The real-world cost-effectiveness of bariatric surgery for the treatment of severe obesity: a cost–utility analysis

Erica L.W. Lester MD MSc, Raj S. Padwal MD MSc, Daniel W. Birch MD MSc, Arya M. Sharma MD DSc, Helen So PhD, Feng Ye MSc, Scott W. Klarenbach MD MSc

Abstract

Background: Severe obesity is associated with adverse health outcomes and increased risk of death. This study evaluates the real-world cost–utility of therapy for severe obesity, from the publicly funded health care system and societal perspectives.

Methods: We conducted a cost–utility analysis using primary data from a prospective observational cohort of adults living with severe obesity (BMI ≥ 35 kg/m² and a major medical comorbidity or BMI ≥ 40 kg/m²) who were enrolled in a regional obesity program over 2 years. We extrapolated 10-year and lifetime Markov models, validated and supplemented with literature sources, to compare medical, surgical and standard care therapies. We performed deterministic and probabilistic sensitivity analyses.

Results: The cohort included 500 adults living with severe obesity, 150 of whom received laparoscopic surgical therapy. From a publicly funded health system perspective, at 2 years, surgical therapy had an incremental cost–effectiveness ratio (ICER) of $54,456 per quality-adjusted life-year (QALY) compared with standard care therapy. Over a lifetime, it had an ICER of $14,056 per QALY. From the societal perspective, at 2 years, surgical therapy had an ICER of $340 per QALY; over a lifetime, it was the dominant option. The results were robust to sensitivity analysis.

Interpretation: From a public health care perspective, surgery for severe obesity is cost effective, and when approached from a societal perspective, it becomes cost saving. Real-world data support using surgical therapy for severe obesity, and our results contribute to the health economic and clinical literature with regard to a robust analysis from a societal perspective.

Severe obesity, defined as a body mass index (BMI) greater than 35 kg/m², is associated with adverse health outcomes, increased health care resource use and reduced home and workforce productivity. Numerous approaches to treat obesity have been developed, including lifestyle modifications, pharmacotherapies and surgical interventions. Bariatric surgery has emerged as one of the most clinically effective options, with numerous studies demonstrating its effectiveness.

Although many studies have examined cost-effectiveness from the health payer perspective, a broader perspective that fully captures all the costs and consequences of obesity and its treatment is lacking. Failing to consider the societal perspective can underestimate opportunity costs and misdirect resource allocation; this is especially relevant when assessing interventions that increase productivity among the working-age population and their quality of life. Further, most analyses use estimates from multiple sources, which may not reflect pragmatic, real-world resource use and outcomes. Failure to incorporate factors such as compliance or outcomes of treatment programs, outside of a rigorous study protocol, may lead to biased estimates of the cost-effectiveness of bariatric treatment programs.

Using prospective, empirical data from a regional bariatric program, we conducted an economic evaluation of surgical and medical therapy compared with standard care from both the public health care system and societal perspectives over 2 years, and extrapolated to 10-year and lifetime time horizons.

Methods

Study design
We used rigorously collected data from a pragmatic clinical treatment program study, the Alberta Population-based
Prospective Evaluation of the Quality of Life Outcomes and Economic Impact of Bariatric Surgery (APPLES) to inform this cost-utility analysis. A detailed study protocol and results of the APPLES study, which was conducted from 2008 until 2010, have been published (Appendix 1, available at www.cmajopen.ca/content/9/2/E673/suppl/DC1).1,3,15–17

In the APPLES study, data were captured prospectively at 6, 12, 18 and 24 months after enrolment, including demographics, anthropometrics, health-related quality of life, medication use, and detailed health care resource use (through linkage with administrative data). Patient surveys also captured costs, including transportation needs, out-of-pocket health-related purchases, attendance at work and enrolment in government support programs.1,5 We used outcomes from the 24-month data collection in our modelling study.

Cohort
We conducted our analysis using data from the participant cohort in the APPLES study.1 This study included adult patients (18–60 yr) enrolled in a regional obesity program, with BMI levels ≥ 35 kg/m² and a major medical comorbidity or with BMI levels ≥ 40 kg/m². Patients were assigned to 1 of 3 groups: medical therapy, surgical therapy or standard care (Table 1).3 The study sample size of 500 adult patients living with severe obesity was determined a priori. The standard care group served as a control group and consisted of patients who were on the program’s waitlist, with no direct weight loss therapy administered. Of the 150 patients who had surgery, 48 had an adjustable gastric band, 51 had a sleeve gastrectomy and 51 had a Roux-en-Y gastric bypass. All surgeries were conducted laparoscopically.1 All patients were deemed surgical candidates, as per the criteria used in Alberta.18

Model structure
We assessed outcomes at 2 years by treatment group using primary data from the APPLES study. We created Markov models using 10-year and lifetime time horizons, applying a half cycle correction. We allocated patients to the health states of death, diabetes, hypertension, diabetes and hypertension or no obesity-related comorbidity, as per the results of the 2-year cohort study; patients could transition between states during each 1-year cycle (Figure 1). We applied baseline characteristics for each treatment arm at the completion of the APPLES study, as well as initial cost and quality-adjusted life-years (QALYs), to the respective health states to account for the changes that occurred during the study.

We obtained transition probabilities between health states from primary data; when unavailable from primary data, we used probabilities informed by focused literature review (Appendix 2, available at www.cmajopen.ca/content/9/2/E673/suppl/DC1). The approach to determining probabilities is outlined in Appendix 1. We assessed costs and effects at a discount rate of 5%, as per convention.19

Model parameters
For the 2-year assessment, we used utility values from the APPLES cohort, measured using the EQ-5D-3L survey and derived via the US valuation system.20 For the 10-year and lifetime models, we extrapolated utility measurements by performing multivariable analysis on the APPLES data (using the mean BMI for each health state) and using the resultant regression equation to calculate the utility associated with each health state by age. We validated our calculated values by comparing them to those collected by the Health Quality Council of Alberta; the values from the APPLES data were minimally and consistently lower than those from the Health Quality Concil. Although the data from the Health Quality Council did not differentiate comorbidities, this finding increased our confidence in the extrapolation.

We defined the health care system and societal perspectives as per the Canadian Agency for Drugs and Technologies in Health guideline (details concerning cost determinations are in Appendix 1).19 The former includes costs to the publicly funded health care system, patients and their families. The societal perspective includes the components of the health care system perspective, combined with the direct costs

| Table 1: Baseline characteristics of treatment groups upon completion of APPLES study at 2 years (entrance into 10-year and lifetime models) |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| Characteristic  | Standard care n = 150 | Medical therapy n = 200 | Surgical therapy n = 150 | All patients n = 500 |
| Age, yr, mean ± SD | 45.6 ± 9.2 | 46.0 ± 10.0 | 45.3 ± 9.5 | 45.7 ± 9.6 |
| BMI, kg/m², mean ± SD | 48.8 ± 8.0 | 46.5 ± 8.2 | 36.4 ± 7.6 | 44.8 ± 9.0 |
| Sex, female | 136 (90.7) | 174 (87.0) | 131 (87.3) | 441 (88.2) |
| Hypertension | 96 (64.0) | 118 (59.0) | 66 (44.0) | 280 (56.0) |
| Diabetes | 71 (47.3) | 82 (41.0) | 33 (22.0) | 187 (37.4) |

Note: BMI = body mass index, SD = standard deviation.
*Unless indicated otherwise.
to other publicly funded services and productivity costs.\textsuperscript{19} We operationalized this definition by adding the cost of the health care system perspective, the costs of income transfer payments (i.e., income for disabled persons, disability and employment insurance benefits), as well as the cost of productivity loss, as measured by the friction method.\textsuperscript{21}

Beyond 2 years, we extrapolated the cost of each health state using APPLES data, assuming costs observed for a health state in the second year would persist. Assumptions associated with the cost extrapolation are outlined in Appendix 2. All costs are expressed in 2016 Canadian dollars. We employed the health-specific Consumer Price Index to adjust costs related to health care, as required.\textsuperscript{22} We adjusted other costs using the Bank of Canada Inflation Calculator.\textsuperscript{23}

**Statistical analysis**

We report QALYs and the incremental cost-effectiveness ratio (ICER) for each treatment group, compared to standard care. For probabilistic sensitivity analyses, we used Monte Carlo simulation with 10 000 iterations. We applied discount rates of 3\% and 0\% in sensitivity analysis. We conducted all modelling using Treeage Pro (2018), and all statistical analysis using Stata 13. We assessed face, internal and external validity.\textsuperscript{24} We report the willingness-to-pay value at which an intervention (medical or surgical) became cost effective for each perspective and time horizon.

In the 2-year model, we conducted a probabilistic sensitivity analysis using patient-level costs and its distributions. Transition probabilities were samples from $\beta$ distributions.

In the 10-year and lifetime model, we assessed all variables via 1-way sensitivity analysis, comparing surgical or medical therapy to the reference group (standard care) using the ranges in Appendices 1, 2 and 3, available at www.cmajopen.ca/content/9/2/E673/suppl/DC1. We also conducted probabilistic sensitivity analysis.

**Ethics approval**

The initial APPLES study and this cost–utility analysis were approved by the University of Alberta Health Research Ethics Board (Pro00003594 and Pro00060692).
Research

The annual costs and utility estimates by health state for the 500 patients in the APPLES study are reported in Appendices 3 and 4, available at www.cmajopen.ca/content/9/2/E673/suppl/DC1, respectively. At 2 years, the absolute QALYs were 1.37, 1.56 and 1.69 for standard care, medical and surgical therapy, respectively (Appendix 5, available at

Figure 2: Cost-effectiveness acceptability curves for the 2-year, 10-year and lifetime time horizons from the (A) publicly funded health care system and (B) societal perspectives. The probability of cost effectiveness represents the proportion of simulations that resulted in an incremental cost effectiveness ratio less than a given willingness-to-pay ratio.

Results

The annual costs and utility estimates by health state for the 500 patients in the APPLES study are reported in Appendices 3 and 4, available at www.cmajopen.ca/content/9/2/E673/suppl/DC1, respectively. At 2 years, the absolute QALYs were 1.37, 1.56 and 1.69 for standard care, medical and surgical therapy, respectively (Appendix 5, available at
Comparing surgery with standard care, the model is sensitive to the cost of having both diabetes and hypertension in the standard care group. The probabilistic sensitivity analysis indicated that, from the health care system perspective, surgical therapy was preferred at willingness-to-pay thresholds above $23,000 (Figure 2).

Using a societal perspective and a lifetime time horizon, surgery was the lowest cost option compared with medical and standard care (Appendix 5). In 1-way sensitivity analysis, when comparing surgery with standard care, the increasing cost of having no comorbidities in the surgery arm altered the results such that surgery had a positive ICER (Appendix 6). In probabilistic sensitivity analysis, surgical therapy is preferred at all willingness-to-pay values (Figure 2).

**Interpretation**

Using prospectively collected, real-world data, we calculated the cost and utility of surgical therapy, medical therapy and standard care for the treatment of severe obesity, from both the public health care system and societal perspectives. We have demonstrated that the perspective taken alters the results of this comparison, with surgical and medical treatment becoming more attractive from the societal perspective. This is most pronounced in the short run; as time progresses, the overarching results begin to converge and surgery emerges as the most cost-effective method across a wide range of willingness-to-pay thresholds. From the societal perspective, surgery becomes the lowest cost option over time, dominating medical therapy and standard care. Surgery also exhibits the highest level of utility gains, and therefore increases quality of life more than medical or standard therapy.

Our finding that bariatric surgery is either cost-effective or cost-saving is congruent with previous published reports. However, many previous studies do not consider the societal perspective; as obesity has effects on many facets of life, failing to take into account the breadth of this impact reduces the applicability of these studies. Moreover, many studies do not consider comparative therapies or real-world data, making our study more informative and valuable for decision-makers. Finally, application of a lifetime time horizon better reflects the long-term impact of obesity interventions, an approach often not pursued in previous studies.

By creating a model that directly reflects prospectively collected data, and using each individual’s data points in custom distributions, the 2-year model incorporates real-world heterogeneity and variability, including the costs of surgical complications. This approach reduces parameter and model uncertainty, and increases the validity of the probabilistic sensitivity analysis. For our long-term models, using real-world data and sources such as the Framingham Risk Score and the Health Quality Council of Alberta to externally validate parameter estimates, we reduced parameter uncertainty.

Conservative assumptions of the long-term effects of surgery and the impact of untreated or ineffectively
treated obesity likely reduced the magnitude of difference in mortality and comorbidity rates between bariatric surgery and the other treatment arms. This was supported by the 1-way sensitivity analysis; although changes in parameters rarely altered the direction of the ICER, the magnitude was often sensitive to mortality rates for different health states. Conservative assumptions likely underestimate the cost-effectiveness of bariatric surgery relative to other treatment approaches.

In general, model results were robust to parameter changes. It is noteworthy that, in every 1-way sensitivity analysis, the cost of having hypertension and diabetes in the standard care group was one of the most sensitive model parameters. Cost estimates for this parameter have a wide range, with a mean value that is notably higher than most other non-death states, which likely contributes to its impact on ICER. From a conceptual standpoint, this emphasizes how the cost of comorbidities can affect overall costs; reduction in the prevalence of hypertension and diabetes via therapy reduces both health care and societal costs.

Limitations
We used a friction approach to calculate societal costs, as per the Canadian Agency for Drugs and Technologies in Health guideline. Relative to the human capital approach, this may underestimate societal costs, particularly in the standard care and medical therapy groups, as there were greater levels of disability and unemployment in those groups. Hypertension and diabetes were chosen as health states because they have been illustrated to be drivers of cost and death. However, we omitted numerous other health states that are common among patients living with obesity, such as sleep apnea or coronary artery disease. It is logical to assume that the costs associated with other common diagnoses are included in the analysis, and exist in each treatment arm at the rate that they would coexist with hypertension or diabetes. Given the multifaceted nature of obesity, it is impractical to label each possible comorbidity, and their iterative combinations, as health states. Furthermore, the number of probability estimations would increase uncertainties in the model.

The APPLES study was conducted before the introduction of pharmaceuticals such as liraglutide and combination naloxone and bupropion. These are being used peripheratively and instead of surgery, and will therefore likely affect both cost and outcomes in medical and surgical groups. Future studies will be needed to understand the long-term clinical and economic impact of these technologies.

This study was conducted using single centre, observational data. Although this institution uses up-to-date, multidisciplinary approaches, with outcomes comparable to those in the literature, this study does not account for variation in practice between centres, which may influence cost and outcome.

Conclusion
Bariatric surgery resulted in the greatest gains in health-related quality of life. From the publicly funded health care system perspective, surgery is cost effective. From a societal perspective, surgery becomes cost saving. These findings, using real-world data, support using surgical therapy for severe obesity, and provide a robust analysis from the societal perspective.

References

**Affiliations:** Departments of Surgery (Lester, Birch) and Medicine (Padwal, Sharma, So, Ye, Klarenbach), University of Alberta, Edmonton, Alta.

**Contributors:** All of the authors contributed to the conception and design of the work. Raj Padwal, Daniel Birch, Arya Sharma and Scott Klarenbach acquired the data, which Erica Lester, Helen So and Feng Ye analyzed and interpreted. Erica Lester drafted the manuscript. All authors revised the manuscript critically for important intellectual content, gave final approval of the version to be published and agreed to be accountable for all aspects of the work.

**Funding:** No funding was received for this study.

**Content licence:** This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY-NC-ND 4.0) licence, which permits use, distribution and reproduction in any medium, provided that the original publication is properly cited, the use is noncommercial (i.e., research or educational use), and no modifications or adaptations are made. See: https://creativecommons.org/licenses/by-nc-nd/4.0/

**Data sharing:** The data used in this study are proprietary and are not available to other researchers.

**Supplemental information:** For reviewer comments and the original submission of this manuscript, please see www.cmajopen.ca/content/9/2/E673/suppl/DC1.