

The clinical effectiveness and cost-effectiveness of bariatric (weight loss) surgery for obesity: a systematic review and economic evaluation

J Picot,* J Jones, JL Colquitt,
E Gospodarevskaya, E Loveman,
L Baxter and AJ Clegg

Southampton Health Technology Assessments Centre, University of Southampton, UK

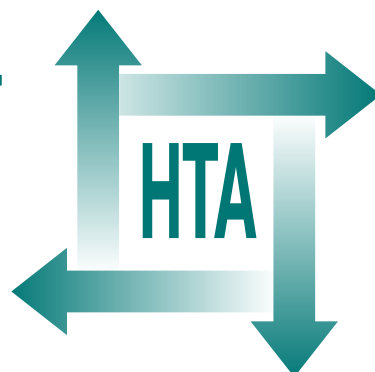
*Corresponding author



Executive summary

Health Technology Assessment 2009; Vol. 13: No. 41
DOI: 10.3310/hta13410

Health Technology Assessment
NIHR HTA programme
www.hta.ac.uk





Executive summary

Background

The prevalence of overweight and obesity among people in England and Wales is increasing. Associated serious health consequences in adults include Type 2 diabetes, cardiovascular disease, musculoskeletal disorders, certain cancers and increased mortality. Childhood obesity is associated with a higher chance of premature death and disability in adulthood. Obesity imposes a considerable economic burden on society. Weight loss improves obesity-related comorbidities and may have a mortality benefit. The intensity of intervention depends on the degree of obesity and presence of comorbidities. Management begins in primary care, but moves to the specialist setting when initial measures have failed and surgery is being considered. Bariatric (weight loss) surgery is increasing, but is not uniformly available across the country and a significant proportion is funded privately.

Objectives

To assess the clinical effectiveness and cost-effectiveness of bariatric surgery for obesity.

Methods

Data sources

Seventeen electronic resources, including MEDLINE, EMBASE and Cochrane, were searched from inception to August 2008. Bibliographies of related papers were assessed and experts were contacted to identify additional published and unpublished references.

Study selection

Titles and abstracts were screened for eligibility by two independent reviewers. Inclusion criteria were defined a priori and applied to the full text of retrieved papers by two independent reviewers using a standard form. The inclusion criteria were as follows:

- *Intervention* Open and laparoscopic bariatric surgical procedures in widespread current use.

- *Comparators* Surgical procedures in current use in comparison with one another; open surgery compared with laparoscopic surgery for the same procedure; surgical procedures in current use compared with non-surgical interventions (medical management, usual care or no treatment).
- *Population* Adult patients fulfilling the standard definition of obese [body mass index (BMI) of 30 or over] and young people who fulfil the definition of obesity for their age, sex and height.
- *Main outcomes* At least one of the following reported following a minimum of 12 months follow-up: measures of weight change; quality of life (QoL); perioperative and postoperative mortality and morbidity; change in obesity-related comorbidities; cost-effectiveness [reporting outcomes as either life-years or quality-adjusted life-years (QALYs)].

The study types that were eligible for inclusion were:

- *Systematic review of clinical effectiveness* Surgery versus surgery–randomised controlled trials (RCTs); surgery versus non-surgical procedure–RCTs, controlled clinical trials and prospective cohort studies (with a control cohort).
- *Systematic review of cost-effectiveness* Full cost-effectiveness analyses, cost–utility analyses, cost–benefit analyses and cost–consequence analyses.

Data extraction and quality assessment

Data extraction was undertaken by one reviewer and checked by two reviewers. Two reviewers independently applied quality assessment criteria. Differences in opinion were resolved through discussion at each stage.

Data synthesis

Studies were synthesised through a narrative review with full tabulation of the results of all included studies.

Economic model

The analysis was developed for three patient populations covered by studies included in the clinical effectiveness review:

- patients with morbid obesity BMI ≥ 40 undergoing adjustable gastric banding (AGB) or gastric bypass (GBP)
- patients with moderate-to-severe obesity (BMI ≥ 30 and < 40) with significant comorbidity at baseline (Type 2 diabetes) undergoing AGB
- patients with moderate obesity (BMI ≥ 30 and < 35) undergoing AGB.

A model developed previously was used for patients with morbid obesity (BMI ≥ 40), with updated assumptions on costs, diabetes incidence, permanency of diabetes remission following surgery, and on the impact of BMI on health-state utility.

A new model, including cardiac heart disease and stroke was applied in the analysis of AGB for moderate-to-severely obese (BMI ≥ 30 and < 40) patients with Type 2 diabetes and for moderately obese (BMI ≥ 30 and < 35) patients. The analysis was initially undertaken for the period of the trial follow-up only, but also included extrapolations up to 20 years following surgery.

Results

Quantity and quality of studies

A total of 5386 references were identified. Twenty-six studies (reported in 52 publications) were included in the review of clinical effectiveness. Three RCTs and three cohort studies compared surgery with non-surgical interventions; 20 RCTs compared different surgical procedures. Two studies focused on patients with a lower BMI (< 35 or < 40). The risk of bias of most of the trials was uncertain, only nine of the RCTs reported adequate sequence generation and only five reported adequate allocation concealment.

Summary of clinical effectiveness Surgery versus non-surgical interventions

The evidence indicates that bariatric surgery is a more effective intervention for weight loss than non-surgical options. Surgery led to a greater reduction in weight in all six studies and the difference was statistically significant in five studies reporting a statistical comparison. In the two RCTs that reported outcomes at two years, mean per cent initial weight loss in the surgical groups was 20%

and 21.6%, whereas the non-surgical groups had lost only 1.4% and 5.5% of their initial weight. In the two cohort studies reporting outcomes at two years, per cent weight change ranged from a weight loss of 16% to 28.6% in the surgical groups, but the non-surgical groups had gained weight with per cent weight change ranging from 0.1 to 0.5%. A large prospective cohort study [Swedish Obese Subjects (SOS) study] found that weight loss was still apparent 10 years following surgery, whereas patients receiving conventional treatment had gained weight. One RCT and one of two cohort studies assessing QoL found greater, and statistically significant, improvements after surgery on some measures, but not others. Two RCTs found that significantly fewer people had metabolic syndrome in the surgical group, and one found significantly higher remission of Type 2 diabetes following surgery. The SOS study found a statistically significant reduction in the incidence of three out of six comorbidities assessed at 10 years follow-up after surgery compared with conventional therapy.

Comparison of surgical procedures

Of the available surgical options assessed by RCTs there is evidence that GBP is more effective for weight loss than vertical banded gastroplasty (VBG) and AGB. Five of the seven included RCTs reported greater weight loss following GBP than VBG with per cent excess weight loss at one year ranging between 62.9% and 78.3% for GBP, and ranging between 43% and 62.9% for VBG. In two studies there was no statistically significant difference in 'success rate' or 'per cent ideal body weight'. One RCT found per cent excess weight loss of 66.6% was significantly greater up to five years following laparoscopic GBP than following laparoscopic AGB, which led to per cent excess weight loss of 47.5% ($p < 0.001$). Evidence from one RCT shows laparoscopic isolated sleeve gastrectomy to be more effective than AGB with greater excess weight loss up to three years (median per cent excess weight loss 66% versus 48%, $p = 0.0025$). GBP and banded GBP led to similar weight loss up to 24 months follow-up among patients with BMI > 50 (57.2% and 64.2%, $p = \text{ns}$) in the single RCT making this comparison. Comparisons of GBP and laparoscopic sleeve gastrectomy (LSG), and of VBG and AGB produced equivocal results. One RCT found slightly greater per cent excess weight loss with LSG (69.7%) than GBP (60.5%, $p = 0.05$) at 12 months, but no statistically significant difference in mean BMI or mean weight loss. Three RCTs found that measures of weight loss at one year follow-up favoured VBG over AGB, but longer-term results were conflicting. All the comparisons of open

versus laparoscopic surgeries (GBP four RCTs; VBG one RCT; AGB one RCT) found that both groups lost similar amounts of weight.

QoL was assessed by only two RCTs. One RCT found that QoL was significantly better following GBP than VBG on some items. The other found that there was no significant difference in QoL following either open or laparoscopic GBP.

Changes in comorbidities after surgery were assessed by five of the 20 RCTs. In general, comorbidities improved in all groups with no significant differences in improvements observed between different surgical interventions.

Adverse events

The extent of reporting of adverse events varied between studies; few were compared statistically and none were powered to do so. Fourteen RCTs reported no deaths. Where deaths were reported separately for each RCT trial arm, mortality ranged from 2% (1/51 patients receiving Open GBP within the first 30 postoperative days) to 10% (2/20 patients receiving Open GBP, one on the fourth postoperative day, one after 13 months). The large SOS study reported mortality of 0.25% in the surgical cohort (5/2010 patients within 90 days of surgery). Adverse events from conventional therapy included intolerance to medication, acute cholecystitis and gastrointestinal problems. Major adverse events following surgery, some necessitating reoperation, included anastomosis leakage, pneumonia, pulmonary embolism, band slippage and band erosion.

Summary of cost-effectiveness

All modelled economic evaluations assessed in this report found that bariatric surgery was cost-effective in comparison to non-surgical treatment although the variability in estimates of costs and outcomes is large. The results of the economic evaluation alongside a clinical trial were inconclusive. However, because of the numerous methodological shortcomings and some poorly justified modelling assumptions the reported results are unlikely to be reliable and generalisable estimates of the incremental cost-effectiveness of bariatric surgery in comparison to non-surgical treatment.

Summary of economic model

Surgical management with GBP or AGB of morbid obesity (BMI > 40) was more costly than non-surgical management, but results in improved

outcomes (in terms of QALYs) over the modelled 20-year time horizon. The incremental cost-effectiveness ratios (ICERs) ranged between £2000 and £4000 per QALY gained. The results were generally robust to changes in assumptions in the deterministic sensitivity analysis, and in all cases the ICERs remained within the range conventionally regarded as cost-effective from an NHS decision-making perspective.

Surgical management (with AGB) of moderate to severe obesity (BMI ≥ 30 and < 40) in patients with Type 2 diabetes was more costly than non-surgical management, but resulted in improved outcomes. The ICER reduced with a longer time horizon from £18,930 at two years to £1367 at 20 years. The results were generally robust to changes in assumptions in the deterministic sensitivity analysis. In the probabilistic sensitivity analysis the probability of surgical management being cost-effective (compared with non-surgical management) was 2.5% at a willingness-to-pay threshold of £20,000 per QALY and 50.6% at a willingness-to-pay threshold of £30,000 per QALY, for a two-year time horizon, and was 100% at both thresholds, for a 20-year time horizon.

Surgical management (with AGB) of moderate obesity (BMI ≥ 30 and < 35) was estimated to be more costly than non-surgical management, but resulted in improved outcomes, though the QALY gain at two years is small (0.08). The ICER reduced with a longer time horizon from £60,754 at two years to £12,763 at 20 years. There was considerable variability in results, in the deterministic sensitivity analysis, with ICERs above the range conventionally deemed acceptable in some scenarios even for longer time horizons. In the probabilistic sensitivity analysis the probability of surgical management being cost-effective (compared with an intensive medical programme) was 64% at a willingness-to-pay threshold of £20,000 per QALY and 98% at a willingness-to-pay threshold of £30,000 per QALY with a 20-year time horizon. In contrast, for a two-year time horizon, the probability of surgical management being cost-effective was zero at both thresholds.

Conclusions

Bariatric surgery appears to be a clinically effective and cost-effective intervention for moderately to severely obese people compared with non-surgical interventions. Uncertainties remain and further research is required, including:

- good-quality, long-term RCTs and cohort studies to provide detailed data on:
 - patient QoL to inform on the gains in utility associated with reduction in BMI
 - impact of surgeon experience on outcome of surgery
 - late complications leading to reoperation
 - more than one weight outcome measure with standard deviation about the mean reported to enable future meta-analysis
 - duration of remission of comorbidities and factors affecting this
 - resource use across the entire patient pathway to develop robust costings
- good-quality RCTs to provide evidence on bariatric surgery for young people and for adults with class I or class II obesity. New research must report on the resolution and/or development of comorbidities such as Type 2

diabetes and hypertension so that the potential benefits of early intervention can be assessed.

- A core set of important adverse bariatric surgery outcomes should be identified so that a standardised approach to describing adverse outcomes can be developed and their impacts on patients' QoL determined.

Publication

Picot J, Jones J, Colquitt JL, Gospodarevskaya E, Loveman E, Baxter L, *et al.* The clinical effectiveness and cost-effectiveness of bariatric (weight loss) surgery for obesity: a systematic review and economic evaluation. *Health Technol Assess* 2009;**13**(41).

NIHR Health Technology Assessment programme

The Health Technology Assessment (HTA) programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The research findings from the HTA programme directly influence decision-making bodies such as the National Institute for Health and Clinical Excellence (NICE) and the National Screening Committee (NSC). HTA findings also help to improve the quality of clinical practice in the NHS indirectly in that they form a key component of the 'National Knowledge Service'.

The HTA programme is needs led in that it fills gaps in the evidence needed by the NHS. There are three routes to the start of projects.

First is the commissioned route. Suggestions for research are actively sought from people working in the NHS, from the public and consumer groups and from professional bodies such as royal colleges and NHS trusts. These suggestions are carefully prioritised by panels of independent experts (including NHS service users). The HTA programme then commissions the research by competitive tender.

Second, the HTA programme provides grants for clinical trials for researchers who identify research questions. These are assessed for importance to patients and the NHS, and scientific rigour.

Third, through its Technology Assessment Report (TAR) call-off contract, the HTA programme commissions bespoke reports, principally for NICE, but also for other policy-makers. TARs bring together evidence on the value of specific technologies.

Some HTA research projects, including TARs, may take only months, others need several years. They can cost from as little as £40,000 to over £1 million, and may involve synthesising existing evidence, undertaking a trial, or other research collecting new data to answer a research problem.

The final reports from HTA projects are peer reviewed by a number of independent expert referees before publication in the widely read journal series *Health Technology Assessment*.

Criteria for inclusion in the HTA journal series

Reports are published in the HTA journal series if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

The research reported in this issue of the journal was commissioned and funded by the HTA programme on behalf of NICE as project number 08/06/01. The protocol was agreed in June 2008. The assessment report began editorial review in January 2009 and was accepted for publication in April 2009. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the referees for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

The views expressed in this publication are those of the authors and not necessarily those of the HTA programme or the Department of Health.

Editor-in-Chief: Professor Tom Walley CBE
Series Editors: Dr Aileen Clarke, Dr Chris Hyde, Dr John Powell,
Dr Rob Riemsma and Professor Ken Stein

ISSN 1366-5278

© 2009 Queen's Printer and Controller of HMSO

This monograph may be freely reproduced for the purposes of private research and study and may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising.

Applications for commercial reproduction should be addressed to: NETSCC, Health Technology Assessment, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk), on behalf of NETSCC, HTA.

Printed on acid-free paper in the UK by Henry Ling Ltd, The Dorset Press, Dorchester.