cost-effectiveness ratios ranged from $ 5,000 to $ 16,000 per QALY for women and from $ 10,000 to $ 35,000 per QALY for men. In general, female patients with a higher initial BMI showed better ICERs per QALY gained. Subgroup analyses showed that the cost-effectiveness ratios were affected by older patients, less obese men, quality of life, complication rates, perioperative mortality and in the variation in parameters like loss of excess weight. The authors concluded that gastric bypass is a cost-effective alternative compared to no treatment for severely obese patients.

- Craig et al. based their estimates for weight loss and complication rates on those from a study by Pories et al., which involved a large sample (608 patients) and 14 years of follow-up with a 96.3% follow-up rate (Craig 2002, Pories 1995). Pories et al. estimated a mean percentage loss of excess weight of about 58% five years after surgery (loss of excess weight was the amount of weight lost divided by the total amount of excess weight before the intervention, and was expressed as a percentage). Excess weight was defined as the weight above a body mass index of 22 kg/m^2. For no treatment, no change in BMI was assumed. Surgical and post-surgical mortality were included in the model. Using a simple linear approximation based on data from the Framingham Heart Study, the authors estimated life expectancy and lifetime QALYs depending on gender, age, and BMI (40 or 50 kg/m^2) (Thompson 1999). For example, it was estimated that a male who was 35 years old with a BMI of 40 kg/m^2 would gain 19.56 QALY with GBP, compared to 18.51 QALY with no surgery. Results were similar for females.

- In the US study of Craig et al., medical costs associated with the initial surgery, treatment of complications, follow-up care, and treatment of obesity-related diseases (e.g. coronary heart disease, stroke, diabetes, hypercholesterolemia, and hypertension) were included (Craig 2002). The Medical Care Component of the Consumer Price Index for All Urban Consumers was used to adjust prices, where necessary. Expected lifetime medical cost estimates were obtained from the published literature (Thompson 1999). For the majority of the remaining costs, estimates of nationally representative hospital charges were obtained from the Healthcare Cost and Utilization Project (HCUPnet). Cost of medications and follow-up visits were obtained from a source on wholesale drug prices (2000 Drug Topics Red Book, Montvale, NJ) and from a local source (R. Atkinson, MD, Clinical Nutrition Clinic, written communication, 2001). The lifetime costs for a male who was 35 years old with a BMI of 40 kg/m^2 was estimated to be $ 68,600 with GBP and $ 38,500 without surgery. Similar costs were found for females. For patients with a higher BMI (50 kg/m^2) the costs difference between surgery and usual care was less pronounced.

<u>Faria 2013</u>

- Faria et al. utilised a Markov model that was built to determine the cost-effectiveness of gastric banding and gastric bypass versus best medical management over a lifetime horizon from a societal perspective (Faria 2013). Results were derived for a representative population of 40 years old, morbidly obese patients with a mean BMI of 49.6 kg/m^2. Costs and utilities were discounted after the first cycle by 3.0%. Subgroup analyses were performed for groups with comorbidities, diabetes mellitus, different ages and BMI. Distributions, probabilities, costs and utilities were obtained from the literature or from an institutional database. Gastric bypass resulted in a dominant strategy with a decrease in lifetime costs and an increase in QALYs. When patients in the gastric bypass group were compared with those in the best medical management group (BMI of > 35 kg/m^2) a QALY gain of 1.9 with € 13,244 costs saved per patient were found. Based on the different BMI cut-offs(BMI 25-30 kg/m^2, 35-40 kg/m^2, 40-50 kg/m^2 and 50-70 kg/m^2) bariatric surgery was cost-saving if compared to conventional treatment. . For a BMI of 30-35 kg/m^2, bariatric surgery was cost-effective with an ICER of € 10,534 per QALY gained. The authors concluded that the patients in the gastric bypass surgery arm gained more QALY and were less expensive if compared to the best medical management treatment. However, this study was not considered transferable because of missing information with regards to year of cost data.

- The effectiveness and utility data used in the study of Faria et al. were retrieved from the literature wherever available, or retrieved from an institutional database. Among the selected literature, data from other cost-effectiveness studies were included (Keating 2009, McEwen 2010, Sjöstrom 2012, Craig 2002, Padwal 2011, Pollock 2011). Unfortunately, the effectiveness assumptions used for the cost-effectiveness modelling were not reported. The estimated overall QALY gained after 40 years follow-up was 16.66 for GBP and 14.48 for conventional medical treatment (Faria 2013).

- The types of costs used in the paper of Faria et al. were retrieved from the literature or from an institutional database (Faria 2013). The cost of surgery, complications, and comorbidities were included. Unfortunately, there was no detailed information about what type of costs were considered for GBP and conventional treatment (i.e. best medical management). The estimated overall costs after 40 years follow-up were € 29,300 for GBP and € 42,500 for conventional medical treatment.

Hoerger 2010

- The study by Hoerger et al. utilised a Markov model to assess the cost-effectiveness of bariatric surgery for severely obese adults with diabetes (Hoerger 2010). This study compared the costs, QALYs, and cost-effectiveness of bariatric surgery (gastric bypass and gastric banding) and usual diabetes care. The perspective of the study was not clear, and so it was assumed to be from a payers' perspective in the US. All costs were inflated to 2005 US dollars. The population was defined as people with a BMI > 35 kg/m^2, who had established or newly diagnosed diabetes. This model was based on the Centers for Disease Control and Prevention–RTI Diabetes Cost-Effectiveness, and the model was expanded to include bariatric surgery. The cost-effectiveness model was a Markov simulation model of disease progression and cost-effectiveness for type 2 diabetes that followed patients from diagnosis to either death or age 95 years. The model simulated development of diabetes-related complications on three microvascular disease paths (nephropathy, neuropathy, and retinopathy) and two macrovascular disease paths (coronary heart disease and stroke). The model outcomes included disease complications, deaths, costs, and QALYs. In the model, progression between disease states was governed by transition probabilities that depended on risk. First, the model allowed for diabetes remission and improvement, and important results of bariatric surgery. Second, the model included an annual probability of relapse from remission to diabetes. Third, the model accounted for perioperative mortality and the long-term effects of surgery on mortality. The results showed that bariatric surgery increased QALYs and costs. Bypass surgery had ICERs of $ 7,000 per QALY and $ 12,000 per QALY for severely obese patients with newly diagnosed and established diabetes, respectively. Banding surgery had ICERs of $ 11,000 per QALY and $ 13,000 per QALY for the newly diagnosed and established diabetes groups, respectively. In the sensitivity analyses, the cost-effectiveness ratios were most affected by assumptions about the direct gain in QoL from BMI loss following surgery. The study showed that gastric bypass and gastric banding are cost-effective methods of reducing mortality and diabetes complications in severely obese adults with diabetes.

- Clinical effectiveness and quality of life in Hoerger et al. was estimated using several sources (Hoerger 2010). The studies of Buchwald et al. (a meta-analysis) and Sjöstrom et al. (the SOS study) in particular provided the majority of the model parameters (Buchwald 2009, Sjöstrom 2004). For bypass surgery, a mean excess weight loss of 63.25% and a BMI loss of 16.17 kg/m^2 were assumed (the time unit for this context was unclear). A diabetes remission rate of 80.3% for persons with newly diagnosed diabetes was applied. For persons with established diabetes, a 40% rate of remission with 40% improvement following bypass surgery was applied. Assumptions for usual care were not stated. Surgical and post-surgical mortality were included in the model. Utility changes reported in five different studies were combined to estimate the mean increase in utility for a 1-unit reduction of BMI. The lifetime QALY gained in the case of newly diagnosed diabetes was 11.76 for GBP, compared to 9.55 for no surgery. Similarly, for established diabetes the lifetime QALY gained was 9.38 and 7.68 for GBP and no surgery respectively.

- In Hoerger et al., first-year bypass and banding surgery costs in the US were based on an analysis of Medstat claims by Eric A. Finkelstein et al. 2008, using unpublished data (Hoerger 2010). This analysis included surgery costs and any complications in the first year. Long-term follow-up costs (care visits, nutritional supplements, and long-term complications) were estimated mainly from using a publication of Parikh et al. and the opinion of an expert panel (Parikh 2006). Surgery was compared to usual diabetes care, which included tight glycemic control similar to that provided in the UK Prospective Diabetes Study (UKPDS). The lifetime costs for patients with newly or established diagnosed diabetes undergoing GBP was $ 86,700 and $ 99,900 respectively. For obese patients with no surgery, the costs were $ 71,100 for newly diagnosed patients and $ 79,600 for patients with established diabetes.

<u>Ikramuddin 2009</u>

- Ikramuddin et al. assessed the cost-effectiveness of Roux-en-Y Gastric Bypass in type 2 diabetes patients in the US, compared to standard medical management (Ikramuddin 2009). The model was designed to run over a 35-year period using a third-party payers' perspective, and used the same structure and predictive health economic model (the CORE Diabetes Model) to predict long-term costs and clinical effectiveness of Roux-en-Y gastric bypass as a treatment for type 2 diabetes, using the prospective observational study as a basis for their clinical effectiveness assumptions. The CORE Diabetes Model used a Monte Carlo simulation with tracker variables to estimate the lifetime costs and clinical outcomes of obese type 2 diabetes patients. A Minnesota cohort demographic was used along with baseline clinical characteristics, and other data was used from the standard "US medical management of diabetes" arm of the Core Diabetes, in both the bariatric surgery cohort and the medical management cohort. Costs and clinical outcomes were both discounted at 3% per annum. Characteristics of patient populations were that patients in the bariatric surgery observational study group were relatively young (mean age 50.1 years, with only 11.8% over the age of 60 years), predominantly female (77.9%) and Caucasian (83.5%). The average time since diagnosis with type 2 diabetes or pre-diabetes was 104 months (8.7 years), with a mean follow-up of 27.6 months (2.3 years). The baseline mean BMIs were high (48.4 kg/m^2, with only 12.8% of the patients having a BMI <40 kg/m^2). The mean discounted life expectancy for bariatric surgery was 11.54 years compared with 10.87 years for medical management of type 2 diabetes, representing an increase of 0.67 years. In terms of discounted QALYs, the mean for bariatric surgery was 6.78 compared with 5.88 for medical management of diabetes, representing an increase of 0.90 QALY gained. Mean lifetime direct medical costs for bariatric surgery were $ 83,482 per patient compared with $ 63,722 per patient for medical management of diabetes, representing an increase of $ 19,760. This led to $ 29,676 per life year gained and to $ 21,973 per QALY gained. This study showed that based in the base-case assumptions, Roux-en-Y gastric bypass was cost-effective in the treatment of type 2 diabetes in the United States with an ICER below $ 50,000 per QALY gained. In addition, it was suggested that bariatric surgery is not cost-effective over a shorter time horizon, or if the negative quality-of-life impact of increased BMI is ignored.

- In the article by Ikramuddin et al., data on the effectiveness of bariatric surgery was drawn from a prospective observational study conducted at an academic medical center in the United States (Minnesota cohort, unpublished data, University of Minnesota Medical Center, Minneapolis) (Ikramuddin 2009). Unfortunately, the effectiveness assumptions used for cost-effectiveness modelling were not reported in the paper (and the electronic appendix was not available). Based on the performed sensitivity analyses, it was assumed that the effects of surgery on BMI reduction were maintained constant for the 35-years of follow-up. Surgical and post-surgical mortality were included in the model. Utilities data came from the CODE-2 study (Bagust 2005). Over a 35-years' time horizon, 6.782 and 5.883 QALYs gained were reported for LRYGB and medical management respectively.

- In the analysis of Ikramuddin et al., direct costs of bariatric surgery costs were divided into surgery, management, and complication costs (Ikramuddin 2009). The same costs types were used for usual medical management (excluding surgery costs). Main sources were the Agency for Healthcare Quality and Research (Healthcare Cost & Utilization Project, HCUP), the MAG Mutual Healthcare Solutions Physicians' Fee and Coding Guide, the Drug Topics Redbook, the DRG Guidebook, and published literature (Parikh 2006). Overall total costs after 35 years follow-up were $ 83,500 for LRYGB and $ 63,700 for medical management.

Jensen 2005

- Jensen et al. utilised a hypothetical decision tree model to assess the costs of non-surgical and surgical weight loss interventions over a lifetime horizon from a social perspective (Jensen 2005). The model was structured with two health states (alive or dead), and if alive patients were transitioned to those who had reduced their BMI or not reduced. Gastric bypass surgery was compared to a non-surgical intervention which incorporated diet and exercise. All patients were 40 years old Caucasian women. Patients were assigned to gastric bypass surgery if they had a BMI of > 40 kg/m². Data for the nonsurgical arm was from a study by Heshka et al., where data from the gastric bypass arm was from the Swedish Obese Subjects (SOS) study (Heshka 2003, Sjöstrom 2000). The affected women in the hypothetical cohort had conventional treatment from the age of 18 onwards. Costs were calculated in US dollars and outcomes were presented as life expectancy and quality adjusted life years (QALY) gained. All costs were inflated to 2004. Gastric bypass surgery costs were based on costs of: surgical room, patient time, caregiver time, nursing, surgical equipment, anaesthesia, pharmaceutical and diagnostic tests. Non-surgical costs were based on a 1-hour per week weight loss program for 2 years and on attending costs. Costs related to complications and diseases of the obesity were included in both arms. The base-case analyses indicated that GBP surgery was cost-effective at $ 7,126 per QALY gained. Patients who had a gastric bypass were $ 4,600 more expensive, but these women had 0.61 years of additional life expectancy. In the best- and worst-case scenario analyses, ICERs of < $ 20,000 and < $ 35,000 per QALY were determined, respectively.

- In the study of Jensen et al., data about effectiveness for the GBP arm came from the Swedish Obese Subjects (SOS) Study, but was not reported in detail (Jensen 2005, Sjöstrom 2000). For the conventional treating group undergoing a commercial weight loss program, a weight reduction of 2.9 kg (6.4 lbs) was reported after 2 years. However, in cases of bad or no compliance, a weight gain at a rate similar to that of the overall population was assumed. Surgical and post-surgical mortality were included in the model. Quality of life calculations were made using data from a study by Hakim et al. that used time trade-off methods to estimate the utility for persons with BMIs of 30 to 34 kg/m², 35 to 39 kg/m², and 40 kg/m², and estimated a gain in utility for each BMI reduction of 1 kg/m² (Hakim 2002). The QALYs gained after GBP and non-surgical care were not reported in the publication.

- In the study by Jensen et al., costs for GBP were taken from a study of Nguyen et al. (Jensen 2005, Nguyen 2001). These costs included operating room, nursing, equipment, anaesthesia, pharmaceutical and diagnostic test costs. Costs associated with obesity-related diseases (i.e., hypertension, hypercholesterolemia, diabetes, coronary heart disease, and stroke) were derived for age and BMI groups using a previous study (Thompson 1999). Costs for non-surgical intervention included a 1-hour weekly weight loss program (including patient time and travel) and were derived from a publication of Heshka et al. and from the US Bureau of Labor Statistics (Heshka 2003). The overall costs of GBP and non-surgical care were not reported in the publication.

Keating 2009

- Keating et al. applied a Markov model that represented a treatment path of type 2 diabetes patients undergoing surgical intervention, compared to conventional therapy. Patients recently diagnosed with type 2 diabetes and with a BMI of 30-39.9 kg/m² were included. Laparoscopic adjustable gastric banding was used for the surgical intervention. A healthcare perspective was used over a lifetime horizon. Intervention costs were derived from 4 500 patients from the Centre for Obesity Research and Education, Monash University, Australia. Lifetime costs were based on published data and author input, and reported in Australian dollars (AUD). Recourse used for surgery was based on data of Keating et al. and treatment costs were extracted from the DiabCo$T study, with 10,600 patients from the Australian Diabetes supplies and services register. Utility weights and transition probabilities were sourced from the DiabCO$T study and SOS respectively, and a discounting rate of 3.0% per annum was applied. Patients with surgical treatment reported higher QALYs gained and lower costs if compared to conventional treatment. Based on the observed results, the authors concluded that surgical intervention was a dominant strategy for the treatment of obese patients (BMI of 30-39.9 kg/m^2) who were recently diagnosed with type 2 diabetes.

- In the study by Keating et al., data on intervention efficacy was derived from a 2-year randomized controlled trial involving 60 obese participants (BMI 30-40 kg/m^2) in Australia (Keating 2009, Dixon 2008). Trial intervention effects were primarily based on diabetes remission rates. The mean number of years in diabetes remission over a lifetime was 11.4 for surgery and 2.1 for conventional treatment (with a relative risk of remission in surgical vs. conventional therapy of 5.5). Surgical and post-surgical mortality were included in the model. Utilities were extrapolated from the DiabCo$t study, which used the EuroQol 5-domain (EQ-5D) questionnaire to assess QALYs (Colagiuri 2003). The lifetime QALY gained was 15.7 for LAGB and 14.5 for conventional treatment respectively.

- In the study by Keating et al., costs of intervention, maintenance, complications, diabetes monitoring/remission, and health care costs to treat diabetes were taken into account (Keating 2009). Surgical costs were sourced from a private hospital and private medical specialists to reflect private provision as is commonly practiced in Australia. For diabetes costs, the main source of unit costs was the Australian Government Department of Health and Ageing (in particular, the Australian 2006 Medicare Benefits Schedule and Pharmaceutical Benefits Schedule). The lifetime total costs for LAGB and conventional treatment were estimated at AUD 98,900 and AUD 101,400 respectively.

<u>Mäklin 2011</u>

- Mäklin et al. employed a combination of a decision tree with a Markov model to investigate the cost-utility of bariatric surgery in Finland from a healthcare provider perspective, with a time horizon of 10 years. The intervention and bariatric surgery incorporated three different techniques (gastric bypass, gastric banding and sleeve gastrectomy). Ordinary treatment as defined by the study incorporated a range of interventions from brief advice given by physicians to intensive conservative treatment, as there is no standard medical treatment for obesity in use in Finland. Parameter values for the model were taken from systematic literature reviews, randomized controlled trials, hospital discharge registers, and from a large representative population survey

(Health 2000) measuring health-related quality of life (HRQoL) and health service. Data on bariatric surgery volumes and costs were obtained from the hospital discharge register and hospital benchmarking database from the National Institute for Health and Welfare. The base-case scenario showed that bariatric surgery was more effective and less costly compared to ordinary treatment, using a 10 year time horizon, with a discount rate of 3.0%. Ordinary treatment costed €50,495, with 7.05 QALYs gained, while gastric bypass costed €33,379, with 7.67 QALYs gained. The gastric bypass resulted as a dominant strategy compared to ordinary treatment. Similarly, it was a cost saving when ordinary treatment was compared to sleeve gastrectomy and gastric banding. In the worst case scenario analyses, gastric bypass confirmed its strong dominance if compared to conventional treatment.

- In the study by Mäklin et al., the effectiveness of surgery was extracted from the literature (Mäklin 2011). In particular, the studies of Buchwald et al. (a meta-analysis) and Sjöstrom et al. (SOS study) provided the main information (Buchwald 2004, Buchwald 2009, Sjöstrom 2007). Starting from a BMI of 47 kg/m^2, the excess weight loss reported for GBP over a 10 year horizon was 60%, with a diabetes prevalence reduction of 82% (from 55% to 10% prevalence). The BMI in the ordinary treatment group was assumed to stay constant, based on the SOS study. Surgical and post-surgical mortality were included in the model. Quality-adjusted life-years was estimated based on life-years gained reported in randomized trials and 15D utilities from the Health 2000 Health Examination survey. The QALY gained over 10 years was estimated at 7.67 for GBP and 7.04 for ordinary treatment respectively.

- In the Finnish study of Mäklin et al., intervention costs and other average annual healthcare costs (including complications) were taken into account, whereas medication costs and productivity loss were excluded from the analysis (Mäklin 2011). Data on bariatric surgery volumes and costs were obtained from the hospital discharge register and hospital-benchmarking database from the National Institute for Health and Welfare. Annual healthcare costs were estimated from the Health 2000 Health Examination Survey data. The total cost over 10 years was estimated at € 33,400 for GBP and € 50,700 for ordinary treatment respectively.

<u>McEwen 2010</u>

- McEwen et al. investigated the cost, quality of life, and cost-utility of bariatric surgery in a managed care population of 221 patients undergoing bariatric surgery between 2001 and 2005 (McEwen 2010). Surgery versus non-surgery was compared using the same sampling (i.e. the costs and utility before surgery were used as a comparator for the costs and utility after surgery). The data used came from 221 patients who were enrolled in a 200,000 member independent practice association model managed care organization and who had underwent bariatric surgery between May 2001 and June 2005. A 2 year horizon and a lifetime horizon were used, with a discount rate of 3.0%. The 2 year time horizon cost-utility analysis showed an ICER of $ 48,662 per QALY gained. The lifetime horizon analysis suggested an ICER of $ 1,425 per QALY gained. In sensitivity analyses, bariatric surgery was more cost-effective in women, non-whites, more obese patients, and when performed laparoscopically. This study showed that the long-term cost-effectiveness appears to be dependent on the natural history and cost of late post-surgical complications and the natural history and cost of untreated morbid obesity. This study was not incorporated into the transferability and adaptation for Switzerland because there was no information about the year of the cost data.

- In the study by McEwen et al., effectiveness was based on the 221 patients included in the study (McEwen 2010). One year after surgery, the mean BMI fell from 51 to 31 kg/m^2 in women and from 59 to 35 kg/m^2 in men with substantial improvements in comorbidities. Long-term effects and assumptions were not described in detail, despite the lifetime horizon used. Post-surgical mortality was not included in the model. Two methods were used to assess the quality of life in the participating patients, namely the EQ-5D and the visual analogue scale. The lifetime QALY gained was 9.95 for GBP and 8.19 for usual care respectively.

- In the study by McEwen et al., total direct medical costs were subdivided into seven mutually exclusive categories including outpatient pharmacy, inpatient, outpatient clinic, diagnostic testing, laboratory testing, and emergency room costs, as well as a residual category for all other costs (McEwen 2010). Costs were determined from paid claims except for outpatient medication costs which were calculated from medications dispensed and average wholesale prices (Red Book: Pharmacy's Fundamental Reference, 2006 Edition). The lifetime costs for LRYGB and usual care were estimated to be $ 83,800 and $ 81,300 respectively.

Michaud 2012

- The cost-effectiveness study by Michaud et al. utilised an already published model (the “Future Elderly Model”) to develop and forecast long-term health and healthcare costs under different scenarios for medical technology development and utilization (Michaud 2012). The features of this model were described in detail in Goldman et al., Briefly, it ran as a microsimulation model, with individual level health trajectories, rather than average or aggregate health characteristics of a cohort. Medical interventions including Roux-en-Y gastric bypass or pharmacotherapy were compared to “no intervention”. The costs were adjusted to 2010 US dollars, with a discount rate of 3.0% over a life time horizon. The discount rate for effects was unclear. The hypothetical cohort was 50 years old, with presence of diabetes, heart disease, hypertension or ADL limitations, and with a BMI between 35 and 40 kg/m^2. Data with regards to costs (which incorporated total economic costs, rather than condition specific costs) and utilities were obtained from several literature sources or an institutional database. In the base-case analyses, gastric bypass showed an ICER below $ 10,000 per life year saved. Based on the observed results, the authors concluded that bariatric surgery generated substantial private value for those treated, in the form of health and other economic consequences.

- The effectiveness inputs in the model by Michaud et al. were mainly extracted from the publications of Dixon et al. (a randomised controlled trial), Picot et al. (HTA), and Sjöstrom et al. (SOS study) (Michaud 2012, Dixon 2008, Picot 2009, Sjöstrom 2004). The authors estimated that bariatric surgery achieved a permanent reduction in weight of 25% among those with a BMI over 40, or among those with a BMI between 35 and 40 with qualifying comorbidities. Post-surgical mortality was not included in the model. The utility of the intervention was quantified in terms of life years gained. Overall life years gained with a lifetime horizon was 30.35 for LRYGB and 28.8 for usual care.

- The US-based analysis undertaken by Michaud et al. included treatment and medication costs, as well as deadweight and income changes (Michaud 2012). Costs inputs were mainly based on the Medical Expenditure Panel Survey (MEPS, prior to age 65), the Medicare Current Beneficiary Survey (MCBS, after age 65), and on published literature (Cremieux 2008, Shekelle 2004). Lifetime costs including indirect costs were estimated at $ 296,400 for LRYGB and $ 283,700 for conventional treatment.

<u>Padwal 2011</u>

- Padwal et al. assessed and reviewed systematically the cost-effectiveness of bariatric surgery (Padwal 2011). Several data sources were used to identify randomized controlled trials (RCTs), cost-utility and cost-minimisation studies comparing contemporary bariatric surgery (adjustable gastric banding, Roux-en-Y gastric bypass, sleeve gastrectomy) to other surgical comparators or a non-surgical treatment, or any study reporting the association between surgical volume and outcome. Out of 1,904 citations, 13 economic analyses met the inclusion criteria. From the payer perspective, the long-term (10 years to lifetime) cost-utility of surgery compared with non-surgical management appeared attractive, with incremental cost–utility ratios (ICURs) ranging from $ 1,000 to $ 40,000 per QALY (in 2009 US dollars). The cost-effectiveness of bariatric surgery compared with non-surgical management appeared to be particularly favourable in patients with type 2 diabetes. Surgery resulted in a long-term incremental cost–utility ratio of $ <1,000 to $40,000 per QALY gained, when compared with non-surgical treatment.

- The study of Padwal et al. provided an overview of cost-effectiveness studies performed before 2011 (Padwal 2011). All cost-effectiveness papers described in their review are part of this report. Effectiveness or utility inputs were not reported in detail.

- Padwal et al. undertook a systematic review of the clinical and economic evidence of bariatric surgery, and simply provided an overview of cost-effectiveness studies performed before 2011 (Padwal 2011). All cost-effectiveness articles described in their review are part of this report (Ackroyd 2006, Anselmino 2009, Campbell 2010, Clegg 2003, Craig 2002, Hoerger 2010, Ikramuddin 2009, Jensen 2005, Keating 2009, McEwen 2010, Picot 2009, and Salem 2008). Costs were not reported in detail.

<u>Picot 2009</u>

- Picot et al. undertook a systematic review to assess the clinical effectiveness and cost-effectiveness of bariatric surgery for obesity (Picot 2009). The review was an update of a previous HTA (Clegg 2003). Seventeen electronic databases were searched from inception to August 2008. Two reviewers independently screened titles and abstracts for eligibility. Inclusion criteria were applied to the full text using a standard form. Interventions investigated were open and laparoscopic bariatric surgical procedures in widespread current use, compared with one another and with non-surgical interventions. The population comprised of adult patients with a BMI ≥ 30 kg/m² and of young obese people. Main outcomes were at least one of the following after at least 12 months follow-up: measures of weight change, quality of life (QoL), perioperative and postoperative mortality and morbidity, change in obesity-related comorbidities, and cost-effectiveness. For comparisons of surgery versus non-surgical procedures, eligible studies were RCTs, controlled clinical trials and prospective cohort studies (with a control cohort). In the economic model, the analysis was developed for three patient populations, namely those with a BMI ≥ 40 kg/m², a BMI ≥ 30 kg/m² and < 40 kg/m² with type 2 diabetes at baseline, and a BMI ≥ 30 kg/m² and <35 kg/m². For the last two groups, a more recent publication of the same authors is available (Picot 2012). Costs in terms of preoperative assessments, hospital admissions for surgery, managing adverse events and post-discharge care were included and adjusted for years 2007-2008. The perspective of the cost-effectiveness analysis was that of the UK National Health Service (NHS) and personal and social services, with maximum time horizon of 20 years, and a discount rate of 3.5 %. Surgical management was more costly than non-surgical management in each of the three patient populations analysed, but gave improved outcomes. For morbid obesity, base case incremental cost-effectiveness ratios (ICERs) ranged between £ 2,000 and £ 4,000 per QALY gained. They remained within the range regarded as cost-effective from an NHS decision-making perspective when assumptions for deterministic sensitivity analysis were changed. Deterministic and probabilistic sensitivity analyses produced ICERs that were generally within the range considered cost-effective, particularly for the long 20 year time horizon.

- In the first UK study by Picot et al., clinical effectiveness was mainly derived from a previous HTA of the same authors (Picot 2009, Clegg 2003). Moreover, effectiveness data from the Swedish Obese Subjects (SOS) study (Sjöstrom 2004), from Australian studies (O'Brien 2006, Dixon 2008), and from a RCT (Angrisani 2007) was used. For the first 5 years after LRYGB, a 36% reduction in initial weight was assumed. For years 5-10, a 17.7% decline in percentage of weight loss was used, as reported for the SOS study. Surgical mortality was included in the model. For the non-surgical group, a stable BMI over time was assumed. The change in utility was estimated using an economic evaluation of Orlistat (Hakim 2002), an Australian study (Dixon 2004), and other publications (Currie 2006, Lee 2005). The QALYs estimated for the 20 years following LRYGB were 12.32, compared to 10.8 for the non-surgical group respectively.

- The costs in the first study by Picot et al., included preoperative assessment (visits, specialist consultations), hospitalisations, complications, and 20 years of follow-up (Picot 2009). Main sources were from a previous HTA performed by the same authors (Clegg 2002), from published literature, and from expert opinions. The overall cost for patients with a BMI > 40 kg/m^2 were £ 19,824 and £ 13,561 for GBP and non-surgical intervention, respectively (for a 20 year time horizon).

Picot 2012

- In the paper published in 2012, Picot et al. investigated the cost-effectiveness of bariatric surgery compared to a non-surgical treatment for adults in the UK with mild (BMI 30-35 kg/m^2) and moderate (BMI 35-40 kg/m^2) obesity (Picot 2012). A systematic review was performed using 17 electronic resources. Studies meeting pre-defined criteria were identified, data-extracted and assessed for risk of bias using standard methodology. The model was populated with outcome data (in terms of weight loss and adverse events) from the systematic review. The perspective of the cost-effectiveness analysis was that of the UK National Health Service (NHS) and personal and social services, with maximum time horizon of 20 years, and a discount rate of 3.5 %. There were only two RCTs included (Dixon 2008, O'Brien 2006) in this systematic review, and data from these two clinical trials informed the modelling exercise. Evidence from both indicated a statistically significant benefit from laparoscopic adjustable banding (LAGB) compared to a non-surgical comparator for weight loss and in obesity-related comorbidity, among adults aged 41.8 to 46.6 years. Both interventions were associated with adverse events. LAGB costs more than non-surgical management. For people with class I or II obesity and type 2 diabetes, the ICER at 2 years was £ 20,159, reducing to £ 4,969 at 5 years and £1,634 at 20 years. Resolution of type 2 diabetes made the greatest contribution to this reduction. In people with class I obesity, the ICER was £ 63,156 at 2 years, £ 17,158 at 5 years, and £ 13,701 at 20 years. Cost-effectiveness results were particularly sensitive to utility gain from reduction in BMI, factors associated with poorer surgical performance and diabetes health state costs. Based on the results, bariatric surgery seems to be costs effective, for people with class I and class II obesity and type 2 diabetes, but not cost-effective for people with class I obesity.

- In the second UK study of Picot et al., clinical effectiveness was extrapolated from the Australian studies of Dixon et al. and O'Brien et al. (Picot 2012, Dixon 2008, and O'Brien 2006). Using the data from Dixon et al., at 2-year follow-up after LAGB, participants (BMI 37 kg/m^2 with diabetes, aged 20-60 years) lost 20.0% of the initial weight compared to a 1.4% reduction in patients with non-surgical therapy (excess weight lost 62.5%, compared to 4.3% respectively). 70% of the LAGB patients experienced diabetes remission (compared to 13% in the control group). Using data from O'Brien et al., at 2-year follow-up after LAGB, participants (BMI 33.7 kg/m^2 with comorbidities, aged 20-50) lost 21.6% of the initial weight compared to a 5.5% reduction in patients with non-surgical therapy (excess weight lost 87.2%, compared to 21.8% respectively). Mortality was not included in the model. The change in utility was estimated using a published model (Hakim 2002). Using a 20-year time horizon, the authors estimated 11.49 QALYs gained for LAGB, and 10.39 QALY gained for conventional treatment respectively.

- In the second study by Picot et al., costs of visits, surgery, hospitalisation, specialist consultation, physiotherapy, and complications were included for the LAGB cost estimations (Picot 2012). Usual care included out-patient visits and medical management for a weight loss program (including diet and pharmacotherapy). Main sources of unit costs were the Finance Department of the Southampton University Hospitals NHS Trust (SUHT), the Department of Health (NHS Reference Costs 2006–2007), and the Unit costs of Health and Social Care. The overall costs for diabetic patients with a BMI of 30-40 kg/m^2 were £ 35,100 for LAGB and £ 33,300 for usual care respectively (with a 20 year time horizon).

Pollock 2013

- Pollock et al. assessed the cost-effectiveness of laparoscopic adjustable gastric banding (LAGB) versus standard medical management (SMM) in obese patients with type 2 diabetes from a UK healthcare payer perspective over a 40 year time horizon. A validated computer model was used for the purpose of this exercise (CORE Diabetes Model, which is a non-product specific health policy analysis tool, comprising several inter-dependent semi-Markov sub-models, each modelling the progression of a diabetes-related complication). A non-parametric bootstrapping approach was used to model the progression of diabetes in a simulated 1'000 patient cohort. Monte Carlo simulation methods were then employed to calculate the mean and standard deviation of costs, life expectancy and quality-adjusted life expectancy over 1'000 iterations. For the base case analyses, cohort characteristics were taken from the Dixon et al. study, where the starting age of participants was 46.9 years. However, the baseline BMI for analysis (42.4 ±4.5 kg/m^2) was taken from the surgical arm of the Swedish Obese Subjects (SOS) study. Costs associated with gastric band placement were taken from the 2010 to 2011 NHS National Tariff using Healthcare Resource Group (HRG) code FZ05B. Additional resource use and information about post-surgical complication were extracted from Picot et al. and Salem et al., and health-related quality of life (HRQoL) utilities were sourced primarily from the UKPDS (Picot 2009, Salem 2008). Future costs and clinical outcomes were discounted at 3.5% annually and all costs were reported in 2010 pounds sterling. In the base-case analyses, the ICER was £ 3,602 per QALY gained (95% CI: £ 2,168 - £ 5,728) over a 40 year time horizon. Assuming a willingness-to-pay threshold of £ 20,000 per QALY gained (in line with commonly quoted thresholds in UK settings), the model projected a 100% likelihood that LAGB would be cost-effective when compared with standard medical management. One-way sensitivity analysis showed that the base case was broadly insensitive to changes in individual input parameters.

- In Pollock et al., treatment effects were taken from Dixon et al., in which diabetes remission was experienced by 73% and 13% of patients treated with LAGB and standard medical management respectively (Dixon 2008). The effects were applied in the first year of the modelling. Thereafter a natural course of risk factor progression based on data of the UK Prospective Diabetes Study (UKPDS) and Framingham studies was used. Surgical mortality was included in the model. Health-related quality of life (HRQoL) utilities were sourced primarily from the UKPDS and supplemented with diabetes-specific utilities from the CODE-2 study where necessary (Pollock 2013, Bagust 2005). The QALY gained over a 40 year time horizon was 10.05 for LAGB and 9.14 for usual care respectively.

- In the UK analysis undertaken by Pollock et al., costs of surgery, complications, diabetes, medication, and visits were included (Pollock 2013). Unit costs of diabetes complications were taken from a cost-effectiveness analysis in UK patients with type 2 diabetes (Baudet 2011). Costs of medications came from a June 2010 NHS Electronic Drug Tariff and a 2010 prescription cost analysis from the Health and Social Care Information Centre. Costs associated with gastric band placement were taken from the 2010 to 2011 NHS National Tariff using Healthcare Resource Group (HRG) code FZ05B. Additional resource use including dietitian visits, clinical psychology consultations, and contacts with a general practitioner and outpatient visits were extracted from Picot et al. (Picot 2009). Costs for post-surgical complication came from Salem et al. (Salem 2008). The overall cost over a 40 year life horizon was £ 23,600 for LAGB and £20,300 for usual care.

Salem 2008

- Salem et al. employed a deterministic decision analytic model to investigate the cost-effectiveness of laparoscopic gastric bypass (LRYGB) and laparoscopic adjustable gastric banding (LAGB), compared to non-operative weight loss interventions from a payer perspective (Salem 2008). The base-case analyses included morbidly obese male and female patients without obesity-related comorbidities, with a BMI of 40, 50, and 60 kg/m^2, and aged 35, 45, and 55 years. Life expectancy and lifetime medical costs were calculated across age, gender, and BMI strata using previously published data from the Framingham Heart Study and Third National Health and Nutrition Examination Survey, over a life time horizon. Usual-care medical costs (in U.S. 2004 dollars) associated with the surgery, including procedural fees, treatment of postoperative complications, follow-up care, and treatment of obesity-related diseases, were included. The medical care component of the Consumer Price Index for All Urban Consumers was used to adjust prices, where necessary. Utilities determined from patient gender, age, and BMI were derived from the 1997 National Health Interview Survey. The QALYs and costs were discounted at 3%. For both men and women, LRYGB was cost-effective at $ 25,000 per QALY gained when evaluating the full range of BMI values and estimates of adverse outcomes, weight loss, and costs. In the base-case analyses for men (aged 35 years with a BMI of 40 kg/m^2), the ICERs were $ 11,604 per QALY gained for LAGB and $ 18,543 per QALY gained for LRYGB when compared to the non-operative arm. For women (aged 35 years with a BMI of 40 kg/m^2), the ICER was $ 8,878 per QALY gained for LAGB and $ 14,680 per QALY gained for LRYGB. In the sensitivity analyses, the ICER per QALY gained for LRYGB was influenced by the rate of operative mortality, extent of weight loss, and operation cost. Based on the derived results, bariatric surgery seems to be cost effective strategy.

- In the study by Salem et al., life expectancy was calculated across age, gender, and BMI strata using previously published data from the Framingham Heart Study and Third National Health and Nutrition Examination Survey (Salem 2008, Thompson 1999). In the first 3 years after LRYBP, an excess weight loss of 71% was assumed. Lifetime assumptions were not clearly reported. For patients not undergoing surgery, a stable BMI was assumed. Surgical mortality was included in the model. Utilities were derived from the 1997 National Health Interview Survey, and were previously published by Craig et al. (Craig 2002). The QALYs gained for surgery and conventional treatment were not reported in the paper (only ICERs were provided).

- In the study by Salem et al., usual care medical costs associated with the surgery, including procedural fees, treatment of postoperative complications, follow-up care, and treatment of obesity-related diseases, such as coronary heart disease, stroke, type 2 diabetes, hypercholesterolemia, and hypertension were included (Salem 2008). The main cost sources were the Framingham Heart Study, the Third National Health and Nutrition Examination Survey (Thompson 1999), estimates of nationally representative hospital charges obtained from the Healthcare Cost and Utilization Project, and expert opinion. Costs of medications and follow-up visits came from a cost analysis of the Veterans Administration healthcare system (Gallager 2003). The overall costs for surgery and conventional treatment were not reported in the paper (only ICERs were provided).

Wang 2014

- Wang et al. used a two-part model, with a deterministic approach for the first 5-year period post-surgery and separate empirical forecasts for the natural history of BMI, with costs and outcomes assessed in the remaining years (Wang 2014). The cost-effectiveness analyses assessed the most common bariatric surgical procedures from a healthcare system perspective, over a life time horizon. The three most common bariatric surgical procedures in the US were laparoscopic gastric bypass (LRYGB), conventional (open) Roux-en-Y gastric bypass (ORYGB), and laparoscopic adjustable gastric banding (LAGB), and were compared to no surgery. The reference case was defined as a 53-year old female with BMI of 44 kg/m^2. A discount rate of 3% for costs and outcomes per annum was considered. The results were adjusted for inflation by converting all costs to 2010 US dollars using the Consumer Price Index. A combination of datasets including Medicare and MarketScan together with estimates from the literature to populate the model was used. In the base-case analyses, bariatric surgery produced additional life expectancy (80–81 years) compared to no surgery (78 years). The ICER for the surgical procedures were $ 6,600 per QALY gained for LRYGB, $ 6,200 for LAGB, and $ 17,300 for ORYGB. ICERs varied according to choice of BMI forecasting method and clinically plausible variation in parameter estimates. In most scenarios, the ICER did not exceed a threshold of $ 50'000 per QALY gained. The one-way sensitivity analyses showed that the parameters with the largest impact were BMI at baseline, age at the time of the procedure and gender. ICERs were decreasing with baseline BMI and the magnitude of weight loss after 5 years, but increasing with higher discount rate and higher rates of early complication. ICERs per QALY gained were also higher for men. Based on the derived results, bariatric surgery compared to no surgery is cost-effective for most patients eligible for bariatric surgery, assuming a willingness-to-pay threshold of $ 50,000 per QALY gained.

- In the US study by Wang et al., the effectiveness of surgery was drawn from a systematic review of literature examining the effectiveness of bariatric surgery (Wang 2014, Picot

2009). Surgical and post-surgical mortality were included in the model. Utility changes from 31 days until year five post-surgery conditional on age, BMI and gender were also estimated using a regression model based on previously published studies (Wong 2012, Arterburn 2008, and Franks 2003). The reported lifetime QALYs gained were 13.4 for LRYGB and 10.6 for no surgery.

- In the US study by Wang et al., direct medical costs included surgery, complications, and follow-up were included. The main source for unit costs was the Medicare claims database (2004–2008) (Wang 2014). The lifetime cost of LRYGB and no surgery was estimated at $ 169,100 and $ 150,900, respectively.