THE ECONOMICS OF ENHANCED RECOVERY

Lawrence Lee, MDCM MSc
Department of Experimental Surgery
McGill University, Montreal
August, 2014

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

This thesis would also not have been possible without the continuing support and leadership from Dr. Liane Feldman, who is the chair of the Steinberg-Bernstein Centre for Minimally-Invasive Surgery and Innovation, and my supervisor. I am amazed by her success as a true surgeon-scientist, and deeply grateful for her mentorship. Dr. Feldman was the first to nudge me towards the path of an aspiring academic that I am traveling today. Throughout the past years she has provided the encouragement and opportunities that have allowed the successes that I have experienced. I owe any current and future success to her.

I would like to thank Dr. Eric Latimer, my co-supervisor in the Department of Epidemiology, for his support in making this thesis possible. I cannot say enough nice things about the thorough and thoughtful manner in which helped me through all steps of this thesis. All of the work presented here would not have been possible without his expertise.

I also must thank everyone who was involved in the conduct of the trials that has provided me with the data and means to complete this thesis. I am also grateful to all of the other members of the Steinberg-Bernstein Centre for Minimally-Invasive Surgery and Innovation and the co-authors. Last but not least, I am grateful to my family and friends, and especially my wife, Katherine McKendy, for their support throughout this journey.

Thank you.

AUTHOR CONTRIBUTIONS

I have made a substantial contribution to all of the co-authored papers contained within this thesis. I developed the original research questions in collaboration with my thesis supervisors, Dr. Liane S. Feldman and Dr. Eric Latimer. I was responsible for the definition of the thesis objectives, which were approved by my thesis committee. With the guidance of my supervisor, I developed the study design and methods used in each of the following manuscripts. The contributions of the co-authors of each manuscript in this thesis are described below:

Lee L, Li C, Landry T, Latimer E, Carli F, Fried GM, Feldman LS. A systematic review of economic evaluations of enhanced recovery pathways for colorectal surgery. *Ann Surg* 2014; 259(4): 670-6

Study conception and design: Lee, Li, Latimer, Feldman

Data acquisition: Lee, Li, Landry

Analysis and interpretation of data: Lee, Li, Landry, Latimer, Carli, Friend, Feldman

Drafting of manuscript: Lee, Li, Landry

Critical revision: Latimer, Carli, Fried, Feldman

Lee L, Tran T, Mayo NE, Carli F, Feldman LS. What does it really mean to "recover" from an operation? *Surgery* 2014;155(2): 211-6

Study conception and design: Lee, Tran, Mayo, Carli, Feldman

Data acquisition: N/A

Analysis and interpretation of data: N/A

Drafting of manuscript: Lee, Tran

Critical revision: Mayo, Carli, Feldman

Lee L, Elfassy N, Li C, Latimer E, AS Liberman, Charlebois P, Stein B, Carli F, Fried GM, Feldman LS. Valuing postoperative recovery: validation of the SF-6D health state utility. *J Surg Res* 2013;184:108-14

Study conception and design: Lee, Li, Latimer, Carli, Fried, Feldman

Data acquisition: Lee, Elfassy, Li, Liberman, Charlebois, Stein

Analysis and interpretation of data: Lee, Li, Latimer, Carli, Fried, Feldman

Drafting of manuscript: Lee, Elfassy, Li

Critical revision: Latimer, Liberman, Charlebois, Stein, Carli, Fried, Feldman

Lee L, Mata J, Augutin B, Carli F, Morin N, Latimer E, Feldman LS. A comparison of the validity of two preference-based health indexes as measures of postoperative recovery. *J Surg Res* 2014 (epub 2014 Feb 15) DOI 10.1016/j.jss.2014.02.016

Study conception and design: Lee, Carli, Latimer, Feldman

Data acquisition: Lee, Mata, Augustin

Analysis and interpretation of data: Lee, Mata, Carli, Morin, Latimer, Feldman

Drafting of manuscript: Lee, Mata, Augustin

Critical revision: Carli, Morin, Latimer, Feldman

Lee L, Li C, Ferri LE, Robert N, Latimer E, Carli F, Mulder DS, Fried GM, Feldman LS. Economic impact of an enhanced recovery after surgery pathway for oesophagectomy. *Br J Surg* 2013; 100(10): 1326-34

Study conception and design: Lee, Li, Ferri, Latimer, Carli, Feldman

Data acquisition: Lee, Li, Ferri, Robert, Mulder

Analysis and interpretation of data: Lee, Li, Ferri, Latimer, Carli, Fried, Feldman

Drafting of manuscript: Lee, Li, Robert

Critical revision: Ferri, Latimer, Carli, Mulder, Fried, Feldman

Lee L, Mata J, Ghitulescu GA, Boutros M, Charlebois P, Stein B, Liberman AS, Fried GM, Morin N, Carli F, Latimer E, Feldman LS. Cost-Effectiveness of Enhanced Recovery versus Conventional Perioperative Management for Colorectal Surgery. *Annals of Surgery* 2014 (under review)

Study conception and design: Lee, Ghitulescu, Fried, Morin, Carli, Latimer, Feldman

Data acquisition: Lee, Mata

Analysis and interpretation of data: Lee, Ghitulescu, Boutros, Charlebois, Stein, Liberman, Fried, Morin, Carli, Latimer, Feldman

Drafting of manuscript: Lee, Mata

Critical revision: Ghitulescu, Boutros, Charlebois, Stein, Liberman, Friend, Morin, Carli, Latimer, Feldman

STATEMENT OF ORIGINALITY

The work presented in this thesis represent original contributions to the evidence base supporting enhanced recovery perioperative management for patients undergoing abdominal surgery. Specifically, it contributes to the areas of psychometrics and health technology assessment in the domain of enhanced recovery after surgery.

While I have received support from my supervisors and committee members, and input from co-authors for each study, the data presented in the following chapters represent my original work.

STATEMENT OF SUPPORT

This thesis was supported mainly through three sources. I am very grateful to have received scholarships from the Fonds de Recherche du Québec – Santé and a Fast Foundation award (previously known as the McGill Surgeon Scientist Program). These scholarships provided me with the necessary support without which this research would not have possible. I would also like to acknowledge the support from the Steinberg-Bernstein Centre for Minimally Invasive Surgery and Innovation, which is funded by an unrestricted educational grant from Covidien. Their support was also instrumental in the completion of the studies contained in this thesis.

The views presented in this thesis are that of the authors and do not necessarily reflect those of the Fonds de Recherche du Québec, the Fast Foundation, or Covidien.

LIST OF TABLES AND FIGURES

ABSTRACT

Enhanced recovery after surgery (ERAS) pathways are multidisciplinary clinical care pathways incorporating multiple evidence-based interventions designed to decrease the surgical stress response, enhance recovery, and improve outcomes. Multiple randomized trials have demonstrated the clinical effectiveness of ERAS over conventional care for elective colorectal surgery, but these pathways require significant resources to design, implement, and maintain. There is little economic evidence to support ERAS, as the existing data are low quality and there are large knowledge gaps regarding post-discharge outcomes and the socioeconomic impact of ERAS. Therefore the objective was to determine the cost-effectiveness of ERAS versus conventional care for patients undergoing elective colorectal surgery.

In order to adequately measure recovery, the postoperative recovery construct was conceptually defined as a multidimensional construct that followed an expected trajectory of immediate postoperative deterioration and then a gradual rehabilitation back to or surpassing preoperative baseline. This definition was used to validate the SF-6D, a multi-attribute utility instrument, as a measure of postoperative recovery and for use as the main outcome measure of the cost-effectiveness analysis. Superior validity evidence was also provided for the SF-6D over the EQ-5D, another utility instrument. A pilot study as performed to estimate the cost impact of ERAS for esophagectomy using deviation-based cost-modeling, a novel method to analyze costs and outcomes for clinical pathways. Results from this pilot study were then used for sample size calculations for the cost-effectiveness analysis comparing ERAS and conventional care for colorectal surgery.

The main study was a multi-institutional prospective cohort study that recruited adult patients undergoing elective colorectal surgery over a one-year period (10/2012 to 10/2013). One

centre utilized ERAS routinely and the other did not. Costs and outcomes were measured over a 60-day time horizon. A total of 190 patients (95 ERAS, 95 conventional care) participated. ERAS was associated with lower length of hospitalization, less productivity loss, less caregiver burden, and decreased outpatient resource utilization. ERAS was also associated with decreased costs from a societal perspective (mean difference -2985 CAN$, 95% CI -5753, -373), but no difference in quality-adjusted life (mean difference: +0.87 quality-adjusted days, 95% CI -1.23, 2.97) compared to conventional care. Uncertainty analysis reported that ERAS was highly probable (>98% at all willingness-to-pay thresholds) to be cost-effective. The base-case results were insensitive to multiple sensitivity scenarios and subgroup analyses.

In conclusion, evidence was provided to support the cost-effectiveness of ERAS over conventional care for patients undergoing elective colorectal surgery. In particular, the analysis addressed many of the limitations of previous economic evaluations and used a validated measure for postoperative recovery as the main outcome measure. Future research should focus on the costs and benefits of ERAS on a population level. High value cost-effective healthcare can be obtained through ERAS, as it lowers costs without compromising outcomes.

RÉSUMÉ

Les approches de soins péri-opératoires selon "enhanced recovery after surgery" (ERAS), qui sont basées sur de multiples études cliniques, ont pour but de diminuer la réponse au stress chirurgical, améliorer la convalescence et les résultats cliniques. Plusieurs études randomisées ont démontré l'efficacité clinique de ERAS par rapport aux soins péri-opératoires conventionnels pour la chirurgie colorectale élective. Cependant, des ressources considérables sont nécessaires pour la création, l'implantation et le maintien de ERAS. Il y a peu de d'évidence économique pour supporter ERAS, et les données existantes sont de pauvre qualité. Il y a surtout un manque de connaissance clair en ce qui a concerne le devenir des patients une fois sortis de l'hôpital et les impacts socio-économiques de ERAS. Donc, l'objectif de cette thèse est de déterminer les coûts-efficacités de ERAS versus les soins péri-opératoires conventionnels pour les patients ayant une chirurgie colorectale élective.

Afin de mesurer adéquatement la convalescence, la notion de convalescence post-opératoire a été définie comme un concept multidimentionnel qui suit la trajectoire attendue de détérioration post-opératoire immédiate suivie d'un retour graduel au niveau de base pré-opératoire, ou supérieur à ce dernier. Cette définition a été utilisée pour valider le SF-6D, un instrument à multiples attributs, mesurant la convalescence post-opératoire et qui représente le principal instrument d'utilité de l'analyse coût efficacité. Des évidences ont démontré la supériorité de SF-6D par rapport au EQ-5D, qui est un autre instrument d'utilité. Une étude pilote, réalisée pour estimer l'impact sur les coûts de ERAS pour l'oesophagectomie, utilise un "deviation-based-cost-modeling", une nouvelle méthode d'analyse des coûts et des résultats cliniques. Les résultats de cette étude pilote ont ensuite été utilisées pour le calcul

d'échantillonage de l'analyse coûts-efficacité comparant ERAS et les soins péri-opératoires conventionnels pour la chirurgie colorectale.

L'étude principale est une étude prospective multi-institutionnelle de cohorte qui a recruitée des patients adultes ayant une chirugie colorectale élective entre octobre 2012 et octobre 2013. Un centre a utilisé ERAS de façon routinière et l'autre non. Les coûts et les résultats ont été mesurés sur une période de 60 jours. Un total de 190 patients (95 ERAS, 95 soins conventionnels) y ont participé. ERAS est associé à une durée d'hospitalisation plus courte, moins de perte de productivité, moins de soins du personnel hospitalier et moins d'utilisation de ressources post congé d'hôpital. ERAS est également associé à une baisse des coûts d'un point de vue sociétaire (différence moyenne -2985$ CAN, 95% IC -5753, -373), mais pas de différence en terme de qualité de vie ajustée (différence moyenne: +0.87 qualité ajustée par jour, 95% CI -1.23, 2.97), par rapport aux soins conventionnels. Des analyses incertaines rapportent que ERAS est très probablement (>98% de seuil) coût-efficace. Les résultats de cette étude sont insensitifs aux multiples scénarios de sensitivité et d’analyses de sous-groupes.

En conclusion, nos évidences supportent l’avantage coût-efficacité de ERAS par rapport aux soins péri-opératoires conventionnels pour les patients ayant une chirurgie colorectale élective. Notamment, les analyses évaluent plusieurs limitations des évaluations économiques antérieures et utilisent une mesure de validité pour déterminer la convalescence post-opératoire. Des recherches futures devraient focusser sur les coûts et les bénéfices de ERAS sur un niveau populationnel. Les soins de santé de haute valeur peut être obtenu grâce à ERAS, comme ce dernier diminue les coûts sans compromettre les résultats.

1. INTRODUCTION

Despite advances in patient selection and operative techniques, major abdominal surgery is still associated with significant morbidity and mortality. Certain procedures, such as colorectal and upper gastrointestinal resections, are especially subject to a high incidence of postoperative complications. Amongst a large population sample of patients undergoing general surgery procedures in the United States between 2005 and 2006, elective colectomy, gastrectomy and esophagectomy were all associated with complication rates of 30% or higher.[1] Part of the pathophysiology behind the development of these complications is the considerable metabolic stress from surgery, which leads to increased demand on the patients' physiologic reserves and organ function.[2] Traditional surgical practices such as preoperative and postoperative fasting, use of high volume crystalloid fluid, opiate-based analgesia, and postoperative bedrest, further increase the stress response and decrease the patients' physiologic reserves.

Recent improvements in anaesthesia, pain control, and minimally invasive surgery have all contributed to improved postoperative outcomes, but the effect of each element is limited if performed in isolation.[3] In the mid-1990s, Dr. Henrik Kehlet began reporting the first results of a multimodal perioperative rehabilitation program incorporating multiple evidence-based interventions into a single clinical pathway aimed at decreasing the surgical stress response and accelerating recovery.[4-7] In these early reports, patients undergoing elective colonic resection were managed using a standardized postoperative care protocol with pain control using using opiod-sparing neural blockade and early enforced mobilization and enteral feeding, and were discharged after a median of two days without increased morbidity. Since then, these multimodal care pathways have been further refined to contain more than 20 elements encompassing all perioperative phases, and are now commonly referred to as enhanced recovery after surgery

(ERAS) pathways (also known as enhanced recovery pathways or fast-track surgery).[8] While specific elements of ERAS pathways may vary between procedures and practice settings, the principal elements of an enhanced recovery pathway includes preoperative patient education and preparation for surgery; attenuation of the surgical stress response, pain, and postoperative nausea and vomiting through anaesthetic, analgesic and surgical techniques; and aggressive early postoperative mobilization, enteral feeding, and avoidance or early removal of drains and tubes.[9]

Since the initial reports from Dr. Kehlet, multiple randomized clinical trials have compared ERAS to traditional perioperative management. Meta-analyses of these trials have demonstrated improved physiologic parameters, such as return of gut function and immunologic markers, shorter hospital stay and decreased complications with enhanced recovery pathways for colorectal surgery.[10,11] Additional randomized trials in other abdominal procedures such as liver and gastric resections have further reinforced the clinical effectiveness of this multimodal perioperative management strategy.[12,13] Despite these promising results, enhanced recovery pathways have yet to be widely adopted. A survey of 295 hospital surgery departments across North America and Europe reported that many enhanced recovery elements, such as avoidance of nasogastric tubes, and early enteral feeding and mobilization, were not commonly practiced.[14] As a result, hospital length of stay after colorectal procedures remained prolonged across all participating departments. There are many reasons for the resistance against the adoption of enhanced recovery pathways, including social, professional, and organizational barriers.[15] These pathways have been met with initial skepticism, as many pathway elements represent major departures from longstanding practices, such as avoidance of preoperative fasting and bowel preparation. Resistance to these pathways can also be for professional reasons, as new competencies are often required (thoracic epidural use, minimally invasive surgical techniques,

etc.), and control ceded (such as specifics of perioperative management that is traditionally decided by the surgeon).[9] Finally, major organizational issues need to be addressed, as support from administration is required in order to devote the necessary resources that are required to develop, implement, and maintain these pathways.[16]

Yet, as healthcare spending continues to grow[17], enhanced recovery pathways may offer enticing economic incentives for their adoption. As of 2012, more than 17.9% of the US gross domestic product is devoted to health care.[18] Furthermore, the rate of increase in health spending outstrips that of the growth in GDP.[17] In Canada, 10.9% of the GDP was spent on healthcare as of 2008, and this percentage has increased over the past few decades.[18] Clearly, there is a need for high-value, cost-conscious health care. Two steps can be taken, either decreasing or eliminating care that provides no benefit, or offering interventions that provide good value for their cost.[19] Enhanced recovery pathways potentially can fulfill both steps by eliminating surgical practices that may be harmful, such as perioperative starvation and prolonged postoperative bedrest, and replacing them with multiple evidence-based interventions within a single perioperative strategy that may reduce waste and variability, and improve outcomes.

It is commonly accepted that enhanced recovery pathways are associated with reduced costs by virtue of shorter duration of hospitalization and decreased complications[16], but these touted economic benefits have not been rigorously evaluated. In addition, there are several reasons for which the previous statement may not hold true. First, the hospital days saved through an enhanced recovery pathway are at the tail end of a hospital admission, which may not be resource intensive, as 40% of variable costs of a surgical admission occur within the first three days.[20] Second, postoperative complications are the main cost driver of surgical admissions[21], and randomized trials comparing enhanced recovery and traditional perioperative

management have not unequivocally demonstrated a decrease in the incidence of postoperative complications.[11] As well, enhanced recovery has no effect on the incidence of severe postoperative complications, that is, those that are most costly. Finally, these pathways may be associated with important design, implementation, and maintenance costs, which are not well described.[22] As well, many enhanced recovery elements may be associated with increased costs, such as esophageal doppler-guided fluid management.[23] In addition, much of the evidence comparing enhanced recovery to traditional perioperative management has focused on traditional audit measures such as length of stay and complications, and few studies have actually investigated whether these pathways actually 'enhance recovery' after surgery and improve health-related outcomes.[24] Prior to widespread implementation of these pathways, a formal economic evaluation assessing both the costs and benefits of these pathways is required in order to determine their cost-effectiveness. Therefore the main objective of this thesis is to determine the cost-effectiveness of enhanced recovery versus traditional perioperative management for patients undergoing elective colorectal surgery.

2. LITERATURE REVIEW

2.1 PREAMBLE

Enhanced recovery pathways have been advocated to improve outcomes and decrease costs. However, there are several reasons for which this may not be the case, which were described in the Chapter 1, including the fact that the hospital days that are avoided are not likely to be resource intensive, complications are the main cost driver of surgical admissions and enhanced recovery pathways have not been unequivocally shown to decrease complications, and the potentially significant costs of design, implementation, and maintenance of these pathways. Furthermore, many of the reviews that allude to the economic benefits of enhanced recovery pathways do not provide references to support this statement.[16,25] It is important to evaluate the current state of the economic evidence in order to assess knowledge gaps, and to establish whether further economic evaluations are necessary.

In this manuscript, all economic evaluations comparing enhanced recovery to traditional perioperative management specifically in patients undergoing elective colorectal surgery were systematically reviewed, with a particular focus on the methodology and quality of the economic evaluations. This manuscript was published in the *Annals of Surgery* (*Ann Surg* 2014; 259(4): 670-6).

2.2 A SYSTEMATIC REVIEW OF ECONOMIC EVALUATIONS OF ENHANCED RECOVERY PATHWAYS FOR COLORECTAL SURGERY

Lawrence Lee[1], Chao Li[1], Tara Landry[2], Eric Latimer[3,4], Francesco Carli[5], Gerald M. Fried[1], Liane S. Feldman[1]

[1]Steinberg-Bernstein Centre for Minimally-Invasive Surgery and Innovation, McGill University Health Centre, Montreal, QC
[2]Montreal General Hospital Medical Library, McGill University Health Centre, Montreal, QC
[3]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, QC
[4]Department of Psychiatry, Douglas Institute, Montreal, QC
[5]Deptartment of Anesthesia, McGill University Health Centre, Montreal, QC

Corresponding author and reprint requests
Liane S. Feldman, MD FACS FRCSC
Director, Division of General Surgery
Director, Steinberg-Bernstein Centre for Minimally-Invasive Surgery and Innovation
McGill University Health Centre
1650 Cedar Ave, L9.303
Montreal, QC
Canada H3G 1A4
Email: liane.feldman@mcgill.ca
Phone: 1 (514) 934-1934 ext. 44004
Fax: 1 (514) 934-8210

Conflicts of Interest and Source of Funding
Dr. Feldman is the principal investigator of an investigator initiated research grant from Ethicon Endosurgery for the economic analysis of fast track surgery. Drs. Lee Li are supported by a scholarship from the Quebec Research Fund for Health Sciences (FRSQ), and the McGill Surgeon Scientist Program. The Steinberg-Bernstein Centre for Minimally Invasive Surgery and Innovation is supported by an unrestricted educational grant from Covidien Canada. Drs. Latimer, Carli and Fried and Ms. Landry have no relevant conflicts of interests to disclose.

Running head
Economic evaluations for ERP

ABSTRACT

Objectives: To perform a systematic review of economic evaluations of enhanced recovery pathways (ERP) for colorectal surgery

Summary background data: While there is extensive literature investigating the clinical effectiveness of ERP, little is known regarding its cost-effectiveness.

Methods: A systematic literature search identified all relevant articles published between 1997-2012 that performed an economic evaluation of ERP for colorectal surgery. Studies were included only if their ERP included all five of the key components (patient information, preservation of GI function, minimization of organ dysfunction, active pain control, and promotion of patient autonomy). Quality assessment was performed using the Consensus on Health Economic Criteria (CHEC) instrument (scored 0-19; high quality ≥ 12). Incremental cost-effectiveness ratios (ICER) were calculated if sufficient data were provided, using difference in length of stay and overall complication rates as effectiveness measures.

Results: Out of a total of 263 unique records identified (253 from databases, 10 from other sources), ten studies met our inclusion criteria and were included for full qualitative synthesis. Overall quality was poor (mean quality 7.8). Eight reported lower costs for ERP. The majority (8/10) of studies were performed from an institutional perspective and therefore did not include costs related to changes in productivity and other indirect costs (e.g. caregiver burden). Five studies provided enough information to calculate ICERs, of which ERP was dominant (less costly and more effective), in all cases for reduction in length of stay; and dominant or potentially cost-effective in 4, and questionable (no difference in costs nor effectiveness) in 1 for reduction in overall complications.

Conclusions: The quality of the current evidence is limited but tends to support the cost effectiveness of ERP. There is need for well-designed trials to determine the cost-effectiveness of ERP from both the institutional and societal perspectives.

INTRODUCTION

In the era of rising health care costs, strategies that both improve clinical outcomes and decrease costs are increasingly attractive. Colonic operations are particularly prone to high morbidity[1], which contribute significantly to increased resource consumption and overall costs. Enhanced recovery after surgery pathways (ERP) are multidisciplinary care pathways that integrate multiple evidence-based interventions in all perioperative phases to decrease the surgical stress response, hasten recovery and ultimately improve outcomes.[26] Meta-analyses of initial randomized trials comparing ERP to traditional perioperative care in colorectal surgery have reported a decrease in overall hospital length of stay as well as a modest decrease in postoperative morbidity.[10,27] Despite these data, adoption of ERP has been slow, particularly in North America.[14]

In view of this resistance to change, it was thought that evidence of economic benefits from ERP may help promote adoption of this paradigm-shifting perioperative strategy.[14] While neither meta-analysis reported economic outcomes, both assumed that ERP would be less costly as a result of the shortened length of stay and decreased morbidity.[14] However there are several reasons why this assumption may not hold true. A reduction in length of stay may not have a significant impact on overall costs, as the hospital days that are avoided at the tail end of an inpatient admission are not likely to be resource intense.[20] More recent and larger randomized trials have failed to demonstrate a decrease in postoperative morbidity[28,29], which is one of the significant cost drivers in colorectal surgery[30]. In addition, ERP require significant time and money to implement and maintain.[31] In light of these factors, it is important that evidence is provided to demonstrate the cost-effectiveness of ERP compared to conventional perioperative care (CC). Therefore the objective of this study was to perform a systematic review and quality

assessment of all economic evaluations comparing enhanced recovery pathways and conventional perioperative care in patients undergoing elective colorectal surgery.

METHODS

Search Strategy

We performed a systematic literature search of all full-text articles published between January 1997 and August 2012. We decided to limit the search to articles published starting in 1997 as this was the year of the first report describing ERP pathways for colorectal surgery.[32] We searched Medline, Embase, HealthStar, PubMed, Web of Science, Scopus, Cochrane Library, CINAHL, National Health Service Economic Evaluation Database (NHS EED), Health Technology Assessment (HTA) database, and the Database of Abstracts of Reviews of Effects (DARE). Electronic search terms were [Colorectal Surgery/ or Digestive System Surgical Procedures/ or Colorectal Neoplasms/Surgery or Colon/Surgery or Rectum/Surgery or Colonic Disease/Surgery or Rectal Diseases/Surgery or Inflammatory Bowel Diseases/Surgery or colorectal surgery.tw or colorectal surgeries.tw] and [Critical Pathways/ or critical path*.tw or clinical path*.tw or ERAS.tw or enhanced recovery.tw or fast track*.tw] and [Costs and Cost Analysis/ or cost*.tw or costs.tw or economics.fs]. In addition, we searched the International Network of Agencies for Health Technology Assessment website (www.inahta.net), which includes publications from health technology assessment agencies from 21 countries, and trial registries (PROSPERO, ClinicalTrials.gov, Current Clinical Trials) for unpublished trials. In addition, the reference lists of key records were assessed for additional relevant studies. Recognized experts in the field were contacted to identify additional studies. Searches were restricted to adult patients only (18+ years). No language limits were applied. Studies were first

screened for relevance based on their title and abstract, and full-text articles were retained if they met the following criteria: if the main study population consisted of patients undergoing elective colorectal surgery that were managed by an ERP; if they used traditional perioperative management as the comparator, and if some form of economic evaluation comparing the two perioperative strategies was performed. Clinical pathways were considered to be an ERP if they documented all five key ERP components as described by Kehlet et al.[25] Two authors (LL, CL) independently assessed each record for eligibility and extracted the data, including quality assessment, from full-text articles that met the inclusion criteria. Disagreements were resolved by consensus, and if no agreement could be made, a third author (LSF) was consulted.

Data Extraction & Synthesis

Details recorded included study design and characteristics, country of origin, currency, patient sample characteristics, number of individual ERP elements (out of a possible 20 elements[25]), type of economic evaluation and methods (in particular if the appropriate statistical methods to compare costs were used), and whether it was a primary or secondary study objective. If a study did not report an incremental cost-effectiveness ratio (ICER) but provided sufficient data, then data were reanalyzed to generate ICERs. ICERs were calculated for differences in complications and mean length of stay. ICERs were interpreted according to the cost-effectiveness plane (Figure 1). We expected that significant study heterogeneity would prevent a valid quantitative synthesis; therefore no attempt was made to carry one out.

The quality of the economic evaluation was assessed using the Consensus on Health Economic Criteria (CHEC) instrument, which contains 19 yes-or-no questions for qualitative assessment of economic evaluations.[33] This instrument was designed for clinical trials and

observational studies, and does not include questions regarding modeling studies, unlike other commonly used quality-assessment tools for economic evaluations.[34,35] As per the CHEC authors, a "no" response was selected if not enough information was available in the article or in other published material to answer a question. Each of the 19-items was accorded one point if present, and summing across items generated a total quality score, with a higher score denoting better quality. Studies were considered "high-quality" if they scored at least 12 points (out of a possible 19).

RESULTS

Study Characteristics & Quality Assessment

Out of a total of 268 titles identified from our systematic search, 54 titles underwent full-text review, of which 10 articles met our inclusion criteria and were included for qualitative synthesis (Figure 2). Five other on-going trials that listed cost as an outcome measure were identified from trial registries, but at the time of review were still in the recruitment phase and therefore were not included.[36-40] The mean number of patients in the ERP group was 148.8 (range 20 – 588), and 167.5 (range 20 – 770) in the CC group. The mean number of ERP elements described was 10.8 (range 6 – 16). Five studies used a "before and after" study design, that is a prospective ERP cohort compared to historical controls. In two studies both groups were studied prospectively, and in one study both groups were studied retrospectively. Two randomized controlled trials were included (Table 1). Only the subgroup analysis of the study by Archibald et al.[41] was included in this review, as it compared patients enrolled to those not enrolled in an ERP, whereas the main analysis was not relevant to this review. Time horizon was

appropriate to include relevant outcomes (at least 30 days follow-up) in nine of ten studies, but was not reported in one publication (Table 1).

Only one study compared ERP to conventional care in a specific patient group (undergoing ileal-anal pouch anastomosis ± proctectomy).[42] In three studies, the patient population only included patients with colorectal malignancy[28,29,43], while in the remainder a varied case-mix of colorectal surgery was studied.[31,41,44-47] A mixture of laparoscopic and open cases were included in four studies[31,41,43,46], and four studies included open resections exclusively.[29,42,45,47] Vlug et al.[28] randomized patients to four possible groups: laparoscopy + ERP, open + ERP, laparoscopy + CC, and open + CC, whereas Bosio et al.[44] compared patients managed with ERP and laparoscopy to controls managed by CC and open resection. Four studies were performed in the United States, four in Europe, one in China and one in New Zealand (Table 1).

Overall quality was poor (Table 2), as assessed by the CHEC instrument (mean quality 7.8). Only 2 studies were considered high quality, and both originated from Europe and performed economic evaluation as a primary study objective.[43,45]

Methods of Economic Analysis

Five studies reported cost as a primary outcome.[31,42,43,45,46] Eight of the ten studies performed the economic evaluation from an institutional perspective, and therefore did not include costs related to convalescence, productivity loss, and informal caregiver burden (i.e. indirect costs) that are included in health care system or societal perspectives (Table 2). In the study by King and colleagues, the stated perspective was institutional (National Health Service), but their analysis included indirect costs, and therefore was considered societal.[43]

Unit costs (i.e. costs of resources consumed) were described in only two studies.[31,43] Only one study included the implementation and maintenance costs of ERP.[31] In four studies, total cost was presented as a single summary measure without explanation of which costs were included. In these studies, costs were calculated as per the "hospital billing system"[28,41,44] or were not specified[29]. For example, it is unclear in the studies by Bosio et al.[44] and Ren et al.[29] whether total medical costs included the cost of readmissions. Overhead costs (i.e. administrative, housekeeping, etc.) were not considered in five studies[28,42,44-46], likely leading to underestimation of the true costs, and unclear in four studies[29,31,43,47]. Only Archibald et al.[41] explicitly reported including overhead costs, however the method of allocation was not specified. In the two studies that included indirect costs[43,45] (referring to non-medical-related costs), productivity losses were calculated using the human capital approach, which assumes the economic value is equal to the cost of their salary and benefits. Informal caregiver burden was not assessed in any studies, which is important if earlier discharge from hospital shifts the burden of care.

In economics, the mean cost is the important summary measure to the decision-maker, as it is more reflective of the overall budget impact, compared to the median.[48] In this review, we identified two studies that reported median costs without reporting the arithmetic mean.[28,42] In the eight studies that reported mean costs, four studies used inappropriate statistical methods to compare costs, one study did not report statistical methods, and one study did not perform any statistical analysis (Table 2). Only King et al.[28] and Folkerson et al.[34] used accepted methods to assess uncertainty in economic evaluations.

Cost-Effectiveness Results

Eight of the ten studies reported a lower cost for ERP compared to CC (Table 2). All four of the studies originating from the United States reported significantly lower direct medical costs associated with ERP.[41,42,44,47] Kariv et al.[42] reported lower anesthesia, nursing, laboratory, and medical services, as well as overall direct medical costs, for ERP. The remaining three US studies did not report breakdowns by types of cost.[41,44,47]

Two of the four European studies did not demonstrate a difference between the two groups.[28,43,45,46] In the study by King et al.[43], there was a significant difference in indirect costs (productivity losses) between ERP and CC, although there was no overall difference when medical costs were included at 90 days. Vlug et al.[28] also demonstrated no difference in direct medical costs between any of the four groups that were included in the study, although there were large differences in magnitude between university and teaching hospitals, which were not explained. Folkerson et al.[45] reported lower direct, indirect, and overall costs for ERP in the base care scenario, as well on sensitivity analyses. In particular, they reported that the difference in productivity losses was almost as large in magnitude as the difference in medical costs. Jurowich et al.[46] analyzed the medical costs for the first five days of hospitalization in uncomplicated patients, and reported a significantly lower overall cost for ERP, although the differences were only found in the first two postoperative days.

Both studies originating from Australasia reported cost-savings for ERP. Sammour et al.[31] was the only study to incorporate the development and implementation costs of ERP into their economic evaluation. Total cost per patient was lower in the ERP group, however no statistical or sensitivity analyses were performed, making it difficult to draw conclusions from their results. In the study by Ren et al.[29], overall costs were lower in the ERP group, and this was mainly due to decreased costs in the postoperative period.

No studies performed an incremental analysis of costs and effectiveness. Sufficient data to calculate ICERs for length of stay and overall complications was available in five studies (Table 3). No studies measured quality-adjusted life years. None of the calculated ICERs were placed in the northwest quadrant (i.e. more expensive and less effective) of the cost-effectiveness plane (Figure 1). In all studies, ERP was associated with a decreased length of hospitalization compared to CC (Table 1). ERP was associated with decreased complications only in 2 of the 5 studies that reported this outcome. The ICER for difference in length of stay as the measure of effectiveness was dominant (less costly and more effective) in all five cases. The ICER using overall complications as the effectiveness measure was dominant in two, cost-effective (less costly and equally effective) in one and questionable (no difference in costs and effectiveness) in two cases.

DISCUSSION

Economic evaluations are an essential part of assessments of new health technologies, and are important for funding decisions made by hospital administrators, insurers, governments and policy developers. This is the first study to systematically review the existing literature and qualitatively synthesize all available evidence on the cost-effectiveness of ERP compared to CC in patients undergoing colorectal surgery.

We identified ten studies with varying levels of methodological quality, of which eight reported cost savings associated with ERP. US studies all reported cost savings, but quality was poor as assessed by the CHEC instrument. European studies were of the highest quality, and the results were more equivocal. This discrepancy may be partly explained by the different perspective from which the economic evaluation was performed, as well as the differences in

health care systems between Europe and the US. However, no studies assessed all relevant costs. Indirect costs associated with productivity loss were poorly assessed, and no studies investigated the impact of ERP on caregiver burden once the patient is discharged home. In addition, this review identified only one study that assessed implementation and maintenance costs required by ERP. Compliance with ERP outside of clinical trials is often poor[49], and improvements in adherence may require significant organizational changes and resources[50]. The economic impact of these processes is unknown. More research is required on these aspects of perioperative care and recovery.

Although no study performed an incremental analysis of costs and effectiveness, sufficient information was available in five studies to calculate an ICER. In all five cases, ERP was dominant over CC for hospital days avoided, and dominant or cost-effective in four of the five cases for complications avoided. However these ICERs must be interpreted with caution. The definition of complications differed between studies, and the severity of the complication was not taking into account, which may also alter the ICER. As well, there is a risk of double-counting in cases where the ICER is expressed as cost per event avoided (in this case, hospitalization days and complications), if the ICER includes the cost of the event in the numerator, as well as the number of events avoided in the denominator.[51] The more appropriate method would be to calculate ICERs using quality-adjusted life years, which incorporates multiple outcomes into a single measure and, as long as health states are valued without regard to productivity, avoids the possibility of double-counting. This is the approach recommended by the Panel on Cost-Effectiveness in Health and Medicine.[52]

However the results reported in this systematic review must be interpreted with caution. The methodology for the calculation of overall costs and unit costs was poorly described. This

information is important as it allows the reader to evaluate the generalizability of the results to their own practice setting. It is also important to report the types of costs that are included. This transparency is required so that the reader can evaluate whether all relevant costs are included. While this is a prevalent problem in published economic evaluations[53], the methodology used to calculate costs can significantly influence unit costs[54], and in this way the overall estimated cost. Published recommendations regarding the methods for deriving costs exist[52,55], which future economic evaluations should follow.

Furthermore, the quality of statistical analysis of costs was poor. Median costs were mostly reported, as the distribution of costs is almost always right-skewed[56], despite the arithmetic mean being the statistic of interest for health-care policy decisions. However, non-parametric statistical testing of medians (e.g. Mann-Whitney U-test) is not appropriate to make statistical inferences on the mean. The assessment of uncertainty, either through statistical or sensitivity analyses, is important when interpreting cost data, but was poorly performed.

The results from this systematic review must be interpreted in view of several limitations. Firstly, the CHEC instrument used for quality assessment may be subject to high inter-rater variability.[57] In this review, two reviewers independently performed the quality assessment with CHEC, and disagreements were resolved by consensus or with an independent third reviewer. While other quality assessment instruments exist[34,35], there is no definitive method on how to evaluate economic evaluations, although the specific instrument used may not affect the conclusions regarding quality[58]. Furthermore, we also did not assess for publication bias due to the small number of included studies. We did however perform a search of trial registries, which did not identify any studies that were registered and completed but never published. Finally, we did not perform any meta-analytic technique to quantitatively aggregate the cost results, but the

significant clinical and methodological heterogeneity of the included trials would have rendered any quantitative synthesis invalid.

In view of the limitations of the existing evidence, more research on the cost-effectiveness of ERPs is required. Future studies should be performed according to published guidelines[35,52,59] to ensure validity and generalizability. In particular, this review identified that transparency about which costs were included and how costs were calculated, and the use of appropriate effectiveness measures, was lacking across most of the studies. There were also few data on the development and maintenance costs of ERPs, which may be significant.[60] We also recommend that future analyses report results from different perspectives that may be important to decision-makers. Finally, studies in specific settings may be required in order to account for the differences in health care systems between countries, as this prevents the generalizability across practice settings. These methodological issues must be addressed in order to provide a convincing and sound economic argument for the adoption of ERPs.

In conclusion, our systematic review reports that the quality of the current evidence is limited but tends to support cost savings for ERP for patients undergoing elective colorectal surgery. However these results must be interpreted with caution, as there is significant study heterogeneity and limited generalizability across countries and institutions. Valuable information can be obtained from economic evaluations beyond the results of clinical effectiveness alone, and there is need for well-conceived trials to determine the cost-effectiveness of ERP from both the institutional and societal perspectives.

Table 1 – Study characteristics of included studies

Study ID	Country	Perspective	Design	ERP elements	F/U	Sample	Clinical effectiveness (LOS and complications)
Archibald et al.[41]	USA	Institutional	Prospective	9	30 days	ERP: 588 CC: 770	Mean LOS: CC 8.6 vs. ERP 4.6 days, $p<0.001$ Incidence of complications not reported
Bosio et al.[44]	USA	Institutional	Historical controls	9	NR	ERP: 20 CC: 20	Mean LOS: CC 8.3 vs. ERP 3.6 days, $p<0.001$ Overall complications: CC 45% vs. ERP 25%, no statistical comparison
Folkerson et al.[45]	Denmark	Societal	Historical controls	13	30 days	ERP: 80 CC: 80	Median LOS: CC 8 vs. ERP 2 days, $p<0.05$ Overall complications: CC 55% vs. ERP 25%, *p<0.05*
Jurowich et al.[46]	Germany	Institutional	Historical controls	10	5 days	ERP: 29 CC: 29	Analysis of first 5 days of hospitalization in uncomplicated patients
Kariv et al.[42]	USA	Institutional	Prospective	7	30 days	ERP: 83 CC: 83	Median LOS: CC 5 vs. ERP 5 days, $p=0.071$ No difference in overall complications (overall incidence not reported)
King et al.[43]	UK	Stated as NHS but societal	Historical controls	16	90 days	ERP: 60 CC: 86	Mean LOS: CC 12.9 vs. ERP 6.3 days, $p<0.001$ No difference in overall complications (overall incidence not reported)
Ren et al.[29]	China	Institutional	RCT	11	30 days	ERP: 299 CC: 298	Mean LOS: CC 6.6 vs. ERP 5.7 days, $p<0.001$ Overall complications: CC 9.4% vs. ERP 9.7%, $p=0.900$
Sammour et al.[31]	NZ	Institutional	Historical controls	13	30 days	ERP: 50 CC: 50	Median LOS: CC 8 vs. ERP 4 days, $p<0.001$ Overall complications: CC 66% vs. ERP 54%, $p=0.221$
Stephen et al.[47]	USA	Institutional	Retrospective	6	30 days	ERP: 86 CC: 52	Mean LOS: CC 6.9 vs. ERP 3.7 days, $p<0.001$ Overall complications: CC 25% vs. ERP 12%, $p=0.058$
Vlug et al.[28]	Holland	Institutional	RCT	14	30 days	ERP: 193 (Lap 100; Open 93) CC: 207 (Lap 109; Open 98)	Median LOS: CC-Lap 6 vs. CC-Open 7 vs. ERP-Lap 5 vs. ERP-Open 7 days, $p<0.001$ Overall complications: CC-Lap 34% vs. CC-Open 41% vs. ERP-Lap 34% vs. ERP-Open 46% days, $p=0.20$

ERP = enhanced recovery pathway; F/U = follow-up; LOS = length of stay; CC = conventional care; NR = not reported; NHS = National Health Service, RCT = randomized controlled trial; NZ = New Zealand; Lap = laparoscopic approach

Table 2 – Cost data reported in included studies

Study ID	Quality (/19)	Costs	ERP	CC	*p*/95% CI	Statistical method
Archibald et al.[41]	6	Total hospital costs (direct + overhead)	US$ 11662*	US$ 21037*	$p<0.0001$	"two-sample hypothesis tests"
Bosio et al.[44]	0	Hospital direct costs	US$ 4993*	US$ 11383*	$p<0.001$	Not reported
Folkerson et al.[45]	16	Direct medical costs Indirect costs Total costs	DKK 17521 DKK 18649 DKK 36170	DKK 21340 DKK 24134 DKK 45474	N/A	Sensitivity analysis (results not shown in this table)
Jurowich et al.[46]	4	Hospital direct costs for the first 5 postoperative days	€ 1628	€ 2391	$p=0.001$	*t*-test for independent samples
Kariv et al.[42]	8	Direct hospital costs	US$5 692†	US$ 6672†	$p=0.001$	Wilcoxon's signed rank test
King et al.[43,+]	15	Total costs Indirect costs	£ 6,545.29* £ 534.39*	£ 7,216.00* £ 1,061.50*	95% CI: -1033.89 to 2433.53 95% CI: 54.00 to 986.67	Bootstrap estimates (10,000 iterations) with CIs taken at 2.5% and 97.5% percentiles
Ren et al.[29]	4	Total costs of the procedure Postoperative costs	CNY 15997* CNY 3594*	CNY 17763* CNY 5268*	$p<0.001$ $p<0.001$	"independent-sample *t*-test"
Sammour et al.[31]	8	Total hospital costs (incl. protocol development and research fellow's salary)	NZ$ 16052*	NZ$ 22939*	Not reported	None performed
Stephen et al.[47]	8	Total hospital costs (excl. surgeon's fees)	US$ 7070*	US$ 9310*	$p=0.002$	1-tailed *t* test
Vlug et al.[28]	9	Direct hospital costs (university hospitals) Direct hospital costs (teaching hospitals)	€ 10594 (lap)† € 12805 (open)† € 5768 (lap)† € 5497 (open)†	€ 11967 (lap)† € 10479 (open)† € 6228 (lap)† € 5650 (open)†	$p=0.56$ $p=0.41$	Kruskal-Wallis/ Mann-Whitney U tests

*Mean cost

†Median cost

+In this study, the confidence intervals in the published manuscript were erroneous, and the author was contacted to provide the corrected confidence intervals.

US$ = US dollars; DKK = Danish Krone; CNY = Chinese Yuan Renminbi; NZ$ = New Zealand dollars

1 DKK = 0.1573 US$; 1 € = 1.2647 US_{2010}, 1.3449 US_{2011}; 1 £ = 1.7460 US$; 1 CNY = 0.1506 USD; 1 NZD = 0.7260 USD; currency exchange rates at date of publication from www.xe.com

Table 3 – Incremental cost-effectiveness ratio analysis of studies with sufficient data

Study ID	Incremental Cost[a] $Cost_{ERP} - Cost_{CC}$	Incremental Effectiveness $E_{CC} - E_{ERP}$	Incremental Cost-Effectiveness Ratio
Length of stay			
Archibald et al.[41]	US$ -9374	4.0 days	ERP dominant
Bosio et al.[44]	US$ -6390	4.7 days	ERP dominant
King et al.[43]	No difference	6.6 days	ERP potentially cost-effective or dominant
Ren et al.[29]	CNY -1776 (US$ -279.75)	0.9 days	ERP dominant
Stephen et al.[47]	US$ -2240	3.2 days	ERP dominant
Overall complications			
Bosio et al.[44]	US$ -6390	20%	ERP dominant
Folkerson et al.[45]	DKK -9304 (US$ -10916.47)	30%	ERP dominant
King et al.[43]	No difference	No difference	Questionable cost-effectiveness
Sammour et al.[31]	NZ$ -6877 (US$ 4992.55)	No difference	ERP potentially cost-effective
Stephen et al.[47]	US$ -2240	No difference	ERP potentially cost-effective

[a]Values in parentheses represent the US$ equivalent at the time of publication; currency exchange rates from www.xe.com

[D]Dominant = less costly and more effective; cost-effective = less costly with no difference in effectiveness; questionable cost-effectiveness = no differences in cost and effectiveness

US$ = US dollars; CNY = Chinese Yuan Renminbi; DKK = Danish Krone; NZ$ = New Zealand dollars

Figure 1 – Cost-effectiveness plane

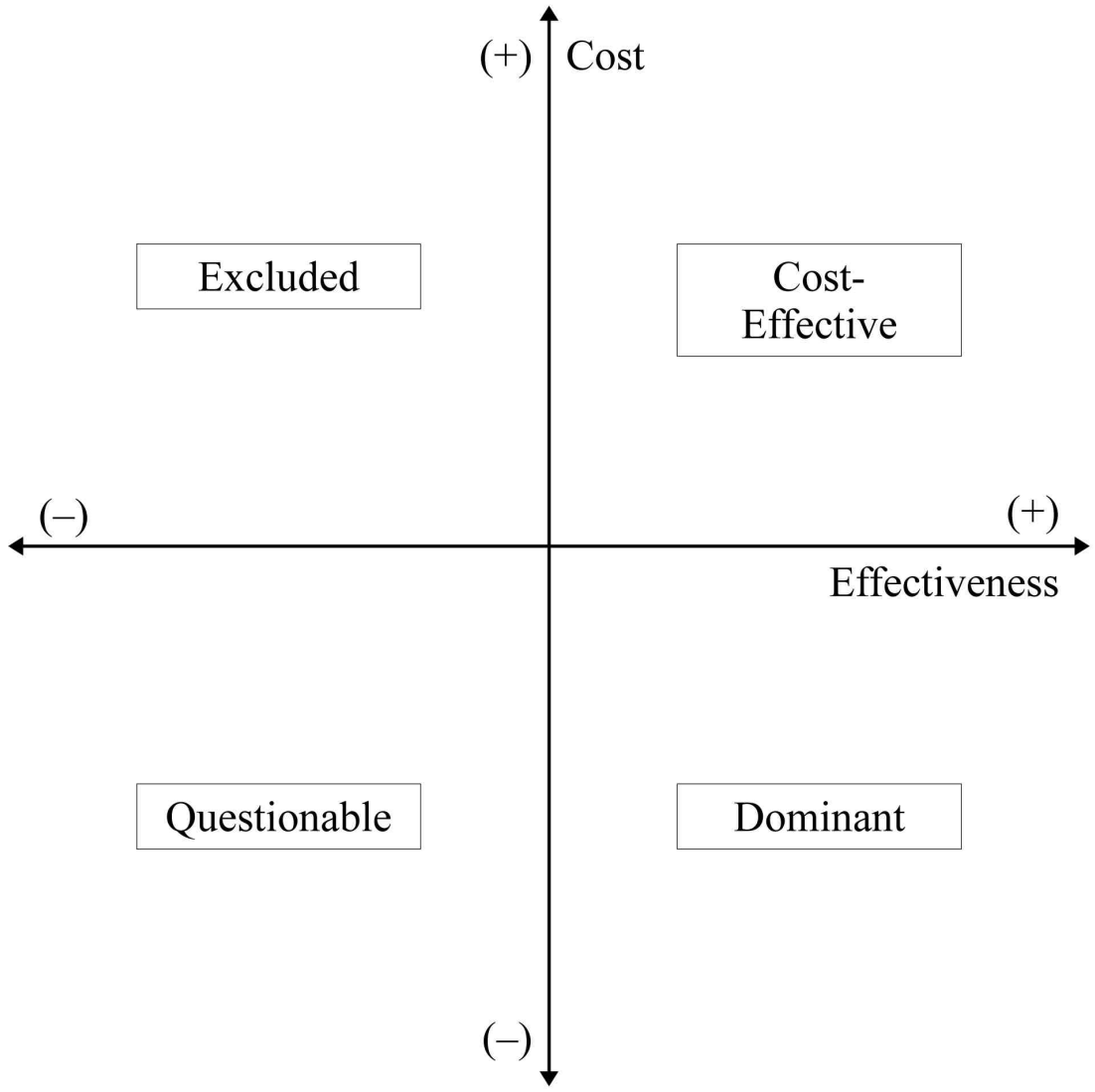
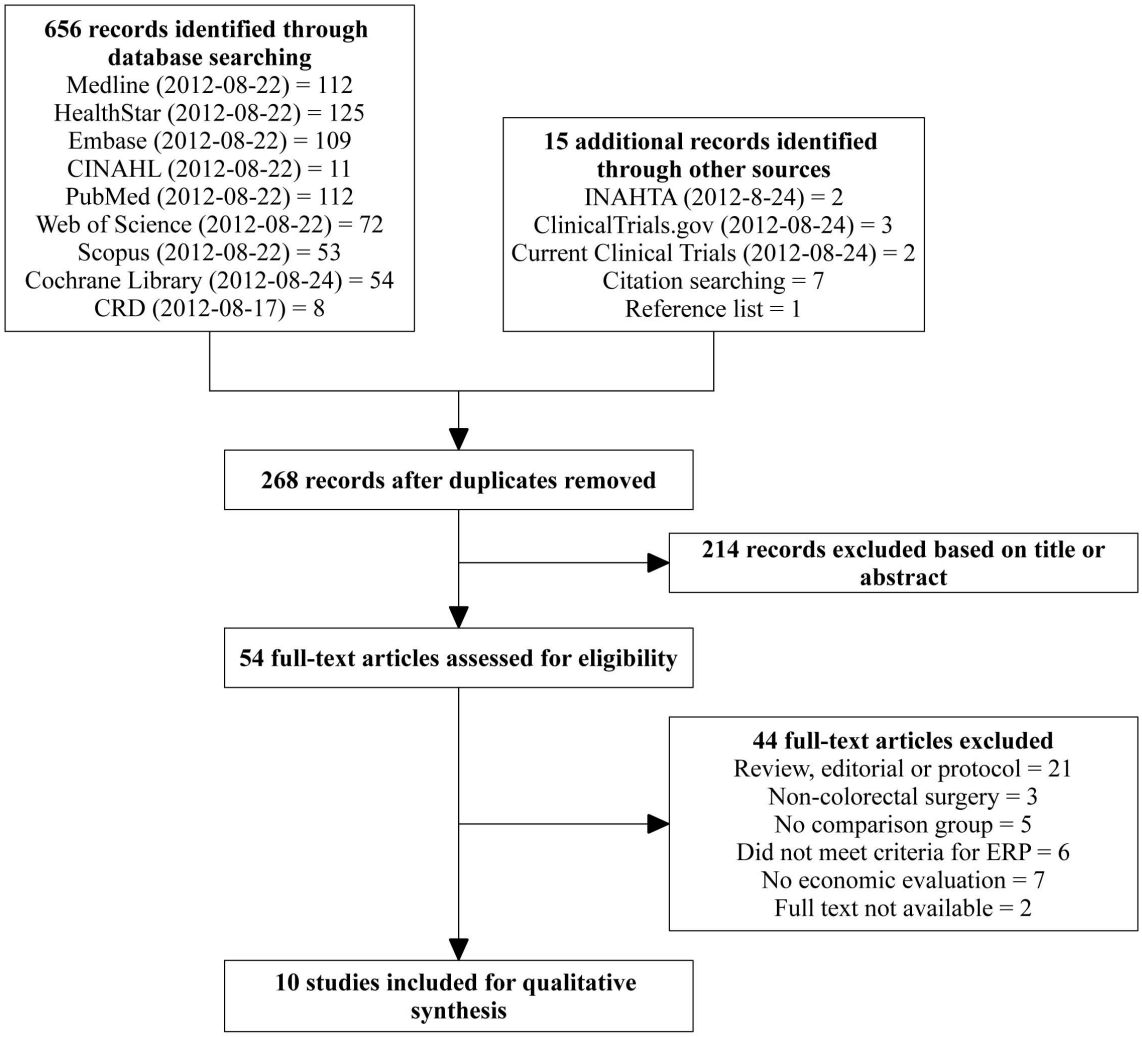


Figure 2 – Study identification and selection flowchart



3. DEFINING 'SURGICAL RECOVERY'

3.1 PREAMBLE

The review of the existing economic literature on enhanced recovery pathways for colorectal surgery identified major evidence gaps. To date, no studies have performed a formal cost-effectiveness analysis comparing enhanced recovery and conventional perioperative management strategies.[22] In addition, there is little evidence on post-discharge functional outcomes after enhanced recovery management. Much of the existing evidence has focused on traditional outcomes, such as length of stay and complications, or short-term biologic variables or symptoms, such as pain, return of gastrointestinal function, immunologic parameters, etc.[24] These outcomes are unlikely to fully capture the complex construct that is postoperative recovery. Therefore it is unclear as to whether enhanced recovery pathways are actually 'enhancing recovery', as recovery is not well measured.

Further compounding the lack of valid measures is the fact that there is no accepted definition of postoperative recovery. It is a complex construct – it may have different meanings for different stakeholders, and it involves different timeframes. Prior to identifying or developing valid measures of postoperative recovery – in particular QALY measures that accurately capture the recovery process, the construct of postoperative recovery must first be clearly defined. The following manuscript is a short research review intended to provide a consistent definition of the complex construct that is recovery after surgery. This manuscript was published in *Surgery* (Surgery 2013; epub 2013 Oct 12).

3.2 WHAT DOES IT REALLY MEAN TO "RECOVER" FROM AN OPERATION?

[1]Lawrence Lee MD MSc, [1]Tung Tran MD MSc, [2]Nancy E. Mayo PhD, [3]Franco Carli MD, [1]Liane S. Feldman MD

[1]Steinberg-Bernstein Centre for Minimally Invasive Surgery and Innovation, McGill University Health Centre, Montreal, Canada
[2]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Canada
[3]Department of Anesthesia, McGill University Health Centre, Montreal, Canada

Corresponding author
Liane Feldman, MD FACS FRCSC
Professor of Surgery
Director, Division of General Surgery
Director, Steinberg-Bernstein Centre for Minimally Invasive Surgery
1650 Cedar Ave, L9-303
Montreal, QC
Canada H3G 1A4
Tel: (514) 934-1934 ext. 44004
Fax: (514) 934-8210
Email: liane.feldman@mcgill.ca

Presentation Information
None

Conflicts of interest and source of funding
The Steinberg-Bernstein Centre for Minimally Invasive Surgery and Innovation is supported by an unrestricted educational grant from Covidien. LL is supported by the Quebec Health Sciences Research Fund (FRQS) and the McGill Surgeon Scientist Program. Neither funding agency was involved in the study design, collection, analysis, interpretation of the data and in the writing of the report.

INTRODUCTION

Many new surgical techniques and innovations are hypothesized to improve recovery. However, much of the effectiveness data of these innovations have focused on audit measures such as length of stay, morbidity and mortality, or biologic or physiologic parameters such as pain, return to gut function, and fatigue. Audit outcomes are, at best, proxy measures of recovery, as length of stay may be affected by external elements such socioeconomic, cultural and institutional factors[61], and complications and mortality are relatively uncommon and often inconsistently measured[62]. These measures are of greatest interest to clinicians, but patients, i.e. those who are actually 'recovering', equate recovery to the absence of symptoms and the return of their ability to perform activities as they could prior to surgery.[63] Biologic and physiologic outcomes, even though considered patient-reported, are difficult to interpret as measures of recovery as they are unlikely to persist beyond the short-term, or are often confounded between disease-specific symptoms and symptoms related to surgery and its potential complications.

Therefore, there is a need for a shift in the emphasis of outcome reporting from these audit measures to more patient- and recovery-centric measures. However there is no consistent definition for postoperative recovery. Recovery may have different meanings for different stakeholders such as administrators, doctors, nurses, and patients. This lack of a consistent definition is further complicated by the fact that postoperative recovery is a complex construct that encompasses multiple domains and timeframes. In order to be able to have reliable and valid measures of postoperative recovery, the construct of recovery must first be well defined. Therefore the purpose of this short research review is to introduce the reader to the concepts that are important to the construct of postoperative recovery, and identify areas where future research should be focused.

WHAT DOES POSTOPERATIVE RECOVERY MEAN?

Postoperative recovery is a complex and multidimensional process that involves multiple domains including physical, physiological, psychological, social, and economic aspects. A comprehensive definition of recovery after surgery has been described by Allvin et al[64], who identified the five defining attributes of recovery after surgery as: 1) an energy-requiring process; 2) a return to a state of normality and wholeness defined by comparative standards; 3) regaining control over physical, psychological, social and habitual functions; 4) returning to preoperative levels of independency/dependency in activities of daily living; and 5) regaining one's optimum level of well-being.[64]

These definitions are not new. In 1958, Dr. Francis D. Moore, a giant of 20th century surgery, wrote that convalescence, or recovery, includes "all the interlocking physical, chemical, metabolic, and psychological factors commencing with injury, or even slightly before the injury, and terminating only when the individual has return to normal physical well-being, social and economic usefulness, and psychological habitus."[65] He also wrote that "since convalescence must be said to terminate somewhere, we have chosen the criteria of the social and economic rehabilitation of an individual, that is, that he is psychologically and physiologically restored to full effectiveness."

These definitions emphasize the multidimensional aspect of recovery. Assessment of any one dimension while ignoring the remainder will not fully capture the whole construct of recovery. For example, consider a physically active patient who undergoes an uncomplicated elective colectomy for cancer. At the three-week post-operative visit, the patient reports no major physical symptoms, but is unable to resume normal sporting activities or work due to fatigue,

which negatively affects the patient's psychological, social, and economic domains. In this case, focusing only on the physical domain and ignoring the other domains will incorrectly describe this patient as 'recovered from surgery'. The natural trajectory of recovery is also implicit in these definitions, and can be described a rapid decline in functioning in all relevant domains immediately post-surgery and persistence in this postoperative state during the deterioration period, which will gradually 'recover' or exceed the baseline value over the rehabilitation period (Figure 1).

In truth, there is no single definition of recovery, nor does there need to be. There are overlapping phases of recovery that are of interest to different stakeholders, and subsequently the outcomes of relevance may vary depending on the phase. It is important that researchers report the timeframe, or phase of recovery, of interest. Table 1 provides a division of recovery into three distinct phases: early, intermediate, and late; each with its relevant outcomes of interests along with examples of validated generic instruments. For example, anesthesiologists often refer to recovery as the time required for patients to sufficiently recover from anesthesia enabling discharge from the PACU to the surgical ward (early phase). Outcomes of interest during this phase of recovery are generally focused on biologic or physiologic processes. The intermediate phase of recovery occurs from the time after transfer from the PACU to the surgical ward until discharge home. Traditionally this phase has been the most relevant to clinicians. Outcomes of interest in this phase tend to be concentrated on symptoms and impairments in the ability to perform activities of daily living, and audit measures such as length of stay and morbidity. There is also a growing body of evidence investigating novel metabolic markers of recovery within the intermediate phase, such as insulin resistance[66], immuno-modulators[67] and other neuroendocrine markers[68], amongst others.

Finally, the late phase of recovery occurs from the time the patient is discharged from hospital until the resumption of usual function or activities. In this phase, the relevant outcomes for recovery include functional status and health-related quality of life (QOL). The late phase is often longer than expected by clinicians. Lawrence et al[69] studied 372 patients aged 60 years and older after major elective abdominal surgery and measured physical ability, functional capacity, and cognitive function. At 6 months after surgery, fewer than 50% of patients had recovered to baseline levels of physical performance (handgrip strength, timed walk), and even more surprisingly, fewer than 20% of patients were able to perform the same activities of daily living as they had before surgery. Similarly, Mayo et al.[70] reported that less than 60% of patients, in whom the mean age was 60 years, had returned to baseline walking capacity at three months after elective colorectal surgery. Even after ambulatory laparoscopic cholecystectomy, more than 50% of patients had not yet reached baseline levels of physical activity by one-month after surgery.[71]

REFINING THE DEFINITON OF A RECOVERY MEASURE

Carli and Mayo[72] developed a causal pathway to evaluate the appropriateness of measures of surgical outcomes (Figure 2). In this model, any short- or long-term outcome measure must be biologically related to the intervention and should not be influenced by external factors. These outcomes must also be related to the short-term changes that occur after surgery. We have adopted this causal pathway in order to develop a conceptual model for the construct of postoperative recovery. In addition to the obligatory relationships with the intervention and the short-term postoperative changes, any recovery outcome measure must also assess the domains relevant to recovery, namely those elucidated by Moore[65] (physical, psychological, social, and

economic), and must correlate with the ability to perform activities of daily living[64]. Furthermore, this measure must be comparative to a baseline or population norm[64], and follow the expected trajectory of recovery, that is, a rapid decline from baseline ("deterioration"), followed by gradual improvement back to baseline or beyond ("rehabilitation"). The time frame has also been modified to reflect the phases of recovery described previously (Table 1). Early recovery from anesthesia allowing transfer out of the recovery room is best measured through biologic and physiologic parameters. The intermediate phase that occurs prior to discharge from hospital is best described with symptoms, such as gastrointestinal function, pain, and nausea, as well as mobility and the ability to perform activities of daily living, as these are criteria that assess the ability to be safely discharged. Long term recovery, or the late phase that occurs in the weeks and months after discharge from hospital, is best estimated with measures of functional status and health-related QOL, as these have been shown to remain significantly impaired in the postoperative period and take the longest to recover.[69] In addition, functional capacity measures, such as the six-minute walk, shuttle, and timed up and go tests, correlate well with the ability to perform activities of daily living, physical and mental health-related QOL, and the ability to perform activities of daily living[73]. The six-minute walk test has also been specifically validated in the context of postoperative recovery[74]. In addition to measures of functional capacity, physical activity can be estimated through validated questionnaires.[75]

The ideal time point at which to perform the assessment is also clearly dependent on the research question, and the course of the disease or intervention under investigation. For example, QOL after cholecystectomy remained poor at 1 month after surgery[71], but was improved compared to baseline by 3 months[76]. Timing is also dependent on the type of measure, as Lawrence et al[69] reported that after major abdominal surgery in the elderly, functional and

physical capacity measures remained significantly below baseline at 6 weeks, but by 6 months had mostly recovered back to baseline values, while QOL had mostly improved by 6 weeks.

Currently, there is no single instrument that evaluates both functional capacity and all of the relevant domains of postoperative recovery. Kluivers et al performed a systematic review on existing recovery specific instruments and found that none of the 12 identified instruments was fully validated for the construct of postoperative recovery.[77] For example, QOL instruments have failed to demonstrate a difference in both short- and long-term QOL after laparoscopic compared to open colectomy[78], despite the assumption of "faster recovery" for the laparoscopic approach. These results suggest that either laparoscopy confers no benefits, or perhaps the QOL instruments that were used do not adequately measure the construct of recovery. In addition, QOL is affected by external factors[79], and may also experience response shift[80], which makes their interpretation difficult.

It is therefore vitally important to use instruments that have validity evidence to measure the construct of surgical recovery. Several instruments, such as the Surgical Recovery Score[81] and the Abdominal Surgery Impact Score[82] have been specifically developed in this context, but have preliminary validity evidence. It is unknown whether these instruments are sensitive enough to detect subtle differences, such as between laparoscopy versus open, or traditional versus enhanced recovery perioperative management.

Until a single comprehensive instrument is developed, the best approach may be to use a combination of complementary instruments to account for all of the suggested criteria. In a study investigating recovery after laparoscopic donor nephrectomy, Bergman et al[83] used a combination of the six-minute walk test, which measured functional exercise capacity, visual analog scales for symptoms such as pain, and the SF-36, which measured mental and physical

health-related QOL, to objectively describe the recovery profile.

This stopgap approach has its limitations as few instruments have been specifically validated in the context of postoperative recovery. This approach also runs the risk of over-burdening patients and may decrease compliance to prospective study protocols, especially if multiple instruments that are used overlap and repeatedly measure similar aspects or domains. Furthermore, the ability to design studies to investigate interventions that are hypothesized to improve recovery is limited if the clinically relevant changes for each instrument are not known.

IMPLICATIONS FOR FUTURE RESEARCH

Therefore, future research on postoperative recovery should first focus on identifying all instruments that are currently used to measure recovery and determine their validity for the underlying recovery construct within specific surgical patient populations. While generic instruments such as the SF-36 have been validated across a wide spectrum of diseases, its psychometric properties have yet to be investigated for many specific surgical populations. Yet, it continues to be one of the most commonly used instruments, despite the fact that it may not be sensitive enough to detect changes between surgical patients (for example, between laparoscopic and open colectomy).[84] It is essential to determine whether these instruments are specifically validated for the patient population and setting in which they are used, as often validity information based on patients with other diagnoses are juxtaposed onto the new setting under study.[85] It is also important to determine the clinically relevant changes for these instruments so that future studies may be adequately powered to detect meaningful changes. Clinically relevant change refers to the minimal change in a measure that is considered meaningful, which can be from the point-of-view of the patient or related to another outcome. Sample size calculations for

"hard outcomes" in randomized studies are heavily scrutinized to ensure that studies are adequately powered to detect a relevant change. However, there are no data that report the relevant changes for existing recovery measures. Elucidation of the clinically relevant changes in recovery measures will provide data to perform the adequate power calculations for studies investigating interventions hypothesized to improve recovery.

For instances in which no valid measure of recovery exists, a valid patient-reported measure that satisfies the definition and trajectory and takes into account multiple stakeholders should be developed. An ideal measure of recovery needs phase-specific (Table 1), multidimensional, responsive to the expected trajectory of recovery (Figure 1), and be able to discriminate between other important outcomes of interest (for example patients with and without complications). We recommend the use modern psychometric methods, such as item-response theory or Rasch measurement theory, to develop, calibrate, and validate an item bank from existing instruments that capture the key health aspects of recovery.[86] Traditional psychometric methods of instrument development often result in a collection of items that are scored on an ordinal scale (for example a 5-point Likert scale: 1 "strongly disagree" to 5 "strongly agree"), that are weighted to form a total score. However this approach has several limitations in that the assumption of "equal differences" between ordinal levels may not hold true. For example, consider the example of another ordinal scale such as cancer staging (graded from I to IV) – one does not assume that the difference between stages I and II is the same as the difference between stages III and IV, yet this assumption, which is mathematically incorrect, is made for many of these instruments.[87] Also, administration of the entire instrument is required as these instruments are based around a total score, which, given the length of many of these instruments, may limit their practicality.[88] On the other hand, both item-response and Rasch

measurement theories estimate the degree to which items, related to an underlying construct, hierarchically fit on a unidimensional, linear continuum (in this case, the trajectory of recovery). Therefore, patients may be situated along a calibrated linear continuum using fewer items, thereby improving validity and ease of administration. However, this approach still requires that the resulting instrument be validated for its intended population and setting.

Another potential roadblock is the confounding between surgery-related and disease-related changes. For example, a visual analog scale to assess pain may not be an entirely useful instrument to measure recovery after surgery to address pain symptoms, as how does one specifically measure improvement in pain due to the surgery and differentiate it from the pain from the surgery itself? Also, recovery after oncologic surgery is further complicated by potential changes due to adjuvant therapy. Therefore, we recommend that initial validation of a measure of recovery be performed in a population of asymptomatic or healthy patients undergoing elective surgery, such as laparoscopic donor nephrectomy, or colonic resection of asymptomatic polyps found on routine colonoscopic screening. With this method, there should be minimal confounding between surgery-related and disease-related changes given the asymptomatic baseline. Subsequent validation, along with the determination of the specific clinically relevant change, in other patient cohorts should then be performed.

Finally, we also recommend that future studies be specific as to which part of the continuum of recovery is under study (Table 1). The exact measures of interest will differ depending on the type of surgery and population under study, but this framework may improve comparability between studies if the timeline of recovery can be standardized.

In summary, recovery after surgery is an important outcome that is inconsistently measured due to the lack of a clear definition. We have divided recovery into three distinct

phases along with their relevant outcomes of interests. Furthermore, we have argued that the recovery construct is patient-centric model that is multi-dimensional, and must include the physical, psychological, social and economic domains; has a comparative standard (either through baseline or population norms); and conforms to the expected trajectory of immediate deterioration, followed by rehabilitation. Any instrument used to measure postoperative recovery must also be related to the intervention or disease process and the immediate postoperative changes, and should be evaluated at time points relevant to the disease or intervention in question. Finally, we have proposed a research agenda to guide future efforts in this field. An instrument that is fully validated for this construct will be of immense utility as an outcome measure after surgery.

Table 1 – Stages of recovery

Phase of recovery	**Definition**	**Time frame**	**Threshold**	**Outcomes**	**Examples of existing instruments**
Early	From OR to discharge from PACU	Hours	Safety (sufficiently recovered form anesthesia and safe to go to floor	Physiologic and biologic	Aldrete Postanesthetic Recovery Score[89]
Intermediate	From PACU to discharge from hospital	Days	Self-care (able to care for self at home)	Symptoms and impairment in ADL	Quality of Recovery score[90] Abdominal Surgery Impact scale[82]
Late	From hospital discharge to return to usual function and activities	Weeks to months	Return to normal (baseline or population norms)	Function and health-related quality of life	Six-minute walk test[74] Community Health Activities Model Program for Seniors (CHAMPS)[71] SF-6D[91]

ADL = instrumental activities of daily living

Figure 1 – Expected trajectory of recovery. The dotted line represents the minimum level of functioning.

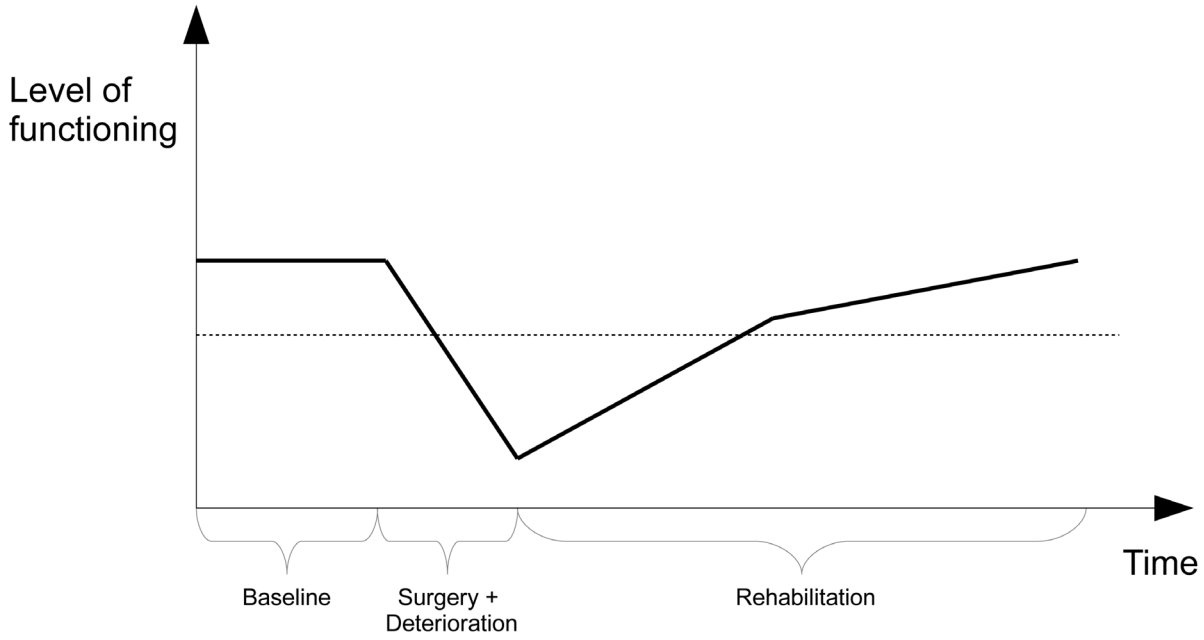
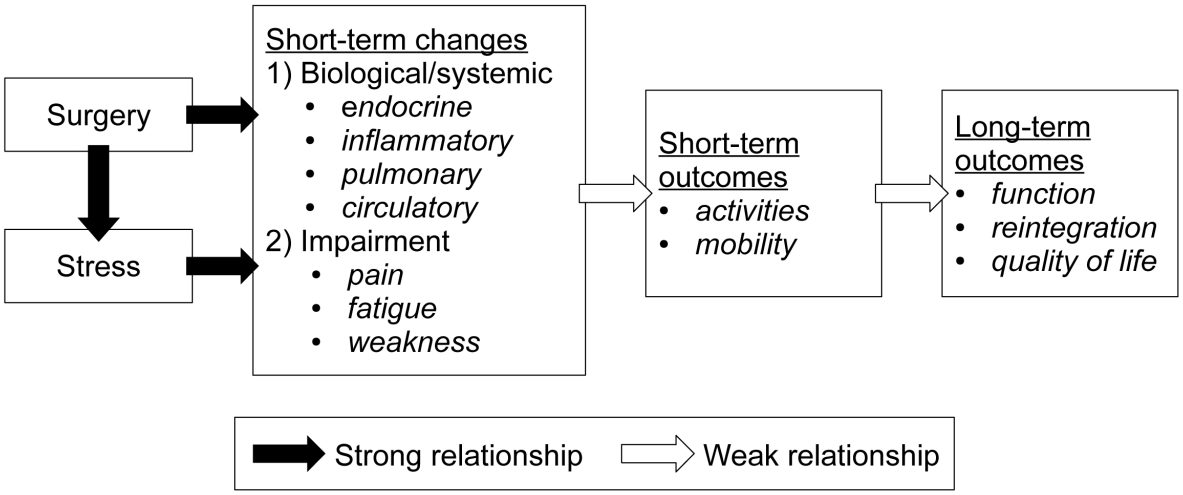


Figure 2 – Causal model for measuring outcomes after surgery proposed by Carli and Mayo[72]



4. THESIS OBJECTIVES

Chapter 2 reviewed the existing economic evidence supporting ERAS and found that the data are limited. In particular, economic analyses are methodologically inadequate, and there were no studies that have investigated the cost-effectiveness of ERAS in the Canadian setting. Given the significant differences between healthcare systems and healthcare costs between Canada, Europe, and the US, a formal economic evaluation should be performed to determine the cost-effectiveness prior to widespread implementation throughout healthcare institutions in Canada.

In order to perform a cost-effectiveness analysis, as outcomes measure that accurately captures the recovery process should be used. Chapter 3 provided a patient-centric definition of recovery. QALY measures appear ideally suited to measure the recovery process as multi-attribute utility instruments can incorporate multiple different domains into a single measure. However, none of these instruments have been validated in the context of postoperative recovery. Given these evidence gaps, the objectives of this thesis are:

1. Identify valid utility instruments for recovery after colorectal surgery so that quality-adjusted life years can be adequately measured; and,
2. To determine the cost-effectiveness of enhanced recovery versus conventional perioperative management for patients undergoing elective colorectal resection.

5. VALIDATION OF THE SF-6D HEALTH STATE UTILITY

5.1 PREAMBLE

Chapter 3 defined postoperative recovery as a multidimensional patient-centric model that follows an expected trajectory of immediate deterioration after surgery, followed by gradual rehabilitation. As well, postoperative recovery can be divided into distinct phases: early, intermediate, and late, each with its own relevant outcomes. It is within this framework that measures must be evaluated in order to be considered valid in the context of postoperative recovery.

Currently there is no gold standard measure of postoperative recovery. Previous studies have employed a strategy combining multiple different instruments that have been validated for a specific dimension relevant to the overall postoperative recovery construct.[92] For instance, Bergman and colleagues utilized the six-minute walk test, a physical performance measure, as well as health-related quality of life, to evaluate recovery after laparoscopic donor nephrectomy.[83] Similarly, Lawrence and colleagues utilized a variety of physical performance, functional capacity, health-related quality of life, and cognitive measures to assess recovery in elderly patients undergoing major abdominal surgery.[69] However this approach runs the risk of overburdening the patient with repetitive measures, and it is also it is difficult to perform a cost-effectiveness analysis if multiple different outcome measures are used.

Multi-attribute preference-based measures of health, such as the SF-6D[93], may be ideal outcome measures for cost-effectiveness analyses of surgical interventions advocated to improve recovery. The multidimensional nature of postoperative recovery may be reflected in the multiple domains that are covered by these instruments, and importantly, each dimension is preference-weighted to provide a single summary measure on the utility scale (0 to 1). However,

the validity of these instruments for postoperative recovery is unknown. The following manuscript investigates the validity of the SF-6D health state utility as a measure of postoperative recovery after colorectal surgery. The responsiveness of the SF-6D to the expected postoperative changes was assessed, and different construct validity hypotheses were tested. This manuscript was published in the *Journal of Surgical Research* (J Surg Res 2013; 184(1): 108-14).

5.2 VALUING POSTOPERATIVE RECOVERY: VALIDATION OF THE SF-6D HEALTH STATE UTILITY

[1]Lawrence Lee MD, [1]Nathaniel Elfassy, [1]Chao Li MD, [2]Eric Latimer PhD, [1]A. Sender Liberman MD, [1]Patrick Charlebois MD, [1]Barry Stein MD, [3]Franco Carli MD, [1]Gerald M. Fried MD, [1]Liane S. Feldman MD

[1]*Steinberg-Bernstein Centre for Minimally Invasive Surgery and Innovation, McGill University Health Centre, Montreal, Qc*
[2]*Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Qc*
[3]*Department of Anesthesia, McGill University Health Centre, Montreal, Qc*

Corresponding author:
Liane S. Feldman, MD FRCSC FACS
Director, Division of General Surgery
Montreal General Hospital L9-303
1650 Cedar Ave
Montreal, QC H3G 1A4
Tel: (514)934-1934 ext. 44004
Fax: (514)934-8210
Email: liane.feldman@mcgill.ca

Subject category:
Outcomes/health services research; Gastrointestinal

Presentation:
This data will be presented at the Outcomes Plenary Session at the 8th Annual Academic Surgical Congress (February 5, 2013) in New Orleans, LA.

Word count: 2851

ABSTRACT

Background: Many surgical innovations are costly but may result in faster patient recovery. Economic analyses of these innovations require utility measures that reflect the construct of "postoperative recovery". We investigated the validity of SF-6D utility value as a measure of postoperative recovery in patients undergoing elective colorectal resection.

Materials and Methods: Patients undergoing elective colorectal resection completed the SF-36 and the 6-minute walk test(6MWT) at baseline(before surgery), and at 4- and 8-weeks postoperatively. SF-6D utilities were derived from the SF-36. Longitudinal validity(responsiveness) was assessed using standardized response means(SRM). Construct validity was assessed by comparing the difference in mean SF-6D between patients with and without complications(discriminant) and by correlating the SF-6D with other measures of recovery(convergent).

Results: A total of 191 patients were included(58% male, mean age 63.0(SD 14.2) years, 81% malignancy, and 54% laparoscopic). SF-6D values dropped significantly from baseline to 4 weeks after surgery(SRM -0.55, $p<0.001$), and returned to baseline by 8 weeks(SRM -0.12, $p=0.111$). At 4-weeks after surgery, the SF-6D was lower in patients with complications than in those without(mean difference -0.043, 95% CI -0.083, -0.002). At all time points, the SF-6D correlated significantly with the physical and mental component scales of the SF-36(Pearson's r 0.47 to 0.80, all $p<0.001$) and the 6MWT(r 0.21 to 0.29, all $p<0.05$).

Conclusions: The SF-6D is a valid measure of postoperative recovery following elective colorectal resection and may be used to measure quality-adjusted life-years for cost-effectiveness analyses of surgical technologies and interventions hypothesized to impact recovery.

Keywords: health economics; utility; postoperative recovery; colorectal surgery; quality-adjusted life years; SF-6D

INTRODUCTION

Many surgical innovations in colorectal surgery, such as minimal access techniques and enhanced recovery after surgery pathways, are costly but may result in faster patient recovery. Cost-effectiveness analyses of new technologies are increasingly important in the era of financial constraints in health-care budgets. These studies require appropriate measures of effectiveness. The US Panel on Cost-Effectiveness in Health and Medicine has recommended the measurement of effectiveness in terms of quality-adjusted life years (QALYs), which are calculated by multiplying the time spent in a health state by the quality of life weight of that health state, which is measured in terms of utilities.[94]

Utility represents the preference of an individual for being in a particular health state, which is defined on a scale from 0 to 1.0, with 0 representing the worst health/death and 1.0 representing perfect health. This allows for the valuation of health status scaled relative to perfect health and death.[95] In particular, the US Panel recommended the use of preference-based health measures to generate utilities. Examples of these include the Euro-Qol-5D (EQ-5D)[96], Health Utilities Index (HUI)[97], and the Short Form-6D (SF-6D)[93]. Preference-based health measures that are used to generate utilities differ from other health-related quality of life instruments, such as the Short Form 36 (SF-36)[98] and the European Organization for Research and Treatment of Cancer questionnaire (EORTC-QLQ)[99], in that the latter have no valuation and therefore cannot be used to calculate QALYs . Utilities may also be generic, for example the EQ-5D and the SF-6D, or disease-specific, such as the King's Health Questionnaire[100] for urinary incontinence. Generic utilities have the advantage of being comparable across studies, but may be insensitive to disease-specific effects of interventions.

The SF-6D is a generic preference-based health measure that is derived from the SF-36, and is used to generate utilities in order to calculate QALYs for cost-effectiveness analyses. First described by Brazier et al.[93], it has rapidly become one of the most commonly used instruments for measuring utilities, partly because of the widespread use of the SF-36 as a generic quality of life instrument. The SF-36 is one of the most commonly reported quality of life measures in colorectal surgery[84], and if proved valid, the SF-6D would be a useful adjunct to the SF-36 for use in cost-effectiveness analyses. While it has been validated for use in endoscopic sinus surgery[101] and carpal tunnel surgery[102], as well as various chronic medical diseases[103,104], no studies have provided evidence for its validity as a measure of postoperative recovery after major abdominal surgery. Given that faster patient recovery is one of the much-touted benefits of many new surgical techniques, it is important to determine whether a generic utility value is sensitive to the complex "construct" of postoperative recovery. Therefore the objective of this study is to determine the longitudinal and construct validity of the SF-6D as a measure of postoperative recovery in patients undergoing elective colorectal resection.

METHODS

Patients: Data previously collected from three clinical trials, including two trials investigating preoperative exercise training[105,106] and one trial investigating thoracic epidural analgesia versus intravenous lidocaine for perioperative pain management[107], were analyzed. The study population consisted of patients undergoing elective colorectal resection at a single university-affiliated institution between 2005 and 2010. Inclusion criteria for the study were: 18 years of age or older, non-pregnant, and elective resection of colon and/or rectum. The choice of operative procedure and of approach (laparoscopy or open) were at the surgeon's discretion.

Postoperative care was standardized for all patients with an enhanced recovery after surgery pathway after 2009, while a portion of patients from 2005-2008 were managed with this pathway. An assessor who was blinded to patient-reported outcomes prospectively recorded patient demographics, operative characteristics, and postoperative outcomes up to 30 days after the initial surgery. Postoperative complications were graded as per the Clavien classification of surgical complications.[108] Complications were defined as minor if they required management at the bedside or pharmacological management without additional interventions (Clavien I-II), and as major if they required additional interventions or intensive care management (Clavien III+).

Outcome measures: All patients completed the SF-36 and the six-minute walk test (6MWT) at baseline (within 1 week before surgery), and at 4- and 8-weeks after surgery. The SF-36 is a widely-used generic quality of life questionnaire.[98] It measures eight different health dimensions: physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, and mental health. Two summary scores, the physical and mental component summary (PCS and MCS, respectively), were calculated and normalized using 1998 US population norms (mean 50, standard deviation 10).[109] The SF-6D is a multi-attribute utility value covering six health dimensions: physical functioning, role limitations, social functioning, pain, mental health and vitality. It is calculated using responses to 11-items on the SF-36, and weighted according to the health-state valuations of a sample of the general population in the United Kingdom. The SF-6D describes a total of 18,000 possible health states, and its valuation ranges from 0.296 (most severe problems in all six health dimensions) to 1.0 (no problems in any of the health dimensions; perfect health). The 6MWT is a validated measure of functional recovery after colorectal surgery[74] in which the patient is asked to walk for six minutes at a pace

that would make them tired along a flat corridor. The distance walked in six minutes is representative of the functional capacity of the patient, and is more reflective of the ability to perform the activities of daily living than maximal exercise tests.[73]

Validity: The validity of the SF-6D as a measure of postoperative recovery was assessed by several methods. Our hypothesis was that the SF-6D utilities would be significantly lower at 4-weeks after surgery compared to baseline, improve significantly from 4- to 8-weeks after surgery, and return to baseline values at 8-weeks after surgery (responsiveness). Criterion validity was not evaluated given the lack of a gold standard measure for postoperative recovery. Instead, construct validity was evaluated in two ways: 1) convergent validity (in which the SF-6D is correlated with other instruments used to measure the construct of postoperative recovery); and 2) discriminant validity (in which the ability of the SF-6D to differentiate between groups hypothesized to differ in the construct of postoperative recovery is evaluated; known-groups validity). We hypothesized that the SF-6D utilities would correlate with the 6MWT, PCS and MCS scores of the SF-36 at all time points (convergent validity) and that the SF-6D would be lower in patients with complications compared to those without at 4-weeks after surgery (discriminant validity). All hypotheses were formulated *a priori*.

Statistical analysis: Descriptive and summary statistics were calculated, as appropriate. Missing data were handled with multiple imputation (10 imputations).[110] In this procedure, missing items are estimated using the appropriate regression models from other complete data, including the outcome variable[111], and repeated ten times to generate ten datasets of imputed values. The between (datasets) and within (variable) variances are then taken into account according to

Rubin's rules[112] in the final analysis to account for the uncertainty in the imputed values. All imputed values were examined to ensure that there were no values outside the range of possible SF-6D values (0.296 to 1.0). The responsiveness of the SF-6D was evaluated using standardized response means (mean change in utilities divided by the standard deviation of the change; SRM) and paired t-tests between the different time points. SRM values are interpreted similarly to effect size: trivial, 0.0 to 0.2; small, 0.2 to 0.5; moderate, 0.5 to 0.8; and strong, greater or equal to 0.8.[113] SRMs are negative for overall deterioration, and positive for overall improvement. Convergent validity was assessed using Pearson's correlation (*r*). Multiple linear regression was performed to assess for the mean difference in SF-6D values at 4-weeks between patients with and without complications, after adjustment for age, American Society of Anesthesiologists (ASA) physical status, pathology and surgical approach (discriminant validity). All statistical analyses were performed with STATA 12 (StataCorp, College Station, TX) and R open-source software v2.13 (www.cran.r-project.org).

RESULTS

A total of 191 patients were included in the study. Baseline patient demographics and operative characteristics are reported in Table 1. The majority of the patients underwent a segmental colectomy and anastomosis. The overall incidence of complications was 46% (88/191), including 38% (72/191) minor complications and 8% (16/191) major complications. Mean length of stay was 7.7 (19.2) days [median 5 days, interquartile range 4-7]. No patients had received adjuvant therapy at 4-weeks after surgery, whereas 12% (22/191) had begun adjuvant therapy at 8-weeks. All patients had complete preoperative SF-6D and 6MWT data, and at least one follow-up value (either at 4- or 8-weeks after surgery). The proportion of missing SF-6D

values was 23% (44/191) and 32% (60/191) at 4- and 8-weeks, respectively. Data for missing items were imputed only if one follow-up value (either at 4- or 8-weeks after surgery) was available. Mean baseline, 4-, and 8-week postoperative SF-6D values were 0.744 (0.127), 0.663 (0.137), 0.722 (0.171), respectively (Figure 1). Mean values for the 6MWT, PCS and MCS scores are displayed in Figure 2. There were significant differences between baseline, 4- and 8-week SF-6D values compared to the SF-6D US population norm[114] that were especially pronounced 4-weeks after surgery (Table 2).

Responsiveness: The SF-6D values differed significantly across the perioperative time periods (Table 3). The degree of change was largest at 4-weeks compared to baseline (SRM -0.55; moderate) and the smallest at 8-weeks compared to baseline (SRM 0.12; trivial), consistent with a measure of postoperative recovery that should decrease at 4-weeks compared to baseline, and increase near baseline at 8-weeks. Floor and ceiling effects were minimal, as only 2 patients scored the maximum SF-6D value of 1.0 (both at baseline), and no patients scored the minimum SF-6D value of 0.296 at any of the time points.

Convergent validity: The SF-6D correlated significantly with the 6MWT, PCS and MCS scores, which are other instruments used to measure the construct of postoperative recovery (Table 4). Correlation with the 6MWT was weak but statistically significant across all time periods. The correlation between the SF-6D and the PCS was strongest at both time periods after surgery, but correlation between the SF-6D and the MCS was similar across all time periods.

Discriminant validity: Mean SF-6D values were lower in patients with complications at 4-weeks compared to those without, after adjusting for age, gender, ASA physical status and laparoscopic approach (Table 5), suggesting that the occurrence of a complication was the main driver of the decrease in SF-6D value at 4-weeks postop. However, there was no difference in SF-6D values 8-weeks after surgery between patients with and without complications on univariate analysis (unadjusted mean difference -0.021, 95% CI -0.065, 0.024), and also after adjusting for the age, gender, ASA physical status, laparoscopic approach, and whether they had received adjuvant therapy at the time of the 8-week assessment (Table 5). There were also no statistically significant differences in SF-6D at 4- and 8-weeks after surgery between patients with minor (Clavien I-II) and major complications (Clavien III+). There were no differences in SF-6D values at 4- or 8-weeks between patients with laparoscopic procedures compared to open procedures in the above analyses.

Given the fact that the presence of complications significantly affected changes in SF-6D over time, the responsiveness of the SF-6D was investigated in patients with and without complications (Table 3). The degree of change during the baseline to 4-week, 4- to 8-week and baseline to 8-week periods were identical between the two groups, however the SRM were larger in magnitude in complicated patients, suggesting that a more pronounced deterioration at 4-weeks after surgery, but also a more rapid increase at 8-weeks.

DISCUSSION

In order to calculate QALYs and perform cost-effectiveness analyses of surgical technologies that may improve recovery after surgery, there is a need for a preference-based health state utility that is a valid measure of the complex construct of postoperative recovery. In

this study, we provide evidence for the responsiveness, discriminant and convergent validity of the SF-6D health state utility as a measure of recovery after major elective colorectal surgery. While generic health-state classification measures such as the SF-6D allow for comparability across disease conditions, they should be validated for each setting in which they are used.[115] The SF-6D was clinically responsive and discriminated between patients with and without complications, even after adjustment for other patient factors that may affect postoperative recovery. The SF-6D also correlated with other validated measures of postoperative recovery, albeit weakly with functional walking capacity (6MWT).

In our study, we studied patients undergoing major colorectal resection only instead of including a wider range of major abdominal procedures. While limiting the generalizability of the results, this approach allowed for a more focused assessment of the SF-6D within a relatively homogenous population, without having to interpret the "noise" of a wide-range of disease or procedure-specific characteristics or postoperative complications. We also chose the SF-6D over other generic health-state classifications, such as the EQ-5D or the Health Utilities Index, due to the richer descriptive capability of the SF-6D, as well as the convenience of obtaining another measure from the widely-used SF-36.

Postoperative recovery can be considered as a period of deterioration immediately after the surgical intervention, followed by a "rehabilitation" phase back to baseline levels of functioning.[116] Therefore we expected the changes in the SF-6D to mirror this trajectory in order to be considered clinically responsive. The SF-6D has the strongest negative change in the period immediately following the operation, followed by a moderate change in the rehabilitation phase. The magnitude of these changes was greater than the mean clinically important difference of the SF-6D (0.041), identified by Walters and Brazier across range of patient groups[117], which further

supports the responsiveness of the SF-6D as a measure of postoperative recovery. There was no difference in mean SF-6D values at 8-weeks compared to baseline, suggesting that patients were recovered back to baseline at this point or that the instrument is no longer able to identify any differences compared to baseline. There were also no floor or ceiling effects at any time points, which is important since they may affect responsiveness.

The SF-6D was also able to discriminate between patients with and without complications at 4-weeks after surgery. There were no differences in SF-6D if complications were stratified by severity, but this may have been due to the small number of major complications. In addition, the impact of Clavien grade on health-related quality of life has not been reported to our knowledge. In light of the fact that complications were such a strong driver of changes in SF-6D, we performed a sensitivity analysis in which we assessed responsiveness stratified by the presence or absence of postoperative complications. This analysis demonstrated that the SF-6D remained a responsive measure, even in uncomplicated patients.

We also did not find a difference in utility as measured by the SF-6D at 4- and 8-weeks after surgery between patients undergoing laparoscopic or open surgery. This was not surprising given that previous studies that have investigated short-term quality of life, using the SF-36, after laparoscopic and open ileocolic resection[118], proctocolectomy with ileal-anal pouch anastomosis[119] and sigmoid resection[120] have failed to consistently demonstrate a difference between the two groups. Other studies that used different quality of life measures have also failed to demonstrate a difference in the short-term.[121-123]

As expected, the SF-6D had strong correlations with the physical and mental component summary scores of the SF-36 at all time points. The SF-36 has been used extensively as a measure of postoperative recovery[124-126]. Correlation with the 6MWT was weak, despite being

statistically significant. This was not unexpected, given that the 6MWT is an indicator of functional walking performance only, and may not correlate with other aspects of recovery such as social and emotional recovery.[74] Ideally we would have liked to correlate the SF-6D with other validated measures that cover specific aspects of recovery other than functional capacity, such as postoperative fatigue[127] or physical activity[71], however due to our study design this was not possible. Future studies that further investigate the validity of the SF-6D in surgical populations should correlate it with these measures.

The present study has several limitations. Firstly, this study was performed in a relatively homogeneous population of patients undergoing elective colorectal surgery, which excluded patients with severe medical comorbidities. Additional studies in other patient populations are required before our results can be generalized. Secondly, our study was performed using an existing dataset instead of designing a specific study to address this objective *a priori*. As a result, not all relevant measures were included. The potential effects of the preoperative interventions studied in those trials were accounted for by using results closest to surgery (i.e. post-intervention). Also, none of the interventions extended into the postoperative period. Thirdly, there may have been informative censoring in that some patients who suffered severe complications (due to debilitation, loss of independence, etc.) may have dropped out at a higher rate than non-complicated patients. These patients likely would have had very poor quality of life[128] (and by corollary scored very low on the SF-6D), and therefore our SF-6D estimates at 4- and 8-weeks after surgery may be artificially high. However this potential effect likely only biased our results towards the null. Given that our results were already clinically and statistically significant, inclusion of these patients would have only further increased the magnitude of the changes in SF-6D over time. Finally, there was a relatively large proportion of missing data,

which were handled with multiple imputation. This method has been shown to reduce bias in quality-of-life studies compared to complete case or last observation carried forward analyses[129], even with large proportions (10-30%) of missing data or if the data are missing not at random.[130,131]

In conclusion, we provide evidence for the validity of the SF-6D health-state utility as a measure of postoperative recovery in patients undergoing elective colorectal surgery. Therefore the SF-6D may be used to calculate QALYs for formal cost-effectiveness analyses of surgical technologies that are hypothesized to improve postoperative recovery.

Table 1 – Patient demographics and operative characteristics

	N = 191
Age, years (SD)	63.0 (14.2)
Male	110 (58%)
Body mass index, kg/m^2 (SD)	27.2 (4.9)
ASA status	
I	21 (11%)
II	130 (68%)
III	40 (21%)
Histology	
Malignancy*	149 (78%)
Stage I	37 (25%)†
Stage II	48 (32%)†
Stage III	45 (30%)†
Stage IV	18 (12%)†
Inflammatory bowel disease	18 (9%)
Other benign‡	24 (13%)
Operative procedure	
Right/left hemicolectomy	71 (37%)
Rectosigmoidectomy	44 (23%)
Low anterior resection	47 (24%)
Abdominoperineal resection	13 (7%)
Total procto- or subtotal colectomy	13 (7%)
Other	3 (2%)
Laparoscopic approach	104 (54%)
New stoma	62 (32%)

SD = standard deviation

*Including one case of colonic lymphoma on final pathology (not included in TNM stage numbers)

†Proportion of malignant cases

‡Includes diverticular disease, benign polyps, and polyposis syndromes

Table 2 – Mean differences in SF-6D utilities of the study sample at different time points compared to the United States population norm

		US Population Norm (0.791)		
Time point	**Mean SF-6D**	**Mean Difference (95% CI)**	**SRM***	**Degree of change**
Baseline	0.744	-0.047 (-0.065, -0.029)	-0.36	Small
4-weeks after surgery	0.663	-0.128 (-0.148, -0.109)	-0.94	Strong
8-weeks after surgery	0.722	-0.068 (-0.094, -0.044)	-0.40	Small

US = United States; CI = confidence interval; SRM = Standardized response mean *Calculated by dividing the mean difference in utility between the study sample and the US population norm by the standard deviation of the change

Table 3 – Responsiveness of the SF-6D to changes after surgery

Time period	Mean Difference (95% CI)	SRM*	Degree of change
All patients			
Baseline to 4-weeks after surgery	-0.080 (-0.102, -0.059)	-0.54	Moderate
4- to 8-weeks after surgery	+0.060 (0.037, 0.082)	0.39	Small
Baseline to 8-weeks after surgery	-0.022 (-0.049, 0.006)	-0.12	Trivial
Non-complicated patients			
Baseline to 4-weeks after surgery	-0.062 (-0.089, -0.034)	-0.32	Small
4- to 8-weeks after surgery	0.057 (0.030, 0.084)	0.31	Small
Baseline to 8-weeks after surgery	-0.005 (-0.039, 0.029)	-0.02	Trivial
Complicated patients			
Baseline to 4-weeks after surgery	-0.103 (-0.134, -0.070)	-0.46	Small
4- to 8-weeks after surgery	0.069 (0.044, 0.094)	0.40	Small
Baseline to 8-weeks after surgery	-0.033 (-0.070, 0.004)	-0.13	Trivial

CI = confidence interval; SRM = standardized response mean

*Calculated by dividing the mean difference in utility between time points by the standard deviation of the change

Table 4 – Pearson's correlations between the SF-6D and other instruments used to measure the construct of postoperative recovery

	Baseline	**4-weeks after surgery**	**8-weeks after surgery**
Six-minute walk distance	0.21 (p=0.004)	0.29 (p =0.003)	0.23 (p =0.014)
Short Form-36			
Physical component summary	0.67 (p <0.001)	0.70 (p <0.001)	0.76 (p <0.001)
Mental component summary	0.71 (p <0.001)	0.80 (p <0.001)	0.71 (p <0.001)

Table 5 – Multivariable analysis of SF-6D values at 4- and 8-weeks

	β estimate	***p* - value**	**95% CI**
Dependent variable: SF-6D at 4-weeks after surgery			
Complications	-0.047	0.038	(-0.088, -0.006)
Age, per year	0.001	0.156	(-0.0004, 0.003)
Male gender	0.037	0.063	(-0.002, 0.076)
Laparoscopic approach	0.031	0.142	(-0.010, 0.072)
ASA III (vs. I/II)	-0.025	0.396	(-0.082, 0.033)
Histology (vs. benign)			
Malignancy	-0.010	0.749	(-0.071, 0.051)
Inflammatory bowel disease	0.044	0.392	(-0.058, 0.145)
Dependent variable: SF-6D at 8-weeks after surgery			
Complications	-0.022	0.242	(-0.064, 0.021)
Age, per year	0.002	0.015	(0.0003, 0.004)
Male gender	0.079	<0.001	(0.036, 0.123)
Laparoscopic approach	-0.007	0.629	(-0.049, 0.034)
ASA III (vs. I/II)	-0.056	0.044	(-0.071, -0.002)
Adjuvant therapy*	-0.066	0.050	(-0.131, -0.001)

ASA = American Society of Anesthesiologists physical status

*Whether the patient had received adjuvant therapy at the time of 8-week assessment

Figure 1 – Change in SF-6D over time (mean, 95% confidence interval). * denotes $p < 0.05$ for the mean difference compared to the previous time point.

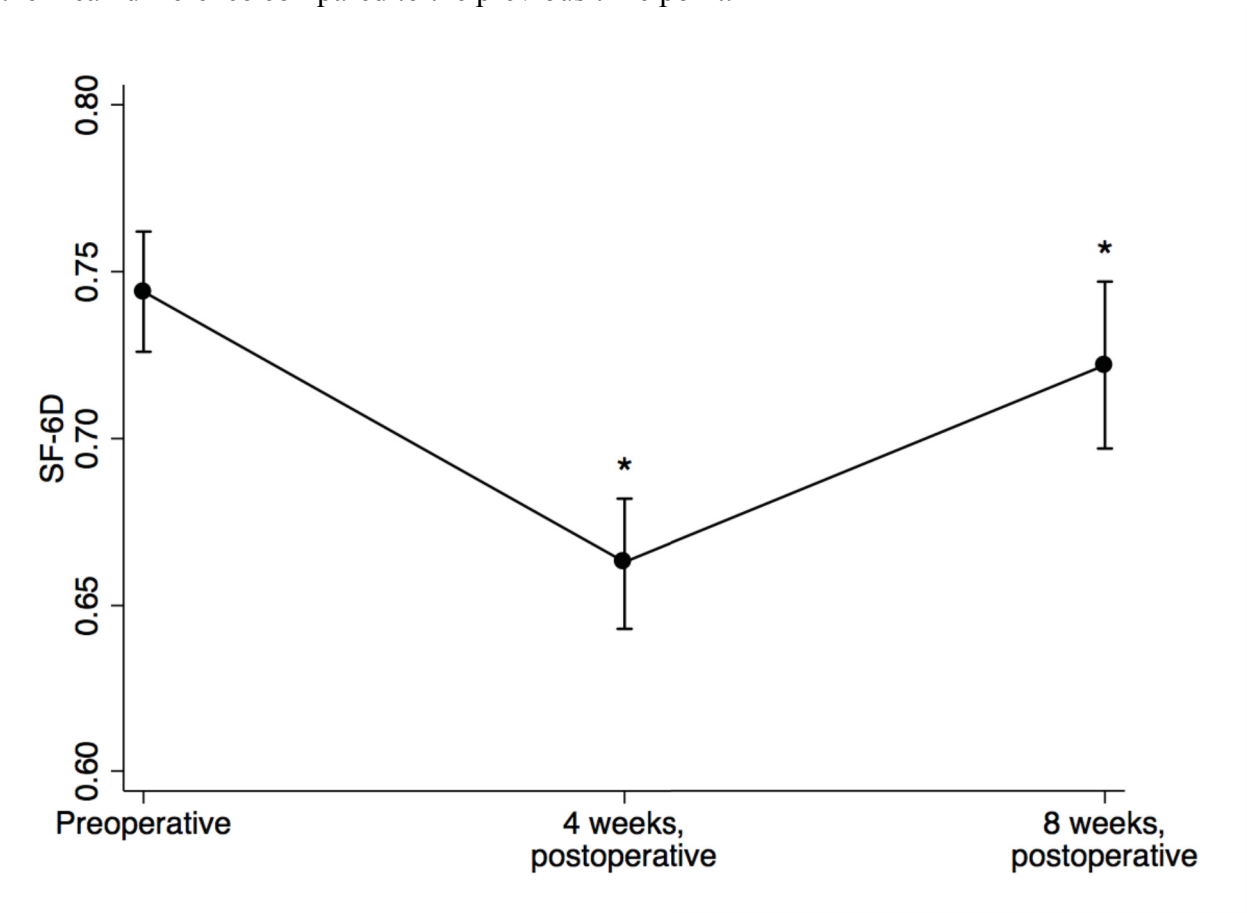
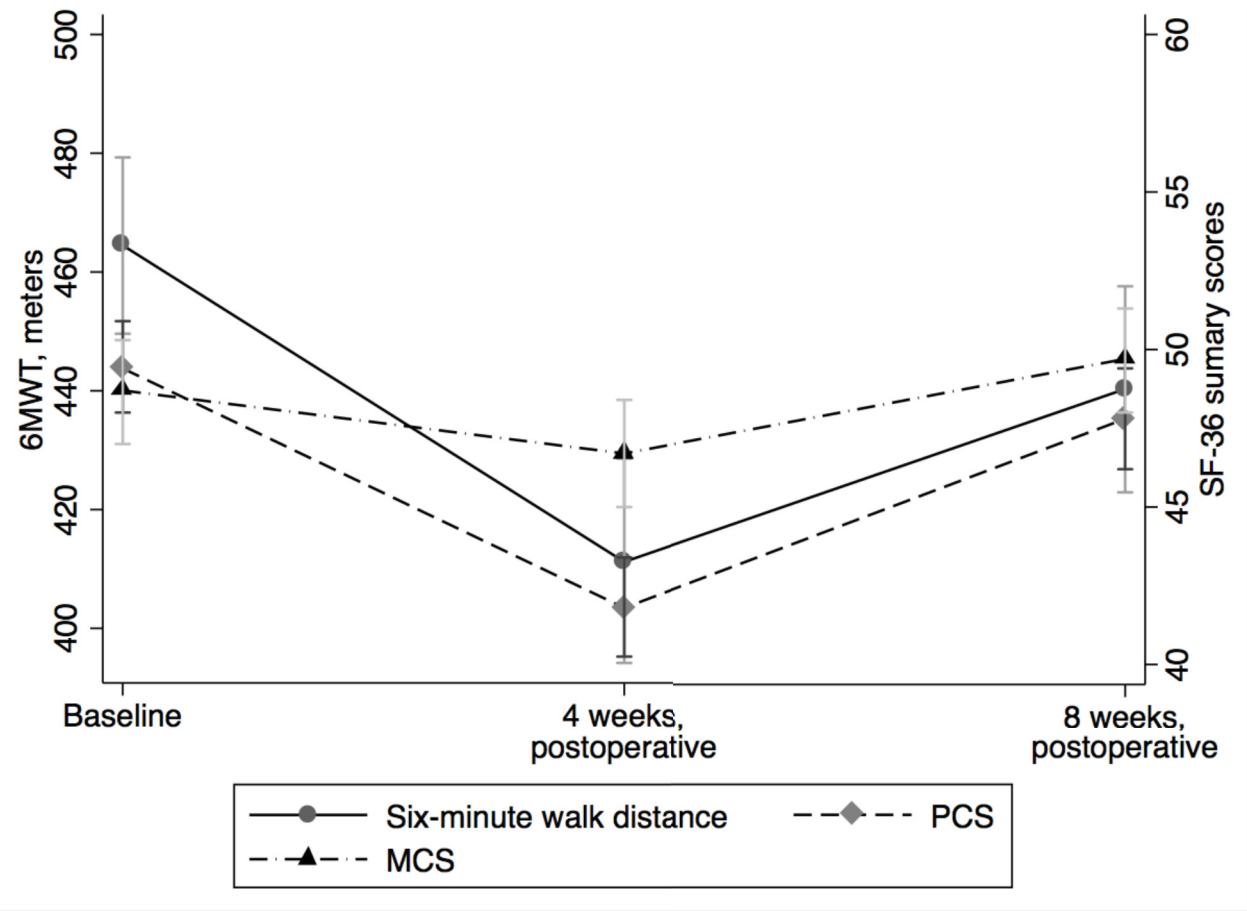


Figure 2 – Changes in six-minute walk distance, and physical and mental component summary scores over time (mean, 95% confidence interval)



6. COMPARING THE VALIDITY OF THE SF-6D AND THE EQ-5D AS MEASURES OF POSTOPERATIVE RECOVERY

6.1 PREAMBLE

Chapter 5 demonstrated the validity of the SF-6D as a measure of postoperative recovery after colorectal surgery.[91] It was responsive to the expected postoperative changes, correlated with a validated physical performance measure, the six-minute walk test[74], and discriminated between patients with and without complications, as well as between patients undergoing colorectal surgery and population norms. However, the SF-6D is one of the many existing indirect utility instruments. Other important instruments include the EQ-5D[96], Health Utilities Index[132], Quality of Well-Being[133], and Assessment of Quality of Life Scale[134]. However, each instrument covers different domains over different timeframes (Table). Furthermore, there is extensive literature to suggest that the results obtained from each instrument may not be interchangeable.[135] It is therefore imperative to choose a utility instrument that is practical, reliable, and valid for the specific setting and population under study.[136]

Postoperative recovery is a multidimensional construct that includes the physical, psychological, social, and habitual aspects.[64,137] These domains are well represented by the SF-6D, EQ-5D, and the Quality of Well-Being scale, but the SF-6D and the EQ-5D are likely to be more practical due to their lower number of items (Table). The EQ-5D is the easiest to administer as it contains only five items (along with a visual analog scale), and it is the most commonly used preference-based measure of health in cost-utility analyses.[138] However, there may be wide variations in cost-effectiveness ratios depending on whether the SF-6D or the EQ-5D is used to generate quality-adjusted life years[139], and the EQ-5D has not yet been validated for postoperative recovery. In the following manuscript, the validity of the SF-6D and the EQ-5D as measures of postoperative recovery after colorectal surgery is compared. This manuscript has

been published in the *Journal of Surgical Research* (J Surg Res 2014 (epub 2014 Feb 15) DOI 10.1016/j.jss.2014.02.016).

Table 6.1 – A comparison of the important multi-attribute utility measures

Instrument	Items	Recall Period	Domains	Utility Range
SF-6D[93]	11*	1 week, 1 month	1. Physical functioning 2. Role limitations 3. Social functioning 4. Pain 5. Mental health 6. Vitality	0.30 to 1
EQ-5D[96]	5	1 day	1. Mobility 2. Self-care 3. Usual activities 4. Pain 5. Anxiety/depression	-0.59 to 1
Health Utilities Index v3[132]	8	1 week, 2 weeks, 4 weeks	1. Vision 2. Hearing 3. Speech 4. Ambulation 5. Dexterity 6. Emotion 7. Cognition 8. Pain	-0.36 to 1
Quality of Well-Being Scale[133]	74	3 days	1. Symptoms 2. Self-care 3. Mobility 4. Usual activities 5. Physical activity	0 to 1
Assessment of Quality of Life Scale (6D)[134]	20	1 week	1. Independent living 2. Mental health 3. Relationships 4. Senses 5. Coping 6. Pain	-0.04 to 1

*Nested within 10 items of the SF-36

6.2 A COMPARISON OF THE VALIDITY OF TWO INDIRECT UTILITY INSTRUMENTS AS MEASURES OF POSTOPERATIVE RECOVERY

Lawrence Lee MD MSc[1], Juan Mata MD[1], Berson Augustin[2], Franco Carli MD[2], Nancy Morin MD[3], Eric Latimer PhD[4], Liane S. Feldman MD[1]
[1]Steinberg-Bernstein Centre for Minimally Invasive Surgery and Innovation, Dept. of Surgery,
[2]Dept. of Anaesthesia, McGill University Health Centre, Montreal, QC
[3]Dept. of Surgery, Sir Mortimer B. Davis Jewish General Hospital, Montreal, QC
[4]Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, QC

Running Head
Validation of the SF-6D and EQ-5D for recovery

Subject Category
Outcomes/health services research; Gastrointestinal

Sources of Financial Support
The Steinberg-Bernstein Centre for Minimally Invasive Surgery and Innovation is supported by an unrestricted educational grant from Covidien. LL is supported by the Quebec Health Science Research Scholarship (FRQ-S) and the McGill Surgeon Scientist Program.

Meeting Information
These data will be presented at the 9th Annual Academic Surgical Congress (San Diego, CA) on February 4-6, 2014.

Corresponding Author
Lawrence Lee, MD MSc
1650 Cedar Ave, E19-125
Montreal, QC H3G 1A4
Telephone: (514) 934-6237
Fax: (514) 934-8475
Email: larry.lee@mail.mcgill.ca

ABSTRACT

Introduction: Cost-effectiveness analyses of surgical interventions require valid measures of postoperative recovery. The objective of this study is to compare the validity of two indirect utility instruments, the SF-6D and EQ-5D, as measures of postoperative recovery.

Materials and Methods: A prospective cohort of patients undergoing elective colorectal resection at two university institutions from 10/2012 to 10/2013 completed the SF-6D and the EQ-5D (including the EQ-VAS) at baseline (before surgery), and at 4- and 8-weeks after surgery. Responsiveness and construct validity were assessed through *a priori* hypotheses.

Results: A total of 165 patients were included. The SF-6D was the most responsive to the expected postoperative changes at 4- and 8-weeks, compared to the EQ-5D and the EQ-VAS. The 4-week SF-6D, EQ-5D, and EQ-VAS discriminated between patients with and without complications after controlling for confounders, with adjusted mean differences of -0.070 (95% CI -0.126, -0.015), -0.133 (95% CI -0.231, -0.030), and -7.91 (95% CI -14.77, -1.04), respectively. Mean SF-6D and EQ-5D values were significantly different from US population norms at all time points, but the magnitude of change was highest for the SF-6D. The strength of correlation between all three instruments was moderate at all time points (r 0.550 to 0.684, all $p<0.05$).

Conclusions: The SF-6D preference-based health index appears to be a more valid measure of postoperative recovery than the EQ-5D and EQ-VAS in surgical cost-effectiveness analyses.

Keywords: Utility; recovery; SF-6D; EQ-5D; colorectal surgery; validity

INTRODUCTION

Many new surgical innovations are advocated to improve recovery, but few have been fully evaluated. Cost-effectiveness analyses of these new technologies require valid measures of postoperative recovery. However, it is difficult to compare studies to one another as multiple different outcomes have been used to measure "recovery".[24,140,141] The Panel on Cost-Effectiveness in Health and Medicine recommends using quality-adjusted life years (QALYs) to measure effectiveness in economic evaluations, as QALYs incorporate multiple different outcomes into a single summary measure.[94] QALYs are calculated by multiplying the duration spent in a particular health state by the quality-of-life weight of that health state, which is measured in terms of "utilities". Utilities are measured on a scale from 0.0 (death) to 1.0 (perfect health). While QALYs can be directed elicited, they are commonly measured using indirect utility instruments, such as the Short Form 6D (SF-6D)[93] or the EuroQol 5D (EQ-5D)[96].

However, there may be significant variability in utility values depending on the instrument used, which can have a large effect on the interpretation of cost-effectiveness results.[142] The choice of which utility instrument to use depends on an instrument's practicality, reliability, and validity for the specific condition under investigation.[136] For conditions such as knee pain, the EQ-5D is superior to the SF-6D[143], but the opposite is true for rheumatoid arthritis[144]. The SF-6D has been previously validated as a measure of postoperative recovery after colorectal surgery specifically for use in cost-effectiveness analyses[91], but the EQ-5D is more commonly used and easier to administer.[138] No previous study has investigated the validity of the EQ-5D to value postoperative recovery. Therefore the objective of this study was to compare the validity of two important indirect utility instruments, the SF-6D and EQ-5D, as measures of postoperative recovery in patients undergoing elective colorectal resection.

MATERIAL AND METHODS

Patients

Consecutive patients undergoing elective colorectal surgery at two university-affiliated institutions between October 2012 and October 2013 were approached for participation at the preoperative clinic (1-2 weeks prior to surgery). Patients were eligible if they were older than 18 years and had a scheduled resection of the colon and/or rectum. Patients were excluded if they did not speak English or French, or had neuropsychiatric conditions or cognitive impairments that interfered with completion of the study questionnaires. Patients who did not undergo resection during the operation were also excluded. Demographics, operative details, postoperative course and final pathology were collected prospectively. Comorbidities were measured using the Charlson Comorbidity Index.[145] Postoperative complications were graded as per the Clavien classification of surgical complications.[108] Outcomes were collected up to 60-days postoperatively.

Measures

All participating patients completed the SF-36[146], from which the SF-6D is derived, and the EQ-5D[147] at baseline (within 2 weeks prior to surgery), and at four- and eight-weeks after surgery. The SF-6D is a multi-attribute indirect utility measure that is nested within the SF-36.[93] It measures six dimensions: physical functioning, role limitations, social functioning, pain, mental health, and vitality, for a possible 18,000 unique health states. Quality of life weights were obtained from a representative sample of 611 members the UK general population using

standard gamble.[93] The SF-6D ranges from 0.296 (most severe impairment in all six dimensions) to 1.000 (perfect health).

The EQ-5D is a five-item instrument measuring mobility, self-care, usual activities, pain, and anxiety/depression. Each item is scored on a three-level scale, for a possible 243 unique health states. Different valuations have been published, but the UK tariff[96] (derived from a representative sample of 3395 members of the UK general population using time trade-off) was used to maintain comparability with the SF-6D. The EQ-5D ranges from -0.594 (most severe impairments in all 5 dimensions) to 1.000 (perfect health). The EQ-5D also contains a visual analog scale (EQ-VAS), which asks respondents to rate their present health on a scale from 0 (worst imaginable health state) to 100 (best imaginable health state).

Validity

All three measures were examined for responsiveness and construct validity. All hypotheses were made *a priori* and were based on previous validity data.[91] In order to be considered responsive, we hypothesized that each measure would be lower at four-weeks compared to baseline, improve from four- to eight-weeks, and return to baseline by eight-weeks. This analysis was also stratified by resection type (colonic versus rectal) to account for differences in postoperative functional outcomes (which in turn are indirectly measured by these instruments). Construct validity was assessed in several ways: first, we hypothesized that each measure would discriminate between patients with and without complications at four-weeks, between colonic and rectal resections at all time points, and between patients receiving adjuvant therapy by eight-weeks and those who did not; and second, each measure would be different

from population norms at all time points. US population mean values for the SF-6D, EQ-5D (using the UK tariff), and EQ-VAS were obtained from a study by Hanmer and colleagues.[148]

Statistical Analysis

Summary descriptive data were expressed as proportion (n), mean (SD), or median [IQR], as appropriate. Correlations between each measure were demonstrated using Pearson's correlation (*r*). Bland-Altman plots were used to demonstrate the agreement between measures.[149] In these plots, the difference in scores is plotted against the mean score, along with the limit of agreement (±2SD). If the limits of agreement between scores are not clinically important, then the two measures can be used interchangeably.[150] Responsiveness was assessed using standardized response means (SRM), calculated by dividing the mean difference between time periods by the standard deviation of the change, and paired *t*-tests. SRM are interpreted in the same manner as effect sizes: trivial, 0.0-0.2; small, 0.2-0.5; moderate, 0.5-0.8; and strong, >0.8.[113] Mean sample and population values at each time period were compared using mean differences (95% confidence intervals) and SRM. Multiple linear regression was performed to determine the mean difference between patients groups for each measure, adjusting for confounders.

Missing data were handled with multiple imputation using chained equations (ten imputations). In this procedure, missing items are estimated using the appropriate regression model (truncated linear regression using the relevant lower and upper values for each measure) from other observed data, taking into account the longitudinal nature of the data, and repeated ten times to generate ten different imputed datasets. Final uncertainty around point estimates

incorporate the between (datasets) and within (variable) variances, according to Rubin's rules.[112] All statistical analyses were performed using STATA 12 (StataCorp, College Station, TX).

RESULTS

A total of 172 patients were enrolled in the study, of which seven patients were excluded from the analysis (three patients underwent open and closure due to peritoneal carcinomatosis, one patient underwent a synchronous liver resection with a right hemicolectomy, and three patients withdrew), leaving a total of 165 patients for analysis. Missing data were present in 13% (21/165) of patients overall (all eight-week assessments). Baseline patient and operative characteristics are reported in Table 1. The cohort included 64 patients (39%) who underwent rectal resections (anterior, low anterior, and abdominoperineal resection, and proctocolectomy). Median length of stay (including readmissions within 60-days) was 6 days [IQR 3-8]. The incidence of 30-day complications was 41% (68/165), and 44% (72/165) at 60-days. The most severe complication was graded Clavien I in 17% (28/165), Clavien II in 17% (28/165), Clavien III in 4% (7/165), and Clavien IV in 4% (7/165). There were two mortalities within 60-days of the initial operation.

Mean SF-6D, EQ-5D, and EQ-VAS values are shown in Figure 1. The correlations between each measure were moderate at all time points (r = 0.550 to 0.684; Table 2), but the agreement between the three measures was poor, especially at the lower end of the utility range. Bland-Altman plots are only shown for baseline values (Figure 2A-C), but were similar at 4- and 8-weeks after surgery (data not shown). There was also significant ceiling effect for the EQ-5D, as 13% of values across all time periods were at the upper limit of the scale, compared to less

than 1% for the SF-6D and 0% for the EQ-VAS. None of the measures demonstrated a floor effect.

Responsiveness

The SF-6D was the only measure that followed the expected trajectory of postoperative recovery, as it exhibited a significant decrease at four-weeks compared to baseline, followed by a significant increase from four- to eight-weeks (Table 3). Mean SF-6D values returned to baseline values by eight-weeks after surgery. The degree of change in SF-6D values was strongest at four-weeks compared to baseline (SRM -0.51, moderate), and smallest at eight-weeks compared to baseline (SRM -0.15, trivial). Comparatively, neither the EQ-5D nor the EQ-VAS were responsive to expected postoperative changes, as no significant differences between baseline and four-week mean values were found. There were only trivial changes in mean EQ-5D values over time. The EQ-VAS exhibited a small degree of change (SRM +0.28) at eight-weeks compared to four-weeks, but trivial changes (SRM -0.10 and +0.08) for the other two time points. For all measures, colonic resections had a smaller decrease at four-weeks compared to baseline, and a larger increase over the four- to eight-week period, although there were no differences in the magnitudes of change between colonic and rectal resections (except for the SF-6D over the baseline to four-week interval).

Construct Validity

At the four-week assessment, all three measures discriminated between patients who experienced a postoperative complication within 30-days and those who did not (Table 4). Even after adjusting for age, gender, comorbidities, laparoscopic approach, type of resection (colonic

versus rectal), and whether a new stoma was fashioned, the mean differences remained significant. In the same multivariate analysis, none of the three measures were able to discriminate between patients who underwent a laparoscopic resection and those who underwent an open procedure (adjusted mean difference for SF-6D: 0.009 (95% CI -0.034, 0.052); EQ-5D: 0.013 (95% CI -0.084, 0.111); and EQ-VAS: 5.33 (95% CI -1.76, 12.43)). Similarly, none of the measures discriminated between colonic and rectal resections at four-weeks after adjusting for age, gender, comorbidities, laparoscopic approach, whether a new stoma was fashioned, and postoperative complications (adjusted mean difference for SF-6D: -0.004 (95% CI -0.053, 0.045); EQ-5D: -0.043 (95% CI -0.136, 0.048); and EQ-VAS: -2.55 (95% CI -9.09, 3.98)). At eight-weeks, none of the measures were able to discriminate between patients undergoing colonic versus rectal resections, or between patients who received adjuvant therapy versus those who did not, after adjusting for age, gender, comorbidities, whether a new stoma was fashioned, and postoperative complications.

Mean SF-6D and EQ-5D values were significantly different from US population norms at all time points, but at four-weeks, the degree of change was moderate for the SF-6D and small for the EQ-5D (Table 5). Mean EQ-VAS values were significantly different compared to population norms at four-weeks, and the degree of change was small. The SF-6D and the EQ-5D exhibited baseline differences between colonic versus rectal resection and population norms (mean difference in SF-6D: colonic -0.009 (95% -0.031, 0.012), SRM -0.08 and rectal -0.045 (95% CI -0.081, -0.010), SRM -0.32; mean difference in EQ-5D: colonic -0.019 (95% CI -0.068, 0.030), SRM -0.08 and rectal -0.083 (95% CI -0.165, -0.001), SRM -0.26), whereas the EQ-VAS did not. No differences in the magnitudes of change between type of resection and population norms were found at four- and eight-weeks.

DISCUSSION

The use of indirect utility instruments, such as the SF-6D and EQ-5D, as outcome measures for clinical trials and economic evaluations has increased. However, there may be important differences between instruments leading to potential variations in utilities. It is therefore essential that the validity of the instrument is confirmed within the specific context and patient population in order to adequately measure the effects of the intervention under study.[136] This study has reported superior validity for the SF-6D compared to the EQ-5D and the EQ-VAS in the context of postoperative recovery after colorectal surgery.

Postoperative recovery is characterized by a period of immediate deterioration after surgery, followed by a gradual rehabilitation period back to or surpassing baseline levels.[151] In the present study, the SF-6D was most responsive to the hypothesized postoperative changes. The degrees of change of the present study were similar to a previous SF-6D validation study, suggesting reproducibility of these results.[91] The statistically significant changes in mean SF-6D values were also clinically relevant, as the differences were greater than the minimal clinically important difference for the SF-6D (±0.041).[117] This may be partly explained by the greater descriptive ability of the SF-6D, which contains a possible 18,000 health states, whereas the EQ-5D contains only 243 possible states. A more detailed version of the EQ-5D exists with more levels per domain, i.e. the EQ-5D-5L[152], but it is not yet widely used. A more descriptive system is likely to capture the multidimensional construct of postoperative recovery with greater precision. Furthermore, the EQ-5D exhibits significant ceiling effect, as it is not able to distinguish patients with low levels of disability[153], which was observed in the present study. The different timeframes of the EQ-5D and SF-6D may also affect their responsiveness, as the EQ-

5D asks respondents to rate their present day health, compared to the SF-6D, which asks respondents to consider the previous week. As the values are likely to vary on a day-to-day basis, the EQ-5D may be less stable. It is unclear whether more assessments would affect the performance of the EQ-5D. Comparatively, the SF-6D is less likely to be affected by 'present day bias'[154] as it asks respondents to consider their health over a longer timeframe. In terms of the EQ-VAS, it is not surprising this measure did not demonstrate responsiveness, as single item rating scales are imprecise, unreliable, likely to be interpreted in different ways by the respondents, and unlikely to representative of a complex construct.[155] These results are in keeping with other studies, as the EQ-5D had returned to near-baseline values by 4-weeks in patients undergoing major hepatic resection.[12]

This study also reported a relatively poor agreement between EQ-5D, EQ-VAS, and SF-6D scores, especially at lower values of utility. These results are not unexpected, as previous studies have also reported large discrepancies in values and minimal clinically important differences between the SF-6D and the EQ-5D.[117,135] Therefore the choice of instrument may significantly affect the interpretation of clinical trials and economic evaluations. For example, there was a two- to threefold increase in QALYs in a trial comparing different exercise regimens for post-menopausal women if the SF-6D were used instead of the EQ-5D.[156] Similar results were reported in a different trial comparing different treatments for knee pain[142]. These findings further reinforce the importance of selecting an instrument that has been validated for postoperative recovery in order to assess the cost-effectiveness of surgical interventions hypothesized to improve recovery.

In the absence of a gold standard measure of recovery, construct validity was evaluated by testing several *a priori* hypotheses. Construct validity was demonstrated in all three measures

to a certain degree, but the SF-6D exhibited more clinically relevant differences. All three measures were able to discriminate between patients with and without complications. Importantly, the mean differences were larger than the minimal clinically important difference for both the SF-6D and the EQ-5D (±0.041 and ±0.074, respectively).[117] None of the three measures were able to discriminate between patients undergoing laparoscopic versus open resection. This finding is not unexpected, as few studies have reported a difference in quality of life between laparoscopic and open resection.[78,91] In addition, none of the measures were able to detect statistically significant differences colonic and rectal resections at any time point, or between patients who received adjuvant therapy by 8-weeks versus those who did not. The absence of significant differences may be partly explained by sample size limitations, as well as the fact that postoperative complications and whether a new stoma was fashioned are more likely to be stronger determinants of short-term quality of life after colorectal surgery than the type of resection.[157] When compared to population norms, both the SF-6D and the EQ-5D were significantly different across all time points, but again the SF-6D exhibited stronger and more clinically significant changes. At baseline, both the SF-6D and the EQ-5D exhibited important differences in the magnitudes of change between colonic versus rectal resections and population norms, which can be explained by the high proportion of patients undergoing rectal resections that received neoadjuvant therapy.

This study has to be interpreted in light of several limitations. First, this study was performed in a relatively homogeneous population of patients undergoing elective colorectal surgery. It remains to be seen if the results can be generalized to other procedures, especially those that are less invasive, as the magnitude of the changes may be different. Second, relatively few assessments were performed in the postoperative period, which may have limited the ability

to clearly delineate the trajectory of recovery. However, the additional information obtained from more frequent assessments must be balanced against pragmatic considerations, such as increased patient burden. There was also a moderate amount of missing data, which were handled via multiple imputation. While the missing-at-random requirement for valid multiple imputation analysis can never be definitively proven, patients in this study with missing data did not have a higher incidence of postoperative complications, which is likely to be the main reason for missingness. Finally, UK valuations of the SF-6D and EQ-5D were used to maintain comparability, as US or Canadian tariffs for the SF-6D do not exist. This may potentially change the values obtained from these instruments, as there may be differences in valuations between countries.[158]

In summary, the SF-6D demonstrated superior validity evidence as a measure of postoperative recovery in patients undergoing elective colorectal surgery, compared to the EQ-5D and the EQ-VAS. However, these findings require further validation in other surgical populations and procedures before they can be widely applied. Within the context of colorectal surgery, our results suggest that the SF-6D should be preferentially used over the EQ-5D and the EQ-VAS to calculate QALYs for cost-effectiveness analyses of surgical interventions hypothesized to improve recovery.

Table 1 – Baseline patient and operative characteristics

	N=165
Age, years (SD)	62.8 (13.6)
Male gender	82 (50%)
Body mass index, kg/m^2 (SD)	27.2 (5.6)
Charlson comorbidity index	
0	38 (23%)
1-2	73 (44%)
3-4	33 (20%)
5+	21 (13%)
Neoadjuvant therapy	21 (13%)
Malignancy	106 (64%)
Laparoscopic approach	99 (60%)
Procedure	
Right hemicolectomy	65 (39%)
Rectosigmoidectomy	48 (29%)
Low anterior resection	22 (13%)
Subtotal/procto-colectomy	9 (6%)
Left hemicolectomy	8 (5%)
Abdominoperineal resection	6 (4%)
Other	7 (4%)
New stoma	43 (26%)
Adjuvant therapy started by 8 weeks*	21 (13%)

*At the time of 2nd follow-up questionnaire

Table 2 – Correlation matrix (Pearson's *r*) of SF-6D, EQ-5D, and EQ-VAS values at all time points. All $p < 0.001$

	Baseline			*4-week*			*8-week*		
	SF-6D	**EQ-5D**	**EQ-VAS**	**SF-6D**	**EQ-5D**	**EQ-VAS**	**SF-6D**	**EQ-5D**	**EQ-VAS**
SF-6D	1.000			1.000			1.000		
EQ-5D	0.596	1.000		0.667	1.000		0.679	1.000	
EQ-VAS	0.550	0.506	1.000	0.641	0.656	1.000	0.684	0.646	1.000

Table 3 – Responsiveness of the SF-6D, EQ-5D and EQ-VAS to hypothesized postoperative changes, stratifying for resection type (colonic versus rectal). Rectal resections include anterior, low anterior, and abdominoperineal resections, and proctocolectomies.

	Mean Difference (95% CI)	**SRM**	**Degree of Change**
SF-6D			
Baseline to 4-weeks			
Overall	-0.082 (-0.107, -0.051)	-0.51*	Moderate
Colon resection	-0.071 (-0.110, -0.032)	-0.46*	Small
Rectal resection	-0.091 (-0.123, -0.059)	-0.57*	Moderate
4-weeks to 8-weeks			
Overall	0.045 (0.019, 0.072)	0.42*	Small
Colon resection	0.058 (0.032, 0.074)	0.48*	Small
Rectal resection	0.033 (0.007, 0.059)	0.31*	Small
Baseline to 8-weeks			
Overall	-0.017 (-0.001, 0.034)	-0.15	Trivial
Colon resection	-0.017 (-0.049, 0.014)	-0.11	Trivial
Rectal resection	-0.017 (-0.061, 0.026)	-0.10	Trivial
EQ-5D			
Baseline to 4-weeks			
Overall	-0.033 (-0.085, 0.020)	-0.10	Trivial
Colon resection	-0.032 (-0.089, 0.027)	-0.10	Trivial
Rectal resection	-0.034 (-0.116, 0.046)	-0.11	Trivial
4-weeks to 8-weeks			
Overall	0.023 (-0.029, 0.074)	0.07	Trivial
Colon resection	0.043 (-0.014, 0.099)	0.15	Trivial
Rectal resection	-0.008 (-0.093, 0.077)	-0.03	Trivial
Baseline to 8-weeks			
Overall	-0.002 (-0.052, 0.046)	-0.01	Trivial
Colon resection	0.012 (-0.046, 0.069)	0.04	Trivial
Rectal resection	-0.024 (-0.118, 0.069)	-0.07	Trivial
EQ-VAS			
Baseline to 4-weeks			
Overall	-2.74 (-7.02, 1.54)	-0.10	Trivial
Colon resection	-2.10 (-7.82, 3.60)	-0.07	Trivial
Rectal resection	-3.67 (-10.32, 2.88)	-0.14	Trivial
4-weeks to 8-weeks			
Overall	5.05 (2.22, 7.87)	0.28*	Small
Colon resection	5.00 (1.50, 8.50)	0.28*	Small
Rectal resection	5.12 (0.02, 10.21)	0.26*	Small
Baseline to 8-weeks			
Overall	2.12 (-1.89, 6.14)	0.08	Trivial
Colon resection	2.85 (-2.32, 8.03)	0.11	Trivial
Rectal resection	0.93 (-6.02, 7.92)	0.03	Trivial

*$p<0.05$

Table 4 – Simple and adjusted mean differences in SF-6D, EQ-5D, and EQ-VAS between patients with and without complications at four-weeks.

	Mean Difference (95% CI)		
Measure	**Simple**	**Adjusted***	**SRM† (degree of change)**
SF-6D (range 0.296 to 1.00)	-0.082 (-0.133, -0.032)	-0.070 (-0.126, -0.015)	-0.21 (small)
EQ-5D (range -0.594 to 1.00)	-0.163 (-0.264, -0.060)	-0.133 (-0.231, -0.030)	-0.21 (small)
EQ-VAS (range 0 to 100)	-10.26 (-17.57, -2.95)	-7.91 (-14.77, -1.04)	-0.18 (trivial)

*Adjusted for age, male gender, comorbidities, laparoscopic approach, resection type (colonic vs. rectal), and whether a new stoma was fashioned
†Adjusted SRM only

Table 5 – Comparison of mean cohort and population values for the SF-6D, EQ-5D, and EQ-VAS

	Baseline	**4-week**	**8-week**
SF-6D (range 0.296 to 1.000)	*US population norm: 0.803*		
Mean difference (95% CI)	-0.023 (-0.041, -0.003)	-0.105 (-0.129, -0.080)	-0.048 (-0.071, -0.024)
SRM (degree of change)	-0.18* (trivial)	-0.66* (moderate)	0.31* (small)
EQ-5D (range -0.594 to 1.000)	*US population norm: 0.786*		
Mean difference (95% CI)	-0.044 (-0.086, -0.001)	-0.074 (-0.124, -0.025)	-0.052 (-0.099, -0.005)
SRM (degree of change)	-0.16* (trivial)	-0.23* (small)	0.17* (trivial)
EQ-VAS (range 0 to 100)	*US population norm: 76.9*		
Mean difference (95% CI)	-2.07 (-5.46, 1.32)	-4.96 (-8.52, -1.39)	0.10 (-2.79, 3.00)
SRM (degree of change)	-0.09 (trivial)	-0.22* (small)	0.01 (trivial)

*$p<0.05$

Figure 1 – Mean SF-6D, EQ-5D and EQ-VAS values at baseline and at 4- and 8-weeks after surgery. EQ-VAS values are shown on a 0.0-1.0 scale rather than the original 0-100 scale. The curves are staggered to better represent the confidence intervals, and do not represent differences in time between assessments.

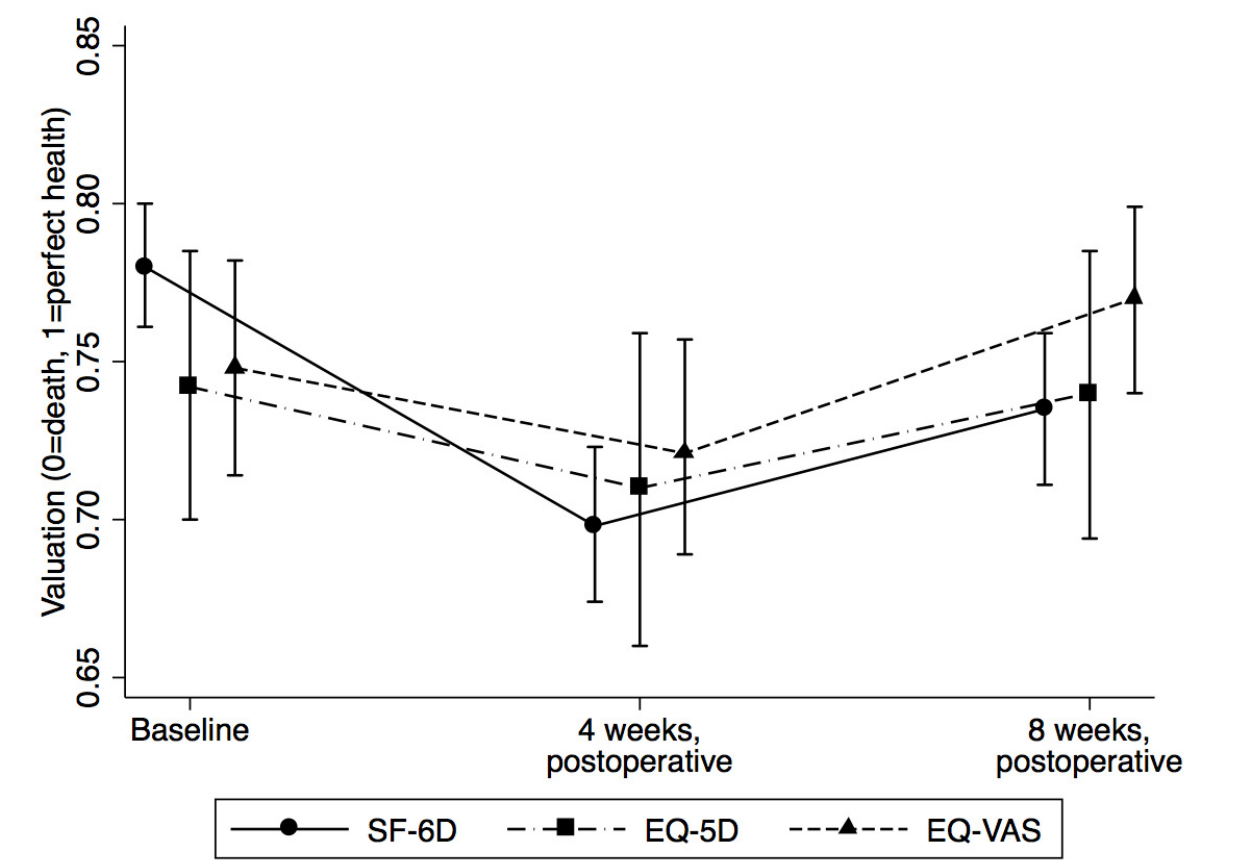
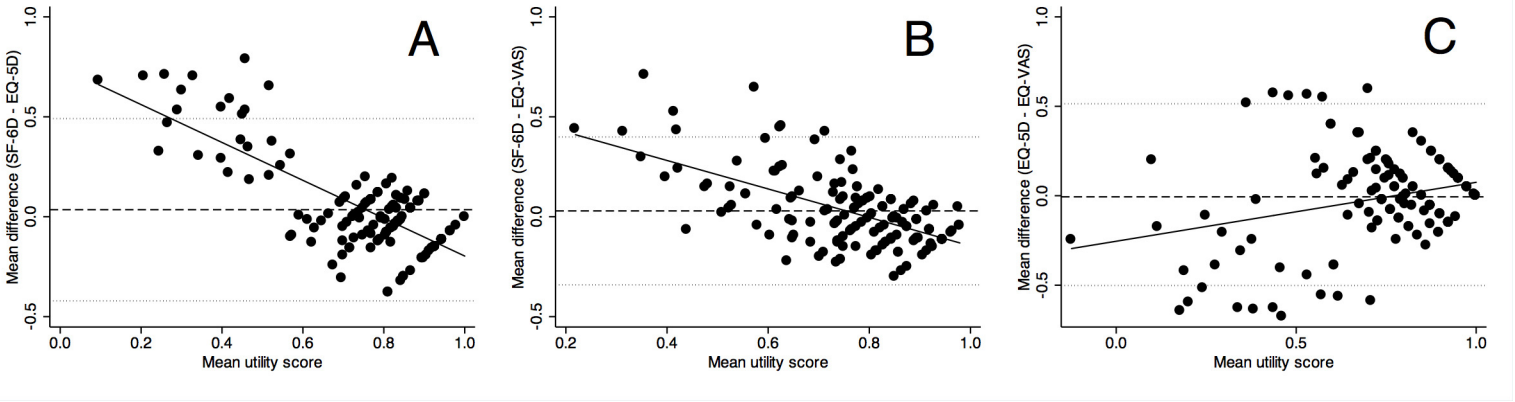


Figure 2 – Bland-Altman plots demonstrating the agreement ± 2SD between A) SF-6D and EQ-5D; B) SF-6D and EQ-VAS; and C) EQ-5D and EQ-VAS at baseline. The solid line represent the regression line between the mean difference and mean utility score (a line closer to horizontal centered at 0.0 denotes perfect agreement), and the dotted lines represent the mean difference (--), and upper and lower confidence interval limits (··).



7. ECONOMIC IMPACT OF AN ENHANCED RECOVERY AFTER SURGERY PATHWAY FOR ESOPHAGECTOMY

7.1 PREAMBLE

Chapter 2 systematically reviewed all economic evaluations of enhanced recovery pathways specifically within elective colorectal surgery, and reported that the literature tended to associate enhanced recovery pathways with cost savings, although the quality of the evidence was poor. While the majority of the literature supporting the clinical and cost effectiveness of enhanced recovery pathways has been performed in the context of colorectal surgery, the fundamentals of enhanced recovery may be applied to other procedures as well. There are increasing data on the benefits in decreasing duration of hospitalization and postoperative complications in orthopedic[159], urologic[160], gynecologic[161], upper gastrointestinal[13,162,163], and hepatobiliary surgery[12,164,165]. However few studies have performed an adequate cost-analysis, and no studies have performed a formal cost-effectiveness analysis of enhanced recovery pathways in non-colorectal surgery.[164]

The following manuscript investigates the economic impact of an enhanced recovery pathway for a high acuity procedure, esophagectomy, by performing a cost analysis using deviation-based cost modeling. Deviation-based cost modeling was specifically developed to analyze the clinical and economic impacts of pathways.[166] In short, patients are classified into four ‘deviation groups’, which are defined using a combination of length of stay and the severity of the complication experienced. The mean cost of each deviation group is then weighted into a single summary measure, which represents the relative contribution of each of the deviation groups. This methodology better reflects the fact that enhanced recovery pathways, and clinical pathways in general, are likely to improve outcomes in patients without complications, or those that experience minor complications only, and do not affect those patients that experience a

major complication, such as an anastomotic leak. This study also served to determine whether enhanced recovery pathways were associated with economic benefits in our local practice environment. As mentioned in the *Introduction*, postoperative complications are the main cost drivers for surgical hospital admissions. The original manuscript comparing clinical outcomes did not report a decrease in complications after implementation of an enhanced recovery pathway for esophagectomy, despite a reduction in length of stay.[167] In the setting of a high-acuity procedure, it was unclear whether an enhanced recovery pathway would be associated with any cost savings without a decrease in the incidence of postoperative complications. This study was also performed to provide an effect estimate for power calculations for a future cost-effectiveness study comparing enhanced recovery to conventional perioperative management in patients undergoing elective colorectal surgery. The cost analysis in this study was originally reported in Euros, as per journal requirements. For the ease of interpretation, the values have been converted back to Canadian dollars. This manuscript was published in the *British Journal of Surgery* (Br J Surg 2013; 100(10): 1326-34).

7.2 ECONOMIC IMPACT OF AN ENHANCED RECOVERY AFTER SURGERY PATHWAY FOR ESOPHAGECTOMY

[1]Lawrence Lee, [1]Chao Li, [1]Nicolas Robert, [2]Eric Latimer, [3]Franco Carli, [4]David S. Mulder, [1]Gerald M. Fried, [1,4]Lorenzo E. Ferri, [1]Liane S. Feldman

[1]*Steinberg-Bernstein Centre for Minimally Invasive Surgery and Innovation, McGill University Health Centre, Montreal, Qc*
[2]*Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Qc*
[3]*Department of Anaesthesia, McGill University Health Centre, Montreal, Qc*
[4]*Division of Thoracic Surgery, McGill University Health Centre, Montreal, Qc*

Corresponding author:
Liane S. Feldman, MD FRCSC FACS
1650 Cedar Ave, L9-303
Montreal, Quebec
H3G 1A4, Canada
Telephone: (514) 934-1934 ext. 44004
Fax: (514) 934-8210
Email: liane.feldman@mcgill.ca

Funding support: Drs. Lee and Li are supported by scholarships from the Quebec Health Sciences Funding Agency (FRQ-S) and the McGill Surgeon Scientist program.

Category: Original article

Presentation information: Data were presented at *Digestive Diseases Week 2013* (Orlando, FL) as a poster presentation on May 18-21, 2013.

ABSTRACT

Background: Data are lacking to support the cost-effectiveness of enhanced recovery pathways for esophagectomy. The aim of this study was to investigate the impact of an enhanced recovery pathway (ERP) on medical costs for esophagectomy.

Methods: This study investigated all patients undergoing elective esophagectomy from June 2009 to December 2011 at a single high-volume university hospital. From June 2010, all patients were enrolled in an ERP. Clinical outcomes were recorded up to 30 days. Deviation-based cost modeling, was used to compare costs between the traditional care and ERP groups.

Results: A total of 106 patients were included (47 traditional care, 59 ERP). There were no differences in patient, pathologic and operative characteristics between groups. Median LOS was lower in the ERP group (traditional care 10 [IQR 9-18] vs. ERP 8 [IQR 7-18] days, p=0.019). There was no difference in 30-day complications (traditional care 62% vs. ERP 59%, p=0.803), and 30-day or in-hospital mortality was low (4%, 4/106). The costs of the on-course and minor deviation groups were significantly lower after implementation of the ERP. Pathway dependent cost saving per patient was CAN$1397. The overall cost saving per patient was CAN$2666. One-way sensitivity analysis demonstrated that the ERP was cost-neutral or more costly only at extremes values of ward, operative and intensive care costs.

Conclusions: A multidisciplinary ERP for esophagectomy was associated with cost-savings without increase in morbidity or mortality.

BACKGROUND

In 2010, over 1.3 billion US$ was spent on treatments for oesophageal cancer in the United States.[168] Despite advances in multimodal therapy, surgical resection remains the mainstay of curative therapy. Oesophagectomy is associated with significant morbidity and mortality[169], even in high-volume centres[170,171] and while patient factors contribute significantly to adverse events and increased costs[172]; structures and processes of care are important factors that further influence surgical outcomes.[173] Failures in structures or processes of care were found to be a major contributing cause in almost a third of all cases of morbidity and mortality after surgery.[174]

Standardized surgical care pathways can improve quality and efficiency of care. Enhanced recovery pathways (ERP) are multi-disciplinary care pathways that incorporate multiple evidence-based interventions that minimize surgical stress, hasten recovery and, improve outcomes.[175] ERPs have been demonstrated to decrease length of stay and complications after colorectal[10], orthopedic[176], and gynecologic[161] surgery, amongst others. In patients undergoing oesophagectomy, moderate morbidity and low mortality was seen in a large series managed by ERP[177]. Two other observational studies have reported lower complication rates and length of stay after ERP compared to traditional management.[178,179]

The present authors have previously reported that an ERP for oesophagectomy was associated with shorter hospital stay, despite no differences in postoperative complication rates.[167] Adoption of ERPs has been poor, as there are many barriers to change.[14] Successful implementation of new evidence into practice requires comprehensive approaches at many levels.[180] Decision-makers and health care purchasers require information about cost when deciding whether to adopt new quality initiatives, and data have been lacking to support the cost-

effectiveness of ERPsfor complex procedures. The aim of this study, therefore, was to investigate the impact of ERP on medical costs for oesophagectomy.

METHODS

This study involved a cost analysis of previously reported data.[167] All patients undergoing elective oesophagectomy for malignancy or high-grade dysplasia at a single high-volume university-affiliated institution from June 2009 to December 2011 were identified from a prospective database. Patients who underwent pharyngo-laryngo-esophagectomy or where oesophagectomy was performed for benign disease or as an emergency were excluded. Two experienced oesophageal surgeons performed all procedures. Feeding jejunostomies were rarely used. Starting in June 2010, all patients were enrolled in a 7-day multidisciplinary ERP incorporating printed patient education materials, structured daily care plans with indications for intensive care admission (intensive care admission immediately after surgery was not routine), early structured mobilization, diet and drain management (Table 1). No patient was managed by ERP prior to June 2010. Patients were divided into traditional care (pre-implementation) and ERP (post-implementation) groups. Clinical outcomes were recorded up to 30-day after surgery. Morbidity and mortality were classified using the Clavien-Dindo system[108], modified for thoracic surgery.[181] As admission to the intensive care unit immediately after surgery was routine prior to the implementation of the ERP, and therefore did not represent an adverse event, these admissions were not defined as a Clavien-IV complication in the traditional care group. Length of stay (LOS) data included duration of primary hospitalisation as well as any readmissions within 30 days of surgery.

The economic evaluation was performed from an institutional perspective, for up to 30 days. Medical costs were calculated by micro-costing, in which the frequencies of each resource consumed were recorded, and multiplied by their respective unit cost to generate the total medical costs. Unit costs were supplied by the hospital finance and accounting department, or were derived by dividing the total direct costs by the total output (e.g. number of patient bed-days occupied for a specific ward; figures obtained from the provincial health ministry[182]) if not directly available. All unit costs were adjusted to include overhead (e.g. administration, information technology, housekeeping, etc.), which was distributed using the direct allocation method. Physician fees were not included, as they are not paid by the hospital within the Canadian health care system. All costs were adjusted to 2011 Canadian dollars using the real health care inflation rate, which is specific for health care services and adjusted for population growth[183]**,** and converted to Euros (1 CAN$ = 0.76 Euros (€) using the exchange rate on December 31, 2011[184]).

The economic impact of ERP was analyzed using deviation-based cost modeling, a validated method to compare the clinical and economic impact of clinical pathways.[185] In this methodology, patients are classified into four deviation groups based on length of stay and postoperative morbidity (Table 2). Therefore, deviations represent departures from the expected hospital course. Deviation-based cost modeling more accurately describes the impact of complications on hospital course, compared to a before-and-after analysis. In its original format, median costs were calculated for each deviation group, and weighted according to the relative proportion of each deviation to provide the weighted-average median cost. The procedure was modified to calculate a weighted-*mean* cost, rather than the median, as the mean is a more important summary measure when assessing budgetary impact.[59] This modification makes the

deviation-based cost model structure more like weighted averaging in decision-analytic models (Figure 1). Overall expected cost savings were defined as the difference in the weighted-average mean costs between the traditional care and ERP groups. Pathway-dependent costs, which describe the impact of patients that deviate off-course, were defined as the weighted-average mean costs of the on-course and minor deviation groups. The effect of the ERP on resource utilization other than hospital bed days was also investigated.

Univariable comparisons between the traditional care and ERP groups were performed using Chi-squared or Fisher's exact tests for categorical variables and Student's t- or Mann-Whitney U-tests for continuous variables. Due to the extreme right-skewedness of cost data, confidence intervals around the mean differences between groups were derived from bootstrap estimates (10,000 iterations) taken at the 2.5^{th} and 97.5^{th} percentiles. One-way sensitivity analyses were performed by varying operative, intensive care, and ward costs across a wide range of values in the ERP group only (while keeping costs constant in the traditional care group), and examining its effect on the weighted-mean cost. Operative costs were included in sensitivity analyses to account for the additional costs of the preoperative visits (as the cost of the preoperative clinic cost was included as in the derivation of operative costs) and also for any potential differences in operative management over time (such as learning curve) that may affect costs. All statistical analyses were performed using STATA 12 (StataCorp, College Station, TX).

RESULTS

A total of 106 patients were included for analysis (47 traditional care, 59 ERP). The results of the original trial have been previously reported[167], but in brief, there were no differences in patient, pathological or operative characteristics between the two groups. Median

LOS was lower in the ERP group (ERP 8 [7-18] vs. traditional care 10 [9-18] days, p=0.019), with no differences in 30-day complications (ERP 59% vs. traditional care 62%, p=0.803) or readmissions (ERP 3% vs. traditional care 4%, p=1.000) between the groups. Overall 30-day or in-hospital mortality was low (4%, 4/106).

There was no overall difference in deviation case-mix between the traditional care and ERP groups (Table 3). However, median LOS was lower in the ERP group for patients who experienced no deviation or minor deviation (traditional care 9 [IQR 8-11] vs. ERP 8 [IQR 7-8] days, p<0.001), while no differences were found for patients who experienced moderate or major deviations (traditional care 20.5 [IQR 18-34.5] vs. ERP 21.5 [18-29] days, p=0.863).

The weighted-average mean cost, or expected cost, for the traditional care and ERP groups was CAN$22835 and CAN$20169, respectively, resulting in an overall expected cost saving of CAN$2666 per patient for the ERP group. There were significantly lower mean costs for the ERP group for patients with no deviations or minor deviations (Table 3), which resulted in pathway-dependent cost savings of CAN$1397 per patient. No differences in mean costs were found for patients with moderate or major deviations. The ERP had no effect on median LOS, postoperative complications, or costs after minimally invasive esophagectomy compared to open approach (eTable 1).

The mean cost of non-hotel resource utilization (radiology, pharmacy, and laboratory) was lower in the ERP compared to traditional care group for patients with similar LOS in the no or minor deviation groups (mean difference -459 CAN$, 95% CI -806, -191). There was no difference in non-hotel resource utilization costs between the ERP and traditional care group for patients that experienced moderate or major deviations (mean difference 2998 CAN$, 95% CI -

1927, 8584). All sensitivity analyses demonstrated that only extreme variations in cost would negate the expected cost savings (Figure 2).

DISCUSSION

There is increasing evidence to support the clinical effectiveness of ERPs for perioperative management. Meta-analyses of multiple randomized clinical trials in colorectal surgery have demonstrated decreased hospitalisation and a lower incidence of postoperative complications compared to traditional care.[10,27,186] A systematic review of economic evaluations of ERP in colorectal surgery noted that many studies demonstrated cost savings for ERP both in North America and Europe.[22] Unlike colorectal surgery, much of the evidence ERP after oesophagectomy is not derived from randomized clinical trials, but rather based observational trials or extrapolated from colorectal surgery.[187] The ERP used in the present study incorporated many elements considered essential to enhanced recovery management, including thoracic epidural analgesia[188,189], intraoperative fluid restriction[190] and early enforced mobilization.[8,191,192] Unrestricted oral intake was delayed until confirmation of anastomotic integrity on the fifth postoperative day, and while earlier oral intake does not cause increased morbidity after major upper gastrointestinal surgery[193], the clinical benefits are equivocal[194,195]. On the other hand, the present ERP did not contain elements that may further decrease surgical stress, enhance recovery and improve outcomes, such as protective ventilation[196] or preoperative carbohydrate loading[197].

A cost analysis undertaken by Zehr et al.[198] reported a 34% decrease in hospital charges after implementation of a clinical pathway. The present study did not report cost savings of the same magnitude (12% decrease), but dissimilarities in the comparison group (i.e. traditional care), as well as secular trends, between the two studies may partly explain this difference, as the

earlier study. was performed between 1991 and 1997. US national trends from the Healthcare Cost and Utilization Project have reported decreasing mean length of stay over time for oesophagectomy (using ICD-9 code 42.4x)[199] and there is evidence of falling risk-adjusted mortality for oesophagectomy.[169]

Many previous cost analysis studies that have compared ERP to traditional care in other procedures have used a before-and-after design.[22,200] Major limitations of this approach are the inability to describe the variable impact of complications after implementation of a pathway, and to account for patients with moderate or severe complications that may deviate from the pathway.[185] In the present study, deviation-based cost modeling was used to overcome some of the limitations of the traditional before-and-after design. This approach more accurately describes the effect of the pathway on hospital course among patients with varying severity of morbidity. In addition, by categorizing clinical course according to deviation groups, the major beneficial effect of clinical pathways can be characterized.[185] Improvements in outcomes for patients who experience severe complications (Clavien III-IV) are likely due to factors beyond the pathway, such as general improvements over time. These patients often represent the tail end of the cost distribution, which are disproportionately represented when non-weighted mean costs are calculated. In comparison, dividing patients into deviation groups can identify in which patients the ERP was most beneficial. In the present study, patients in the ERP group who experienced no complications, or Clavien I-II complications had a significantly lower length of stay compared to the patients who experienced the same complications in the traditional care group. This translated into significantly lower costs in patients with no deviations or minor deviation managed by ERP compared to those managed by traditional care, while no difference

in mean costs was found for patients who experienced a moderate or major deviation between the two groups.

The ERP was associated with an overall expected cost savings of 2013 € per patient, which included pathway-dependent cost savings of 1055 €. There are several reasons for this cost difference. Decreased length of stay in the ERP group clearly resulted in lower costs. However, among patients with similar length of stay, there were lower costs for patients who experienced minor complications (Clavien I-II), suggesting that the ERP lessened the impact of these complications. Given that postoperative morbidity is a substantial driver for increased hospital costs[201], this effect is important. One explanation for this finding may be that resource utilization was lower for patients managed by ERP, as this process encouraged minimal laboratory and radiological investigations.

Only 23 patients in the entire cohort underwent minimally invasive oesophagectomy and there were no differences in clinical outcomes for patients undergoing this approach between the traditional care or ERP groups. While there may be further benefit by using a minimally invasive approach within an ERP, as has been shown in colorectal surgery[28], the small number of patients undergoing this approach means that no inference can be drawn regarding either clinical or economic benefit.

This study has a number of limitations. Although the majority of the data were obtained from a prospectively maintained database, some information related to resource utilization was collected afterwards, from comprehensive electronic medical records. As an observational study, there may have been selection bias, although this may be in part mitigated by the fact that all patients were enrolled in the ERP after implementation in June 2010, rather than selection of patients that were most likely to succeed on the pathway. There were no changes in patient

referral patterns during the study and no differences in patient characteristics that might affect outcomes, such as co-morbid status and disease stage, between the two groups. Compliance with the ERP was not available, although failure must be interpreted with caution, as many ERP processes are dependent on the clinical state of the patient. Non-compliance to certain elements may represent deviations in clinical course, rather than true failures to comply.

The present analysis was also limited to an institutional perspective with a time horizon of 30 days. It is unknown whether increasing time or including indirect costs (i.e. different perspectives) would affect the overall results. Although there were no differences in proportion of patients that were discharged to rehabilitation or extended-care facilities between the two groups, it is not known if inclusion of these costs (using a health-care system perspective) would affect the overall conclusions. Development, maintenance and implementation costs were not included. Given the significant time and effort required to develop these pathways, opportunity costs may be significant.[202] Patient-reported outcomes or quality-adjusted life years were also not measured, therefore a cost-utility analysis could not be performed. Future studies should measure these outcomes. Finally, the costs used for the analysis were derived from institutional and provincial financial data, which may not be generalizable to other practice settings or health care systems. This limitation is common to most, if not all, economic evaluations. By varying the important contributing costs across a wide range of values the results of the present study might be generalized to a wide range of other practice settings.

The present data support the need for further economic evaluation of ERPs and provide financial incentives for decision-makers to adopt this approach for patients undergoing oesophagectomy.

Table 1 – Comparison of traditional and Enhanced Recovery perioperative management for esophagectomy

Traditional Care	**Enhanced Recovery Pathway**
Preoperative phase	
- Medical evaluation, medical and anaesthesia consultation at the discretion of the surgeon	- Medical evaluation, including automatic anaesthesia consultation if clinically relevant comorbidities present in order to avoid intensive care postoperatively - Preoperative education regarding preoperative optimization, daily in-hospital treatment goals, expected discharge date - Preoperative optimization including smoking cessation counseling, respiratory muscle strengthening - Assessment of potential difficulties with target discharge date, if present automatic referral to discharge planning specialist
Intraoperative phase	
Anaesthesia - Insert thoracic epidural catheter (bupivacaine 0.25%) - Fluid management as per treating anaesthesiologist - Extubation in the operating room at the operative team's discretion, routine intensive care admission if intubated **Surgery** - Antibiotic and DVT prophylaxis - Tailored surgical approach according to patient's performance status and location and extent of neoplasm	**Anaesthesia** - Prophylactic antiemetics for PONV - Insert thoracic epidural catheter (bupivacaine 0.25%) - Avoid fluid overload (crystalloid infusion 4-6 mL/kg/hr) - Maintain normothermia - Extubation in the operating room, observation in the post-anaesthesia care unit for 6 hours, then transfer to the surgical ward **Surgery** - Antibiotic and DVT prophylaxis - Minimally-invasive approach encouraged for HGD or early (cT1-2, N0) cancer - Tailored surgical approach according to patient's performance status and location and extent of neoplasm - Avoid blood loss; blood product transfusion only if blood loss > 20% circulating blood volume
Postoperative phase	
Standardized pre-printed orders - Barium esophagram on POD7 - Start oral intake only after confirmation of anastomotic integrity on barium	**Daily treatment plan** Day of surgery: NPO; ensure patient sits in chair, use incentive spirometry every hour when awake

esophagram - Remove tube thoracostomy only after solid diet started - Thoracic epidural analgesia until removal of chest tube, multimodal analgesia (acetaminophen and COX-2 inhibitors) once epidural removed - Discharge at the surgeon's discretion when clinically indicated	POD1: NPO; ensure patient sits in chair for at least 30 minutes, ambulate ½ length of hallway (35 metres) twice, use incentive spirometry every hour when awake POD2: NPO; ensure patient sits in chair for at least 30 minutes, ambulate ½ length of hallway (35 metres) three times, use incentive spirometry every hour when awake; remove urinary drainage catheter, remove chest tube if cervical anastomosis POD3: Sips of water allowed; ensure patient sits in chair for at least 60 minutes and ambulates full length of hallway (70 metres) three times, use incentive spirometry every hour when awake; remove nasogastric tube POD5: Sips of water allowed; ensure patient sits in chair for all meals and ambulates full length of hallway (70 metres) five times, use incentive spirometry every hour when awake; remove chest tube if thoracic anastomosis, remove epidural and start oral multimodal analgesia (acetaminophen and COX-2 inhibitors), stop intravenous fluid infusion; barium esophagram performed POD6: Diet as tolerated started if anastomotic integrity confirmed on barium esophagram; ensure patient sits in chair for all meals and ambulates full length of hallway (70 metres) five times, use incentive spirometry every hour when awake; remove remaining cervical and abdominal drains; ADL assessment POD7: Anticipated discharge date if discharge criteria* met; review postoperative care instructions; 3-week postoperative follow-up scheduled

POD = postoperative day; NPO = nil per os; ADL = activities of daily living
*Discharge criteria include ability to tolerate solid oral intake and to perform basic ADLs independently, and absence of fever, pain, nausea and vomiting.

Table 2 – Definitions of deviation-based cost modeling groups (adapted from Vanounou et al. 2007[185])

Deviation	Hospital course	Clinical impact	Examples
On-course	LOS ≤ 50th percentile	No complications Clavien I and II	Uncomplicated course, or surgical site infection treated by incision and drainage with no effect on LOS
Minor	LOS = 50th to 75th percentile	No complications Clavien I and II	Paralytic ileus of four day duration treated by nasogastric tube drainage only
Moderate	LOS > 75th percentile Any hospital duration	Clavien I and II Clavien IIIa (requiring surgical, radiological or endoscopic intervention not under general anesthesia)	Pneumonia requiring extended length of stay Anastomotic leak requiring reoperation but no admission to ICU
Major	Any hospital duration	Clavien IIIb-V (requiring surgical, radiological or endoscopic intervention under general anesthesia; or death of a patient)	Septic shock secondary to anastomotic leak requiring ICU management

LOS = length of stay; ICU = intensive care unit

Table 3 – Clinical and economic impact of an enhanced recovery pathway for esophagectomy using deviation-based cost modeling

	Traditional Care (n=47)	**Enhanced Recovery Pathway (n=59)**	***p*-value/ mean difference (95% CI)***
Deviation mix, % (n)			0.573
On-course	47% (22)	56% (33)	
Minor deviation	19% (9)	13% (8)	
Moderate deviation	15% (7)	12% (7)	
Major deviation	19% (9)	19% (11)	
Mean cost, € (95% CI)*			
On course	12533 (11629, 13411)	11137 (11243, 12315)	- 742 (-1784, -335)
Minor	18700 (15870, 21872)	13925 (13168, 16204)	-4120 (-7834, -803)
Moderate	19936 (17244, 22362)	24780 (21392, 29476)	5497 (-379, 11358)
Major	54410 (30611, 85699)	48153 (28202, 76086)	-5242 (-44345, 33689)
Weighted-mean cost†, CAN$	22835	20169	
Pathway dependent cost savings‡, CAN$	1397		
Overall cost savings, CAN$	2666		

LOS = length of stay

*95% confidence intervals derived from bootstrap estimates (10,000 iterations), taken at the 2.5th and 97.5th percentiles. Results are statistically insignificant if the 95% confidence interval of the mean difference crosses 0.

†Weighted average for all deviation groups

‡Weighed average for no and minor deviation groups only

eTable 1 – Comparison of 30-day clinical and economic outcomes for minimally invasive oesophagectomy managed by traditional care versus enhanced recovery pathway. Deviation-based cost modeling was not applied given the small sample sizes, therefore traditional analyses were performed.

	Traditional Care (n=10)	**Enhanced Recovery Pathway (n=14)**	***p*-value/ mean difference (95% CI)***
Median LOS, days [IQR]	14.5 [8-19]	8 [7-25]	0.492
30-day complications	6 (60%)	8 (57%)	0.611
30-day readmissions	0 (0%)	1 (7%)	0.983
Mean cost, CAN$ (95% CI)	34063 (17011, 64748)	32715 (15452, 62386)	-1348 (-41315, 38053)

*95% confidence intervals derived from bootstrap estimates (10,000 iterations), taken at the 2.5th and 97.5th percentiles. Results are statistically insignificant if the 95% confidence interval of the mean difference crosses 0.

Figure 1 – Graphic representation of deviation-based cost modeling using a decision-analysis model framework. Square node represents a decision node, circle node represents a chance node, and triangles represent endpoints.

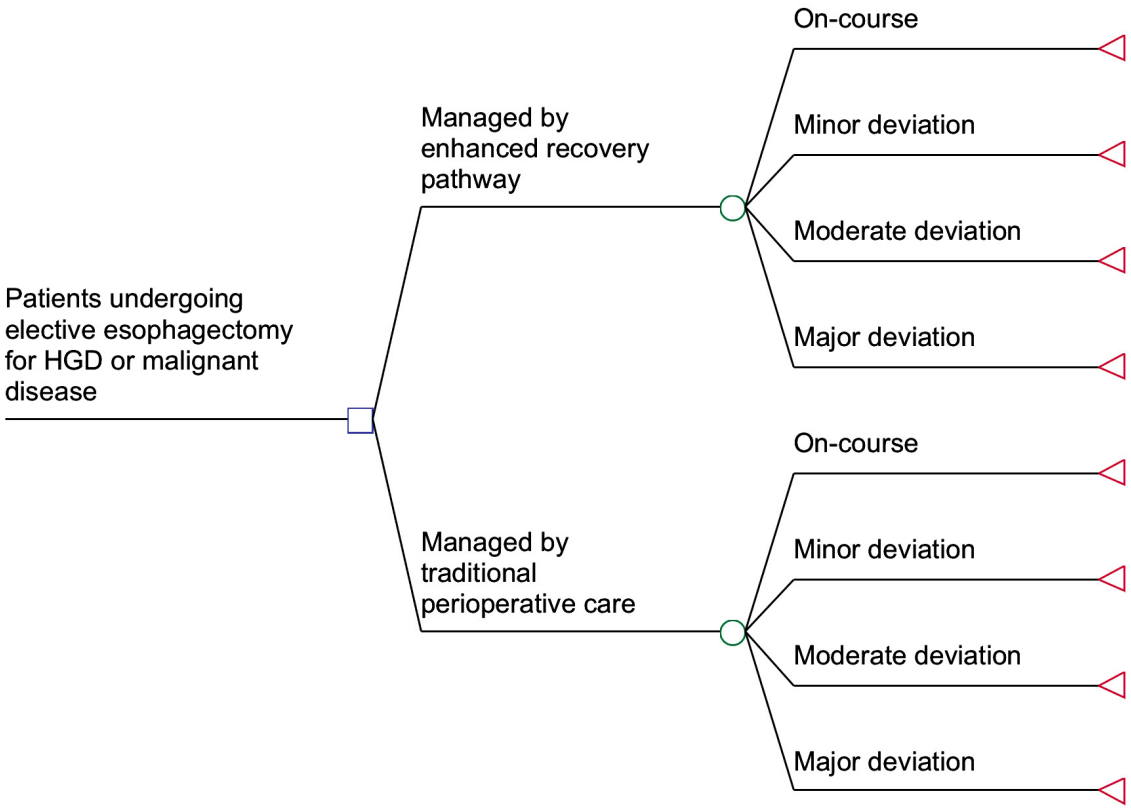
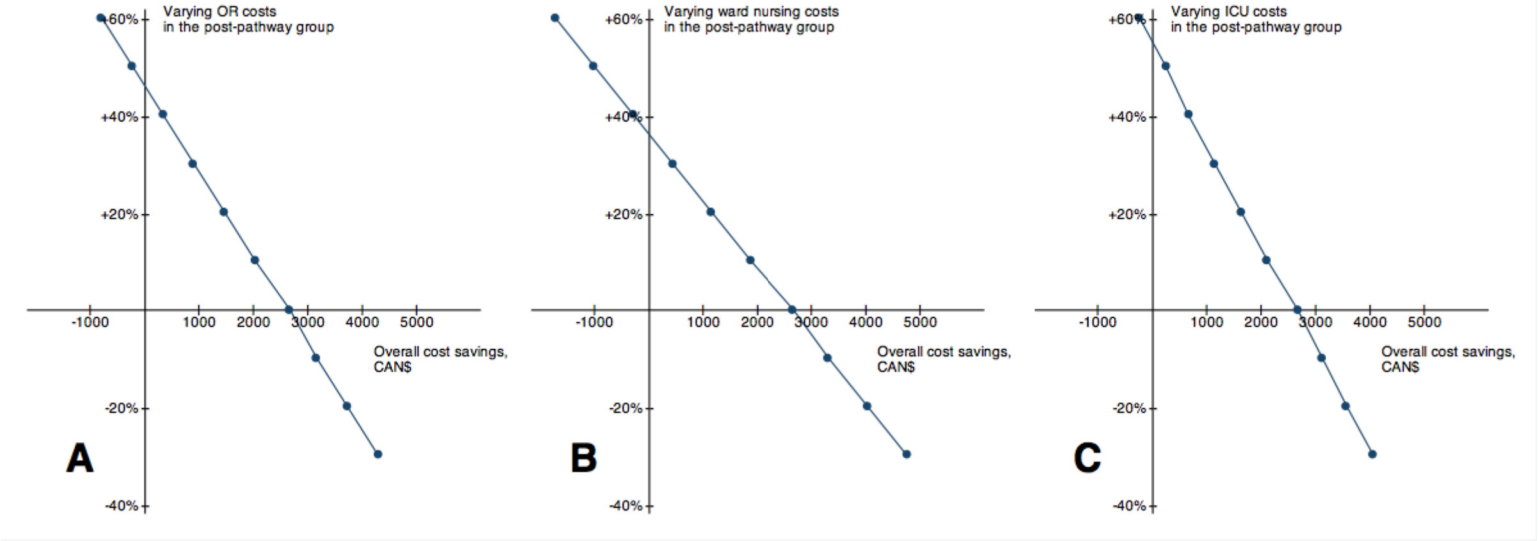


Figure 2 – One-way sensitivity analysis of key variables and their effects on overall cost savings. A) Varying operative (OR) costs will result in cost neutrality only when operative costs in the enhanced recovery pathway (ERP) group are 46% higher compared to traditional care operative costs. B) Varying ward nursing costs will result in cost neutrality only when ward nursing costs in the ERP group are 36% higher compared to traditional care nursing costs. C) Varying intensive care unit (ICU) costs will result in cost neutrality only when the daily ICU costs are 52% higher in the ERP group compared to the traditional care group.



8. COST-EFFECTIVENESS OF ENHANCED RECOVERY

8.1 PREAMBLE

The main objective of this thesis was to determine the cost-effectiveness of enhanced recovery versus conventional perioperative management in patients undergoing elective colorectal surgery. As reported previously in the literature review (Chapter 2), the economic evidence supporting enhanced recovery management for colorectal surgery is limited. As well, there are no economic evaluations performed within the Canadian setting. There are also few data to demonstrate the post-discharge outcomes and cost after ERAS.

All of the previous chapters have contributed to the design of the following study. Chapter 2 identified knowledge gaps in the economic literature. Chapters 3 provided a clear definition of postoperative recovery so that multi-attribute utility instruments could be psychometrically evaluated for the recovery construct. Chapters 5 and 6 provided evidence that the SF-6D was superior to the EQ-5D as a measure of postoperative recovery. Finally, Chapter 7 provided an effect estimate for sample size calculations. However, these sample size calculations were performed using the estimated difference in costs alone. Appendix 1 reports an estimated difference in QALYs based on applying the minimal clinically important difference of the SF-6D (± 0.041[117]) to the data from Chapter 5, and provides a sample size calculation specific to economic evaluations[203], taking into account the estimated differences of both costs and effectiveness (i.e. QALYs).

A cost-effectiveness analysis was performed using a multi-institutional prospective cohort study design. Eligible participants were recruited over a one-year period from two institutions in Montreal, Canada, of which one routinely utilized ERAS whereas the other did not. The analysis was performed from three different perspectives over a 60-day time horizon. The results show

that ERAS was associated with decreased costs from a societal perspective, with a statistically insignificant difference in quality-of-life over 60 days. ERAS was also highly likely to be cost-effective at all willingness-to-pay thresholds. These results were also robust to a wide range of sensitivity and subgroup analyses. An additional uncertainty analysis using the net monetary benefit (NMB) framework is reported in Appendix 2. This analysis also corroborated the subgroup analyses that were included in this manuscript. This manuscript was published in the *Annals of Surgery* (in press).

8.2 COST-EFFECTIVENESS OF ENHANCED RECOVERY VERSUS CONVENTIONAL PERIOPERATIVE MANAGEMENT FOR COLORECTAL SURGERY

[1]Lawrence Lee MD MSc, [1]Juan Mata MD, [2]Gabriela A Ghitulescu MD, [2]Marylise Boutros MD, [1]Patrick Charlebois MD, [1]Barry Stein MD, [1]A Sender Liberman MD, [1]Gerald M Fried MD, [2]Nancy Morin MD, [3]Franco Carli MD MPhil, [4]E Latimer PhD, [1]Liane S Feldman MD

[1]Steinberg-Bernstein Centre for Minimally Invasive Surgery and Innovation, Department of Surgery, McGill University Health Centre, Montreal, QC
[2]Division of Colorectal Surgery, Sir Mortimer B Davis Jewish General Hospital, Montreal, QC
[3]Department of Anaesthesia, McGill University Health Centre, Montreal, QC
[4]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, QC

Correspondence and requests for reprints to:
Lawrence Lee, MD MSc
1650 Cedar Ave E19-125
Montreal, QC H3G 1A4
Tel: (514) 934-2637
Fax: (514) 934-8475
Email: larry.lee@mail.mcgill.ca

Disclosures & Sources of financial support:
This study was funded by an investigator-initiated grant from Ethicon. The Steinberg-Bernstein Centre for Minimally Invasive Surgery and Innovation is supported by an unrestricted educational grant from Covidien. ASL is a paid speaker and a surgical proctor for Covidien. PC is on the advisory board for Merck. GMF receives research and fellowship funding from Covidien. LSF receives research and fellowship funding from Covidien and research support from Ethicon. LL, JM, GAG, MB, BS, NM, FC, and EL have no conflicts of interests to declare.

Running head:
Cost-effectiveness of ERP

STRUCTURED ABSTRACT

Objective: To determine the cost-effectiveness of enhanced recovery pathways (ERP) versus conventional care (CC) for patients undergoing elective colorectal surgery.

Summary Background Data: ERPs for colorectal surgery are clinically effective, but their cost-effectiveness is unknown.

Methods: A multi-institutional prospective cohort cost-effectiveness analysis was performed. Adult patients undergoing elective colorectal resection at two university-affiliated institutions from 10/2012-10/2013 were enrolled. One centre used an ERP, the other did not. Postoperative outcomes were recorded up to 60-days. Total costs were reported in 2013 Canadian dollars. Effectiveness was measured using the SF-6D, a health utility measure validated for postoperative recovery. Uncertainty was expressed using bootstrapped estimates (10,000 repetitions).

Results: A total of 180 patients were included (95 CC, 95 ERP). There were no differences in patient characteristics except for a higher proportion of laparoscopy in the ERP group. Mean length of stay was shorter in the ERP group (6.5 vs. 9.8 days, p=0.017) but there were no differences in complications or readmissions. Patients in the ERP group returned to work quicker and had less caregiver burden. There was no difference in quality of life between the two groups. The cost of the ERP program was $153 per patient. Overall societal costs were lower in the ERP group (mean difference -2985, 95% CI -5753, -373). The ERP had a greater than 99% probability of cost-effectiveness. The results were insensitive to a range of assumptions and subgroups.

Conclusion: Enhanced recovery is cost-effective compared to conventional perioperative management for elective colorectal resection.

INTRODUCTION

Recent improvements in anaesthesia, pain control, and minimally invasive surgery have all contributed to improved postoperative outcomes after colorectal surgery, but the effect of each element is limited if performed in isolation.[3] Enhanced recovery pathways (ERP) incorporate numerous evidence-based elements from all perioperative phases into a single multidisciplinary care pathway aimed at decreasing the surgical stress response and accelerating recovery after surgery.[25] While specific elements of an enhanced recovery pathway may vary between procedures and practice settings, the principal elements of an enhanced recovery pathway includes preoperative patient education and preparation for surgery; attenuation of the surgical stress response, pain, and postoperative nausea and vomiting through anaesthetic, analgesic and surgical techniques; and aggressive early postoperative mobilization, enteral feeding, and avoidance or early removal of drains and tubes.[9]

ERPs have been shown to improve short-term physiologic outcomes, and decrease hospitalization and complications after colorectal surgery.[10,11] Despite these promising results, ERPs have yet to be widely adopted.[14] As healthcare spending continues to grow[17], the cost savings associated with ERPs may further encourage their adoption. It is commonly accepted that enhanced recovery pathways are associated with reduced costs by virtue of shorter duration of hospitalization and decreased complications[16], but these touted economic benefits have not been supported by rigorous evaluation, as the currently available evidence is limited.[204] Furthermore, few studies comparing ERP to conventional perioperative management have investigated implementation and maintenance costs, post-discharge patient-reported outcomes, and socioeconomic impact, which may significantly affect the cost-effectiveness of ERPs.[204] Prior to widespread implementation of these pathways, a formal economic evaluation assessing

both the costs and benefits of these pathways is required in order to determine their cost-effectiveness. Therefore the main objective of this study is to determine the cost-effectiveness of enhanced recovery versus traditional perioperative management for patients undergoing elective colorectal surgery.

METHODS

A cost-effectiveness analysis was performed using a multi-institutional prospective cohort study design. Adult patients undergoing elective colorectal surgery from 10/2012 to 10/2013 at two university-affiliated tertiary teaching institutions located in Montreal, Canada were eligible. One centre managed all patients undergoing colorectal surgery using an ERP, while the other did not. Both institutions were large teritary teaching centres with fellowship-trained colorectal surgeons. A comparison of perioperative management strategies between conventional care and ERP is shown in Table 1. Patients were ineligible if they only had a small bowel procedure planned, were undergoing a synchronous procedure with their colorectal resection, did not speak English or French, lived outside the Montreal metropolitain area, or had a neuropsychiatric condition that precluded completion of the study questionnaires. All eligible patients were approached for participation at their preoperative clinic visit. The study protocol was approved by the ethics review board at both institutions.

Patients were enrolled at their preoperative clinic visit (within two weeks before surgery). Patient demographic, operative characteristics, and clinical outcomes were prospectively recorded by two evaluators unaware of the study hypothesis. All outcomes were collected up to 60 days after the initial surgery. Postoperative complications were evaluated using the

Comprehensive Complication Index, a 0 to 100 scale incorporating the number and severity of complications that occur in a single patient. [205]

Health Outcomes

All participating patients completed the study questionnaires at enrolment (within 2 weeks before surgery), and at four- and eight-weeks after surgery. Health-related quality of life was evaluated using the SF-6D, a health utility index that has been previously validated for recovery after colorectal surgery.[91] It has been shown to be more responsive to the expected postoperative changes than other instruments, such as the EQ-5D.[206] The SF-6D is derived from 11 items of the SF-36, and covers six dimensions: physical functioning, role limitations, social functioning, pain, mental health, and vitality. Each dimension has four to six levels, and responses are weighted according to an algorithm to obtain a utility score ranging from 0.296 (worst level in all dimensions) to 1.000 (best level in all dimensions). Therefore the relative importance of each of the dimensions can be expressed on a single summary measure. The SF-6D has been demonstrated to be more responsive to the construct of postoperative recovery than other utility instruments, such as the EQ-5D.[206] The area under the curve method (using the trapezium rule) was used to calculate quality-adjusted life years (QALY).[207] The mean difference in QALYs was adjusted for the baseline SF-6D scores.[208]

Costs

Medical costs were calculated by micro-costing, in which the frequency of resource consumption is multiplied by their unit costs. Unit costs were supplied by the respective hospital finance departments or were derived from provincial health ministry records.[209] All unit costs

were adjusted to include overhead costs (e.g. administration, housekeeping, sterilization services, etc.), which were allocated by the stepdown method (see Table, Supplemental Digital Content 1, which reports important unit costs from both institutions). ERP program costs, which includes the development, implementation, and audit of the ERP, were calculated by dividing the cost of the yearly multidisciplinary steering group by the total number of patients managed by ERP over the study period. Physician fees were obtained from the provincial fee schedule.[210]

Outpatient resource utilization, out-of-pocket expenses, caregiver burden, and productivity losses were measured using a questionnaire designed for the Canadian setting,.[211,212] Utilization of outpatient services was approximated by asking the patient to recall the number of post-discharge visits to community health service centers (local health centers where patients can receive outpatient nursing and other basic medical care), surgeon and other specialist follow-ups (excluding readmissions). Out-of-pocket expenses were estimated by the patient and included medical (i.e. medications and medical supplies) and non-medical (e.g. hospital parking, hired homecare, etc.) related to the surgery. Caregiver burden was estimated by asking the patient if assistance was required from a non-compensated friend or relative in the postoperative period, and the time missed off work by this person, if applicable. The monetary value of caregiver burden was estimated based on the median Canadian salary (days missed from work × salary/365).[213] Productivity losses experienced by the patient was estimated based on the date of return to work, and the monetary amount was estimated based on the patient's yearly salary, which was recorded at enrollment. Patients who did not want to disclose their yearly salary were assigned the Canadian median salary for the analysis. All productivity losses were calculated from a societal perspective using the human capital approach.

Total costs were calculated from the institutional, healthcare system, and societal perspectives. Institutional perspective costs included medical costs borne by the hospital, which include operating room, ward, pharmacy, laboratory, radiology, emergency department, ambulatory clinics, allied health professional, and chemotherapy costs. Within the Canadian healthcare system, physician remuneration is paid by provincial health ministry and not by the institution. Therefore these costs were included in the total healthcare system costs. Other costs included in the healthcare system perspective included assisted-care facility (e.g., rehabilitation, convalescence, or long-term care facilities) and community health services centre costs. The societal perspective included productivity loss, caregiver burden, and out-of-pocket expenses. All costs were expressed in 2013 Canadian dollars (CAN $1= US $0.81, using purchasing power parity[214]).

Base Case and Sensitivity Analysis

Incremental costs from each perspective were calculated, and divided by the incremental QALYs to obtain the incremental cost-effectiveness ratios (ICER). Uncertainty around ICERs were assessed using bootstrap replications (10,000 repititions). In this method, random sampling with replacement was performed x times, where x = original sample size, for each group. Incremental costs and QALYs were calculated for each random sample, and repeated 10,000 times. Each of the bootstrapped ICERs were plotted on the cost-effectiveness plane. The probability of cost-effectiveness of the ERP at different willingness-to-pay thresholds (i.e. the amount a decision-maker is willing to pay to gain an additional QALY) was also calculated and displayed on cost-effectiveness acceptability curves.

Sensitivity analyses were also performed to test the robustness of the base-case results. Three different scenarios were analyzed: complete-case analysis (using only patients without missing data), uniform preoperative salary using the annual Canadian median personal income (for those previously employed), uniform unit costs, and friction valuation of productivity loss (assuming an elasticity of 0.8[215] and a friction period of 14.6 weeks[216]). Subgroup analyses were also performed to investigate the cost-effectiveness of the ERP based on employment status (employed and non-employed), approach (laparoscopic and open surgery), whether a new stoma was fashioned (stoma and no stoma), and postoperative complications (patients who experienced postoperative complications and those who did not).

Statistical Analysis

Sample size calculations were based on an analysis of the impact of an ERP on elective esophagectomy surgery at one of the study institutions.[217] In this study, the mean difference in total costs was $2571 (CAN$). This estimate compared favorably to the mean difference in costs after laparoscopic and open colorectal surgery at 12 weeks.[218] The pooled standard deviation in colorectal surgery was estimated to be $5000, given the incidence of severe postoperative complications after colorectal surgery and the high variability of the associated costs of treating these complications. These estimates result in an effect size of 0.51. Using a conservative estimate of an anticipated effect size of 0.4, the sample size required to detect an effect size of 0.4 at alpha = 0.05 and power = 0.8 is 100 patients per group. This sample size was adequate to account for variability in costs and effects (see Appendix 1).

Univariate comparisons were performed using Student's *t*- or Mann-Whitney U-tests for continuous data, and chi-squared test for categorical variables. Confidence intervals around

skewed cost data were obtained from the 2.5th and 97.5th percentile values from bootstrap replications. Missing data were handled with multiple imputation using chained equations (ten imputations). In this procedure, missing values are estimated using the appropriate regression model (truncated linear regression using the relevant lower and upper values for the SF-6D, and a lower limit of zero for costs) from other observed data, taking into account the longitudinal nature of the data, and repeated ten times to generate ten different imputed datasets. Final uncertainty around point estimates incorporate the between (datasets) and within (variable) variances, according to Rubin's rules.[112] All statistical analyses were performed using STATA 12 (StataCorp, College Station, TX).

RESULTS

A total of 190 patients (85 CC and 85 ERP) were included for the final analysis (Figure 1). Missing data was present in 11% (21/190) of patients overall (all eight-week assessments), with 9% (9/95) and 11% (12/95) missing in the CC and ERP groups, respectively. There were no differences in patient and operative characteristics, except for a higher proportion of patients undergoing a laparoscopic procedure and fewer receiving a stoma in the ERP group (Table 2). The ERP group had significantly better compliance to all enhanced recovery elements compared to CC, although several elements were already in use in the CC group (see Table, Supplemental Digital Content 2, which demonstrates adherence to enhance recovery elements between the two groups).

Clinical and post-discharge outcomes are reported in Table 3. Length of stay was shorter in the ERP group (mean difference: 3.3 days, 95% CI 0.6, 6.1; median 4 [IQR 3-7] vs. 7 [5-9], p<0.001). There were no differences in the incidence of 60-day complications, emergency

department visits, and readmissions, or in mean complication severity between the two groups. Postoperative primary ileus was the sole postoperative complication in eleven patients in the ERP group and five patients in the CC group. A total of three patients required admission to a rehabilitation or extended facility after primary hospitalization (two CC, one ERP). The proportion of patients returning to work by eight-weeks after surgery was similar between the two groups, but the patients in the ERP group experienced fewer days off work (mean difference: 9.0 days, 95% CI 0.7, 17.3; p=0.035). There was no difference in the mean number of postoperative days that patients required assistance with basic and independent activities of daily living, but caregivers experienced fewer days off work in the ERP group (mean difference: 3.7 days, 95% CI 1.2, 6.3; p=0.004). Patients in the ERP group also required fewer visits to community health service centers.

The total yearly cost of the enhanced recovery program, which is tasked with the design, implementation, audit, and maintenance of ERPs for multiple procedures, was $108,770 per year (see Table, Supplemental Digital Content 3, which reports the cost breakdown of the ERP program). During the study period, a total of 708 patients were managed by ERP, including 112 patients undergoing gastro-esophagectomy, 266 pulmonary resection (wedge resection and lobectomy), 85 prostatectomy, and 235 colorectal surgery (including 54 patients undergoing ileostomy closure), resulting in a mean enhanced recovery programme cost of $153 per patient.

Total and incremental costs comparing ERP and CC from the institutional, healthcare system, and societal perspectives are reported in Table 4. The ERP was associated with signficantly lower total societal costs, as productivity losses and caregiver burden were signficantly decreased compared to CC. There were no statistically significant differences in total costs from the institutional and healthcare system perspectives. Figure 2 compares the

evolution over time of quality of life between the two groups. Baselines SF-6D values were similar between the two groups (CC 0.775 (SD 0.152) vs. ERP 0.795 (SD 0.093), p=0.268). The difference in incremental improvement in quality of life over time between ERP and CC was not statistically significant (mean difference: +0.87 quality-adjusted days, 95% CI -1.23, 2.97; p=0.416). On base-case analysis, ERP was dominant (i.e. less costly and more effective) over CC in 78% of bootstrap estimates (Figure 3). Figure 4 demonstrates the probability that the ERP is cost-effectiveness at different willingness-to-pay (WTP) thresholds (i.e. the amount a decision maker is willing to pay to gain an additional unit of effectiveness). Even at conservative WTP thresholds of $0 and $50,000 per QALY, the ERP had a greater than 99% probability of cost-effectiveness

Across the five sensitivity analysis scenarios, the ERP remained dominant (Table 5) with a high probability of cost-effectiveness (Figure 4). The friction valuation of productivity loss resulted in the lowest probability of cost-effectiveness across the five sensitivity scenarios, but the overall probability remain greater than 90%. In the uniform unit costs scenario, the mean difference in costs from the institutional perspective was -2598 CAN$ (95% CI -6642, -1037) in favor of the ERP. Results from subgroup analyses also did not differ from the base case, as the ERP remained dominant. However, the probability of cost-effectiveness was lower for patients that were unemployed, undergoing open surgery, and with stomas (see Figure, Supplemental Digitial Content 4, which demonstrates the probability of cost-effectiveness of the different subgroup analyses).

DISCUSSION

Despite increasing evidence supporting the beneficial effects of ERPs, uptake remains low.[14] Economic benefits may provide a further argument for their adoption, but this has not yet been backed up with strong data. This study is the first to perform a comprehensive cost-effectiveness analysis comparing enhanced recovery to conventional perioperative care in patients undergoing colorectal surgery that incorporated post-discharge and socioeconomic outcomes and used standard economic methodology in a North American setting. The results demonstrate that an ERP was associated with shorter hospitalization, decreased resource utilization and lower societal costs compared to CC.

Much of the current literature supporting ERP has focused on clinical or physiologic outcomes in the immediate post-operative period.[219] Despite strong evidence that ERPs reduce length of stay, the economic benefits of these pathways have not yet been clearly demonstrated. Reductions in length of stay may not result in significant cost savings, as hospitals days saved at the tail of admission are not likely to be resource intensive.[20] Furthermore, individual studies have not consistently demonstrated lower complication rates with ERPs[11], which is an important cost driver for surgical admissions.[201] Other economic evaluations also have not included implementation costs. These factors partly explain why institutional costs were not different between ERP and CC in this study.

However, little is known regarding post-discharge outcomes and the socioeconomic impact of ERPs. Few studies have investigated patient-reported outcomes and patient-centered outcomes after ERP, especially beyond 30 days after surgery.[219] The lack of data has led to speculation, from both healthcare professionals and patients alike, that earlier discharge afforded by ERPs would simply shift the recovery process into the outpatient setting and increase burden

on caregivers.[220,221] In the present study, patients managed by ERP experienced less productivity loss, caregiver burden, and fewer visits to community health centers. This contributed to the decreased total societal costs in favor of the ERP.

We did not find a statistically significant difference in quality of life over 60 days between patients managed by ERP versus CC, although uncertainty analysis indicates that 78% of bootstrap estimates show increased QALYs for patients managed by ERP. This suggests that a larger sample size may have resulted in a statistically significant result, but the base estimate demonstrated a minimal difference (+0.86 quality-adjusted days = +0.002 QALYs). Large differences in quality of life are unlikely to be seen in the following months after surgery, as patients who "recovered" faster may have received adjuvant more therapy quickly. This may partly explain why the trajectory of postoperative recovery between the two groups converged at eight weeks (Figure 2). Other factors may include the finding that the dimensions measured by the SF-6D return to baseline by 8-weeks after abdominal surgery, inadequate sample size or limited instrument (SF-6D) sensitivity to detect subtle differences.[69,91] Whether faster postoperative recovery will result in improved long-term survival due to better tolerance of adjuvant therapies deserves investigation. We chose a time horizon of 60 days because previous studies have demonstrated that patients do not return to their pre-surgical baseline until at least two- to three-months after major abdominal surgery.[69,74,91,105]

The ERP had a greater than 99% probability of cost-effectiveness, even at a $0 per QALY threshold. Several factors need to be taken into account to the generalizability of these findings to other settings. Most importantly, there are differences in medical costs between the Canadian and other healthcare systems. For example, the average cost per hospital day in the US has been estimated to be US$1960 per day[222], compared to CAN$608 for the ERP institution in

Quebec, Canada. While it may seem that the disparity in unit costs would only further increase the mean difference in overall costs, other aspects such as dissimilarities in physician renumeration and hospital reimbursement need to be considered. Hospital case volume must also be considered, especially when considering the costs of designing, implementing, and maintaining an ERP. The ERP program in this study affected a wide variety of procedures, thus spreading out program costs across a large number of patients. ERP program costs per patient may be significantly higher at a smaller volume centre, as one study reported that ERP implementation costs were €1011 per patient.[223]

Extensive sensitivity and subgroup analyses were performed to address the limitations of this study. Most importantly, the cohort study design may have resulted in selection bias. While patient characteristics were similar between the two groups, there were still several imbalances that may have affected the cost-effectiveness of the ERP. There were more patients in the CC group with preoperative employment and in the highest salary bracket, although this did not reach statistical signficance. Nevertheless, this could have affected the mean overall costs in the CC group. A sensitivity analysis scenario in which all patients were assigned the same preoperative salary (median annual Canadian salary) did not differ from the base case result. A subgroup analysis examining only patients with preoperative employment as well as those without also did not differ from the base case result, although the probability of cost-effectiveness of the ERP was lower for non-employed patients. There were also more patients in the ERP group who underwent a laparoscopic approach. Laparoscopy has higher operative costs, but is associated with improved clinical and patient-reported outcomes, even within a conventional perioperative care program.[224] More recent evidence has shown that the benefits of laparoscopy were further increased when performed within an ERAS program, and vice

versa.[28,225] Given these findings, the higher incidence of laparoscopy in the ERP group could have biased the results of the present study, but subgroup analyses also did not demonstrate a difference from base case results, although the ERP in patients undergoing open surgery had a lower probability of cost-effectiveness, albeit still greater than 70% at a $0 per QALY threshold. There were also more stomas in the conventional care group (although the difference was not statistically significant), which may negatively affect postoperative quality of life.[226] However, subgroup analyses of patients with and without stomas still favored the ERP. The study design also recruited patients from two different institutions. Despite the fact that the ERP institution had higher unit costs than the CC institution, the ERP remained cost-effective. Not surprisingly, the ERP was even more cost-effective in a sensitivity analysis in which uniform unit costs were used across both participating institutions, and the difference in costs became significant at the institutional perspective. Conventional perioperative management may also differ from one institution to another. However, two recent surveys of perioperative care strategies reported similar processes of care as the conventional care group in the present study.[14,227] For example, the majority of cases in those studies still underwent bowel preparation and less than half received preoperative patient education. Lastly, the amount of missing data was moderate, and they were handled with multiple imputation. While multiple imputation is a valid method of handling missing data[131], the missing at random assumption that is required for unbiased estimates can never be definitively proven. Therefore, the effect of missing data was assessed by performing a complete-case analysis, in which the ERP remained cost-effective. These analyses suggest that our results were robust across a range of assumptions and subgroups.

Other limitations include the fact that there were relatively few assessments, performed at four-week intervals, which increases the potential for poor recall and recall bias. However the

need for additional assessments was balanced against the increase in patients' study burden, which may negatively affect response rate. The quality of life instrument that was used in this study had a recall period of one week.[91] Other instruments with shorter recall, such as the EQ-5D, could have been used, but have not been shown to be responsive to the construct of postoperative recovery.[206] Lastly, several outcomes, such as return to work and out-of-pocket expenses, are influenced by external factors such as employment type, and socioeconomic and insurance status, which were not measured in this study, but may have affected the results. The study questionnaires also only assessed for absenteeism and did not include 'presenteeism' (i.e. decreased productivity while at work), which may further impact productivity loss after surgery.

In summary, enhanced recovery perioperative management was associated with improved clinical and post-discharge outcomes, which resulted in lower overall costs when compared to conventional management in patients undergoing elective colorectal resection. ERPs did not shift the burden of care from the in-patient to outpatient setting, as ERPs were associated with faster return to work, and reduced caregiver burden. The initial upfront costs of development, implementation, and maintenance of ERPs did not increase the overall cost associated with colorectal surgery. These results were robust across a range of assumptions and subgroups. This study provides further evidence to support the adoption of ERPs.

ACKNOWLEDGEMENTS

LL received salary support from the Quebec Health Science Research Fund (FRQ-S) and the Fast Foundation (as part of the McGill Surgeon Scientist Program).

Table 1 – Comparison of perioperative management between conventional care and enhanced recovery pathway

Conventional Care	Enhanced Recovery Pathway
Preoperative Management	
Medical optimization of risk factors	Medical optimization of risk factors
No formal patient education	Psychological preparation for surgery and postoperative recovery: oral and written explanation of perioperative pathway, diet and ambulation plan, presence of drains, expectation about duration of hospital stay (3-4 days)
No instructions for preoperative exercise given	Physical preparation with exercises at home: anaerobic and resistance 1-2h/day, gradual increase from 50 to 80% of maximum capacity, breathing exercises
Nil per os starting midnight prior to surgery	Carbohydrate loading drink the night before and the morning of surgery. Clear fluids allowed until 2 hours prior to general anesthesia
Variable use of oral bowel preparation	Oral bowel preparation only if diverting ileostomy or intraoperative colonoscopy planned
Variable use of sedation routine	No preoperative sedation
Intraoperative Management	
Antibiotic and DVT prophylaxis	Antibiotic and DVT prophylaxis
Maintain normothermia	Maintain normothermia
IV infusion as per anesthesia	Avoid over-hydration (6 mL/kg/hr)
Insert epidural catheter at the discretion of the anesthetist	Insert epidural catheter routinely for open or laparoscopic rectal procedures, and local anesthetic infiltration of surgical incision
No routine nasogastric or abdominal drainage	No routine nasogastric or abdominal drainage
No routine prophylactic antiemetics	Routine prophylactic antiemetics
Postoperative Management	

Clear fluids starting after flatus, daily progression to diet as tolerated	Fluids (including two cans of Ensure®) on POD 0, diet as tolerated starting POD1
No structured mobilization plan	Structured mobilization: patient encouraged to sit in chair on POD0. Mobilization goal of at least 6 hours starting POD1
Urinary drainage catheter kept until epidural analgesia removed	Urinary drainage catheter removed on POD1
IV fluids discontinued only after resumption of oral intake	IV fluids discontinued upon arrival onto surgical ward on POD0
Thoracic epidural analgesia or patient-controlled analgesia until resumption of oral intake, then transition to oral analgesia (opioids).	Thoracic epidural analgesia or patient-controlled analgesia until POD2, then transition to multimodal analgesia (minimizing opioids)

Table 2 – Comparison of patient and operative characteristics between conventional care and enhanced recovery pathway

	Conventional Care (N=95)	**Enhanced Recovery (N=95)**	***p*-value**
Mean age, years (SD)	61.6 (13.4)	63.9 (13.1)	0.226
Male gender	50 (53%)	49 (52%)	0.995
Mean body mass index, kg/m^2 (SD)	27.7 (6.1)	26.9 (4.8)	0.308
Mean Charlson comorbidity index, points (SD)	2.3 (2.0)	2.5 (2.2)	0.514
Mean POSSUM[228] physiologic score, points (SD)	17.0 (5.3)	17.8 (5.3)	0.253
Mean POSSUM[228] operative score, points (SD)	11.9 (3.5)	11.3 (2.3)	0.130
Current smoker	10 (11%)	9 (9%)	0.809
Diabetes	17 (18%)	14 (15%)	0.556
Previous abdominal surgery	35 (37%)	39 (41%)	0.552
Diagnosis			0.833
Malignancy	63 (66%)	61 (64%)	
T1-T2	19 (30%)	12 (20%)	0.178*
T3-T4	44 (70%)	49 (80%)	
Inflammatory bowel disease	8 (8%)	11 (12%)	
Diverticular disease	7 (7%)	5 (5%)	
Other benign	17 (18%)	18 (19%)	
Neoadjuvant therapy	14 (14%)	15 (14%)	0.928
Procedure			0.441
Right hemicolectomy	36 (38%)	42 (44%)	
Rectosigmoidectomy	26 (27%)	29 (31%)	
Low anterior resection	14 (15%)	13 (14%)	
Proctocolectomy ± IPAA	8 (8%)	7 (7%)	
Left hemicolectomy	5 (5%)	3 (3%)	
Abdominoperineal resection	6 (6%)	1 (1%)	
Laparoscopic approach	45 (47%)	71 (75%)	<0.001
New stoma	33 (35%)	22 (23%)	0.056
Procedure duration, minutes (SD)	238 (146)	218 (105)	0.277
Adjuvant therapy by eight-weeks	11 (12%)	14 (15%)	0.557
Preoperative employment	50 (53%)	42 (44%)	0.246
Yearly salary†			0.185
$0 to $30,000	14 (28%)	6 (14%)	
$30,000 to $60,000	15 (30%)	12 (29%)	
$60,000 to $90,000	8 (16%)	9 (21%)	
$90,000+	11 (22%)	8 (19%)	
Did not disclose‡	2 (4%)	7 (17%)	

*Malignancy only

†Amongst those with preoperative employment

‡Patients who did not disclose their yearly salary were assigned the lowest income bracket for the analysis. The distribution of salary did not change ($p = 0.922$) which this modification.

Table 3 – Comparison of clinical and post-discharge outcomes between patients managed by enhanced recovery versus those managed by conventional care.

	Conventional Care (n=95)	**Enhanced Recovery (n=95)**	***p*-value**
Clinical outcomes			
Mean total hospitalization, days (SD)	9.8 (12.2)	6.5 (6.0)	0.017
Mean primary hospitalization, days (SD)	9.4 (11.8)	5.7 (5.5)	0.007
60-day emergency room visits	17 (18%)	19 (20%)	0.711
60-day readmissions	10 (11%)	12 (13%)	0.650
60-day complications	41 (43%)	38 (40%)	0.659
Mean complication severity, (SD)	10.7 (174)	10.2 (14.3)	0.814
Post-discharge outcomes			
Returned to work by eight-weeks[*]	29 (58%)	30 (71%)	0.181
Mean lost days from work,[*] days (SD)	34.8 (19.7)	25.8 (17.8)	0.035
Mean postoperative days requiring help with activities of daily living[†], days (SD)	14.7 (17.3)	14.9 (12.1)	0.901
Mean caregiver lost days from work, days (SD)	5.0 (12.0)	1.3 (2.6)	0.004
Mean postoperative community health service centre visits, visits (SD)	3.7 (8.7)	1.4 (4.6)	0.026
Mean postoperative surgeon follow-up visits, visits (SD)	1.1 (0.5)	1.0 (0.4)	0.064
Mean postoperative specialist or family medicine visits, visits (SD)	1.2 (1.6)	0.9 (1.4)	0.099

[*]Amongst those patients that were previously employed
[†]Includes both basic or instrumental activities of daily living

Table 4 – Comparison of institutional, healthcare system, and societal costs between conventional care and enhanced recovery pathway.

	Conventional Care (N=95), $ (95% CI*)	**Enhanced Recovery Pathway (N=95), $ (95% CI*)**	**Mean Difference, $ (95% CI*)**
OR costs	5123 (4621, 5641)	4800 (4491, 5154)	-323 (-889, 284)
Hotel costs	5137 (3888, 6955)	4263 (3428, 5258)	-874 (-3033, 677)
Radiology costs	65 (37, 110)	81 (55, 116)	+16 (-32, 61)
Pharmacy costs	826 (657, 1084)	747 (617, 891)	-79 (-355, 159)
Laboratory costs	123 (86, 182)	73 (44, 84)	-50 (-126, -15)
Allied health professional costs	53 (44, 63)	24 (11, 40)	-29 (-45, -11)
Ambulatory clinic costs	175 (150, 203)	145 (122, 170)	-30 (-66, 5)
Emergency department costs	70 (38, 101)	44 (26, 65)	-26 (-63, 10)
Chemotherapy costs	246 (105, 428)	336 (159, 532)	+90 (-151, 351)
ERP-specific costs (design, implementation, audit)	–	153†	+153†
Institutional costs	**11818 (10287, 14217)**	**10668 (9621, 12042)**	**-1150 (-3487, 905)**
Community health service centre costs	287 (162, 456)	137 (66, 222)	-150 (-304, -2)
Assisted-care facility costs	90 (0, 248)	65 (0, 195)	-25 (-224, 181)
Physician billing costs	1976 (1789, 2206)	1699 (1586, 1816)	-277 (-535, -54)
Healthcare system costs	**14171 (12436, 16671)**	**12569 (11482, 13968)**	**-1602 (-4050, 517)**
Productivity losses	2724 (1918, 3542)	1662 (1077, 2316)	-1062 (-2054, -74)
Caregiver costs	389 (213, 581)	142 (69, 248)	-247 (-463, -49)
Out of pocket expenses	270 (173, 380)	198 (121, 322)	-72 (-220, 77)
Society costs	**17556 (15456, 20154)**	**14571 (13395, 16230)**	**-2985 (-5753, -373)**

*Derived from 2.5th and 97.5th percentiles of 10,000 bootstrap replications
†No confidence interval

Table 5 – Results of sensitivity and subgroup analyses. Note that incremental costs are calculated from a societal perspective. Confidence intervals were derived from the 2.5th and 97.5th percentile bootstrapped estimates.

	Incremental Costs (95% CI)	**Incremental Quality-Adjusted Days (95% CI)**	**ICER**
Sensitivity analyses			
Complete case analysis	-2987 (-6040, -108)	+0.58 (-0.90, 2.05)	Dominant
Uniform unit costs	-5421 (-8842, -2443	+0.87 (-1.23, 2.97)	Dominant
Uniform pre-employment salary*	-2311 (-5226, 130)	+0.87 (-1.23, 2.97)	Dominant
Friction valuation†	-2773 (-5608, -164)	+0.87 (-1.23, 2.97)	Dominant
Subgroup analyses			
Employment status			
Employed	-4194 (-7556, -608)	+1.33 (-1.23, 3.89)	Dominant
Unemployed	-1094 (-6216, 2420)	+0.84 (-2.41, 4.09)	Dominant
Surgical approach			
Laparoscopic	-2251 (-5072, -566)	+1.19 (-1.64, +4.04)	Dominant
Open	-1346 (-7501, 4371)	+0.55 (-1.09, 2.20)	Dominant
New stoma fashioned			
Yes	-942 (-5516, 3604)	+0.02 (-3.93, 3.99)	Potentially cost-effective
No	-2503 (-6486, 657)	+0.86 (-1.47, 4.14)	Dominant
Complications			
No complications	-2500 (-4334, -718)	+0.86 (-2.57, 4.30)	Dominant
Any complications	-3102 (-8938, 1768)	+0.78 (-1.80, 3.36)	Dominant

*Canadian median salary
†Elasticity factor of 0.8

Supplemental Digital Content 1 – Description of important unit costs at the enhanced recovery pathway and conventional care centres. All costs include overhead costs (allocated by step-down method) and are expressed in 2013 CAN$ (=0.86 US$[214]).

	Conventional Care Institution	**Enhanced Recovery Pathway Institution**
Operating room time cost*, per hour	$899.37	$913.77
Ward cost, per day	$493.05	$608.15
Intensive care unit cost, per day	$1527.40	$2002.57
Emergency room visit cost, per visit	$319.26	$229.43
Physiotherapist cost, per hour	$61.85	$69.96
Radiology costs, per unit†	$2.43	$2.59
Laboratory costs, per unit†	$1.06	$1.21
Ambulatory clinic cost, per visit	$76.64	$70.86
Common unit costs		
CLSC costs, per visit	$80.24	
Rehabilitation facility, per day	$206.98	

*Not including equipment costs

†Radiology and laboratory tests are each ascribed a specific number of units, akin to Relative Value Units (RVUs), which are standard across all public health centres within the province of Quebec, Canada. For example a computed tomography scan of the abdomen and pelvis with contrast is worth 27 units.

Supplemental Digital Content 2 – Comparison of the adherence to enhanced recovery pathway elements between patients managed by enhanced recovery versus those managed by conventional care.

	Conventional Care (n=95)	**Enhanced Recovery (n=95)**	***p*-value**
Preoperative management			
Preoperative written patient education	0 (0%)	95 (100%)	<0.001
Preoperative mechanical bowel preparation	63 (66%)	34 (36%)	<0.001
Preoperative sedative	54 (57%)	0 (0%)	<0.001
Preoperative carbohydrate drink	0 (0%)	46 (48%)	<0.001
Intraoperative management			
Antibiotic prophylaxis	95 (100%)	95 (100%)	1.000
Mean intraoperative IV crystalloid administration, mL (SD)	2475 (1368)	1707 (1122)	<0.001
Mean intraoperative IV colloid, mL (SD)	429 (405)	305 (385)	0.038
Use of abdominal drains	13 (14%)	4 (4%)	0.022
Nasogastric tube left in situ	5 (5%)	1 (1%)	0.097
Normothermia (>36^{o}C) at end of operation	91 (96%)	91 (96%)	0.710
Thoracic epidural analgesia	61 (64%)	56 (59%)	0.456
Postoperative management			
Perioperative thromboprophylaxis	95 (100%)	95 (100%)	1.000
Median days to mobilization > 2h/day, days [IQR]	2 [1-2]	1 [1-2]	<0.001
Median days to discontinuation of intravenous fluids, days [IQR]	3 [2-5]	1 [1-1]	<0.001
Median days to passage of first flatus, days [IQR]	3 [2-3]	1 [1-2]	<0.001
Median days to receiving fluids, days [IQR]	2 [1-3.5]	0 [0-0]	<0.001
Median days to toleration of solid diet, days [IQR]	4 [3-5]	1 [1-2]	<0.001
Median days to removal of indwelling bladder catheter, days [IQR]	2 [1-3]	1 [1-1]	<0.001

IV, intravenous

Supplemental Digital Content 3 – Breakdown of enhanced recovery pathway (ERP) program costs over the study period.

	Cost (2013 CAN$)
Full-time ERP nurse coordinator (yearly salary)	81 225
Opportunity costs of ERP steering group (1 hr/meeting × 26 meetings)	14 320
Nurse specialists and managers, nutritionist, physiotherapist, librarian, clinical leaders from surgery and anaesthesia ($550 per meeting)	
Patient education material (operating costs of work performed by a medical informatics centre)	13 225
Total	108 770

Figure 1 – Study flowchart

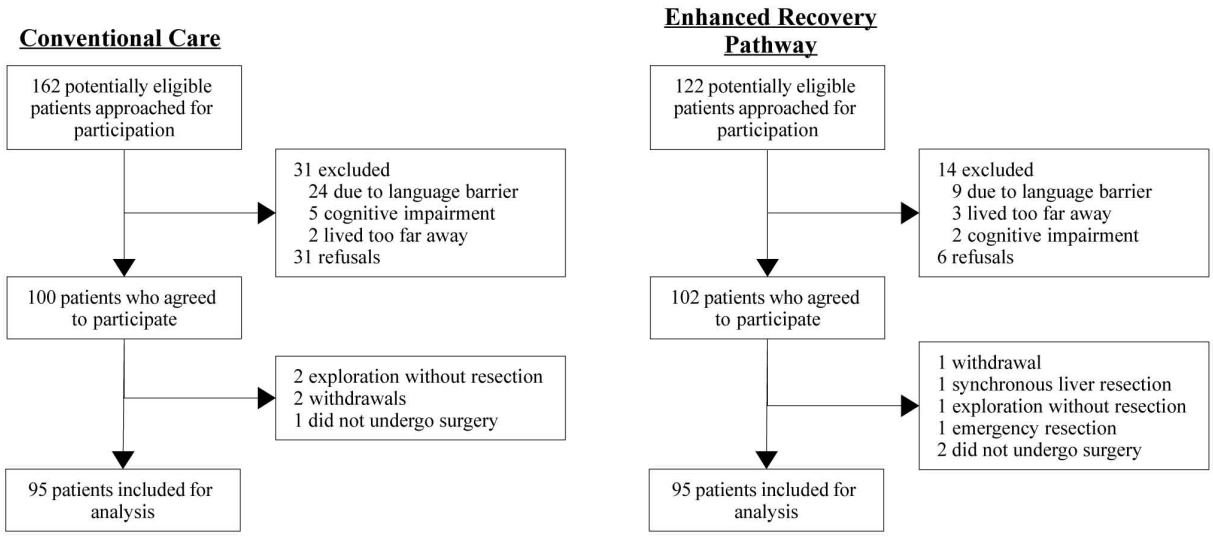


Figure 2 – Comparison of mean SF-6D scores between patients managed by conventional perioperative management and patients managed by an enhanced recovery pathway

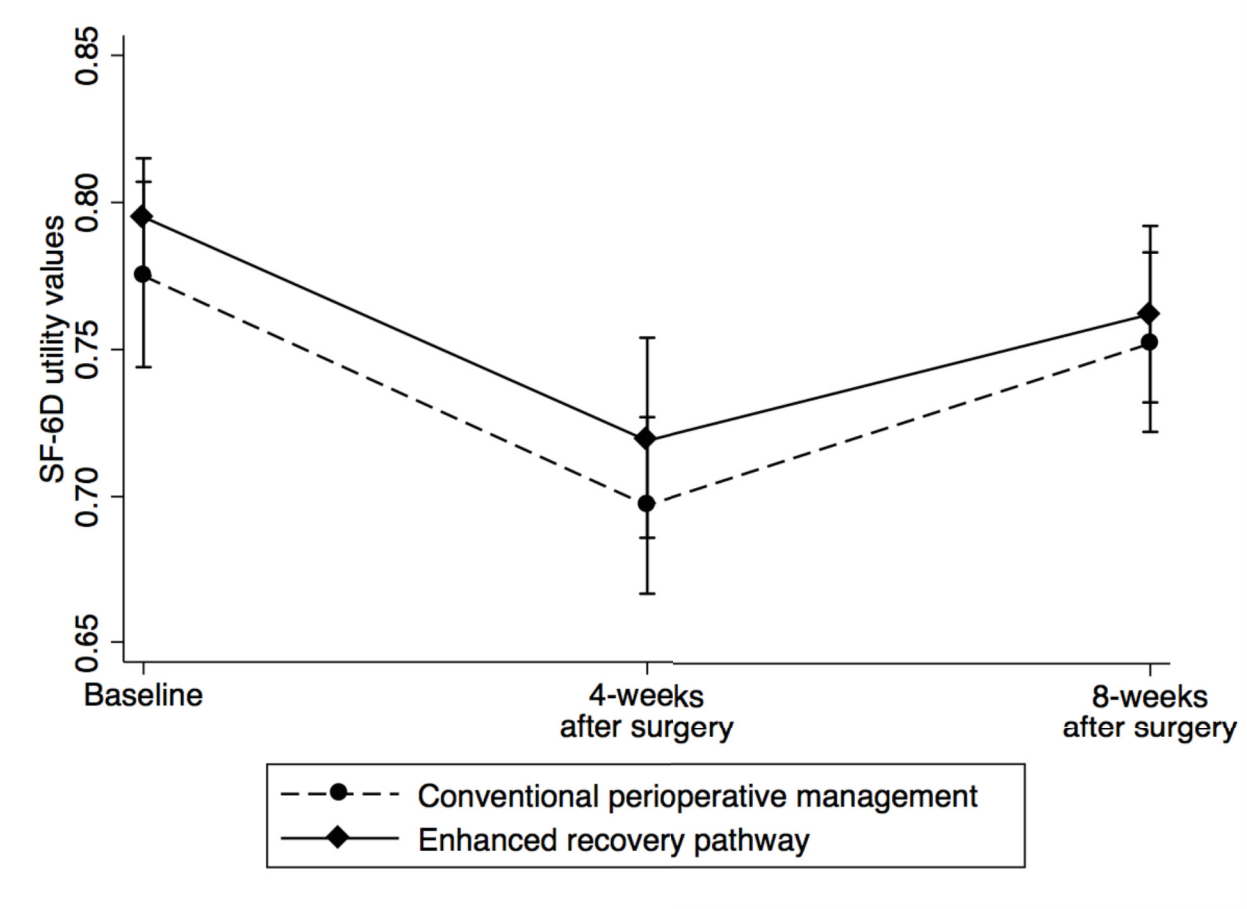


Figure 3 – Incremental societal costs and QALYs (each point represents one bootstrap estimate) plotted on the cost-effectiveness plane

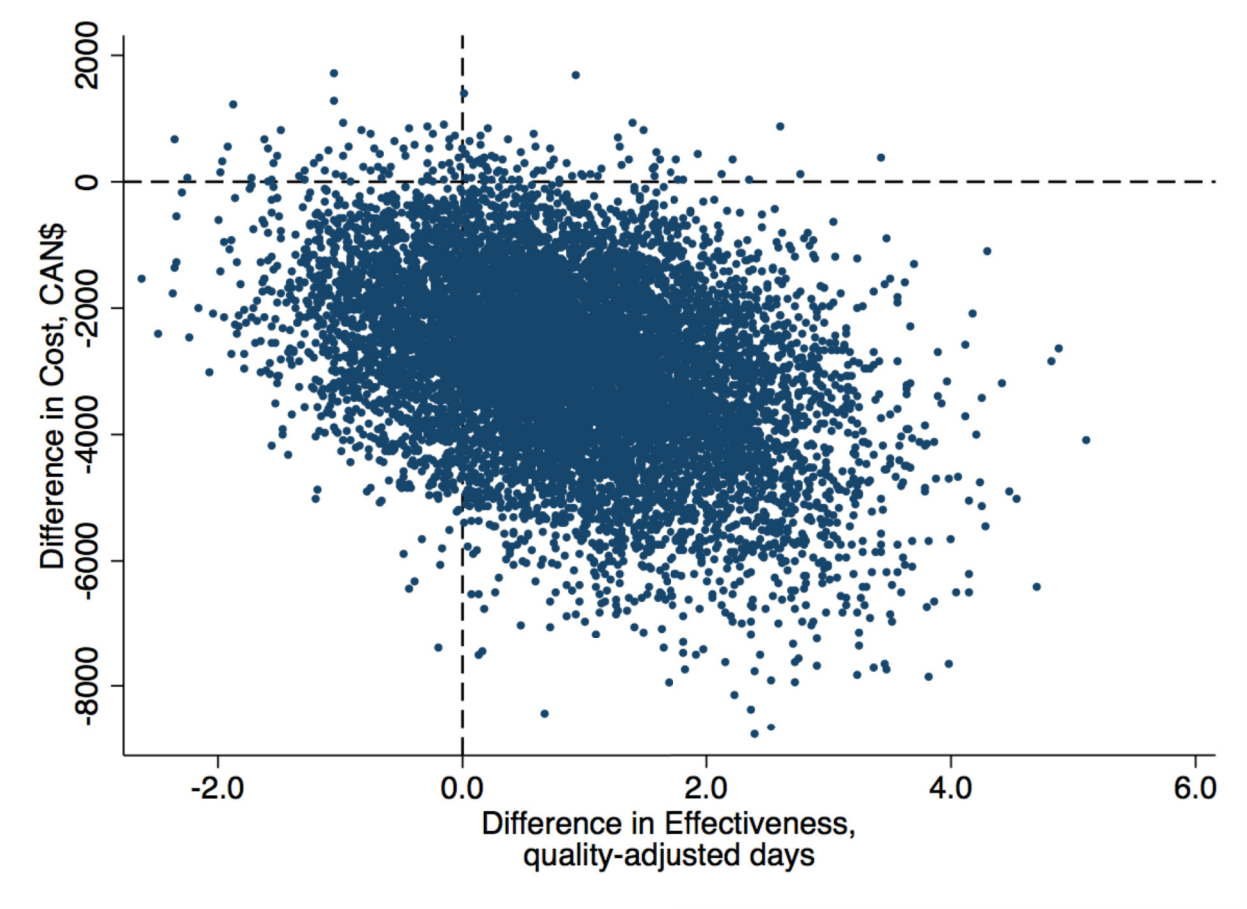
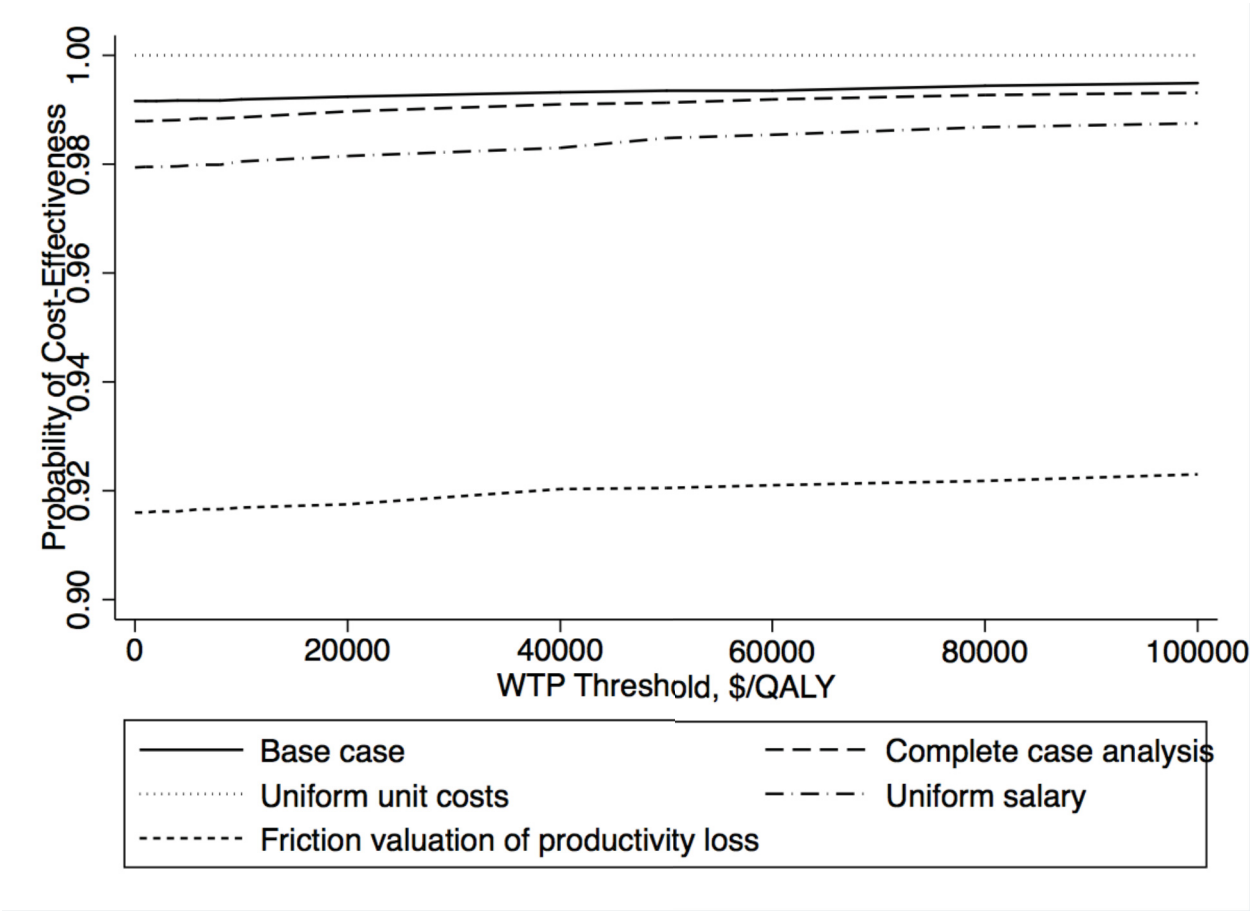
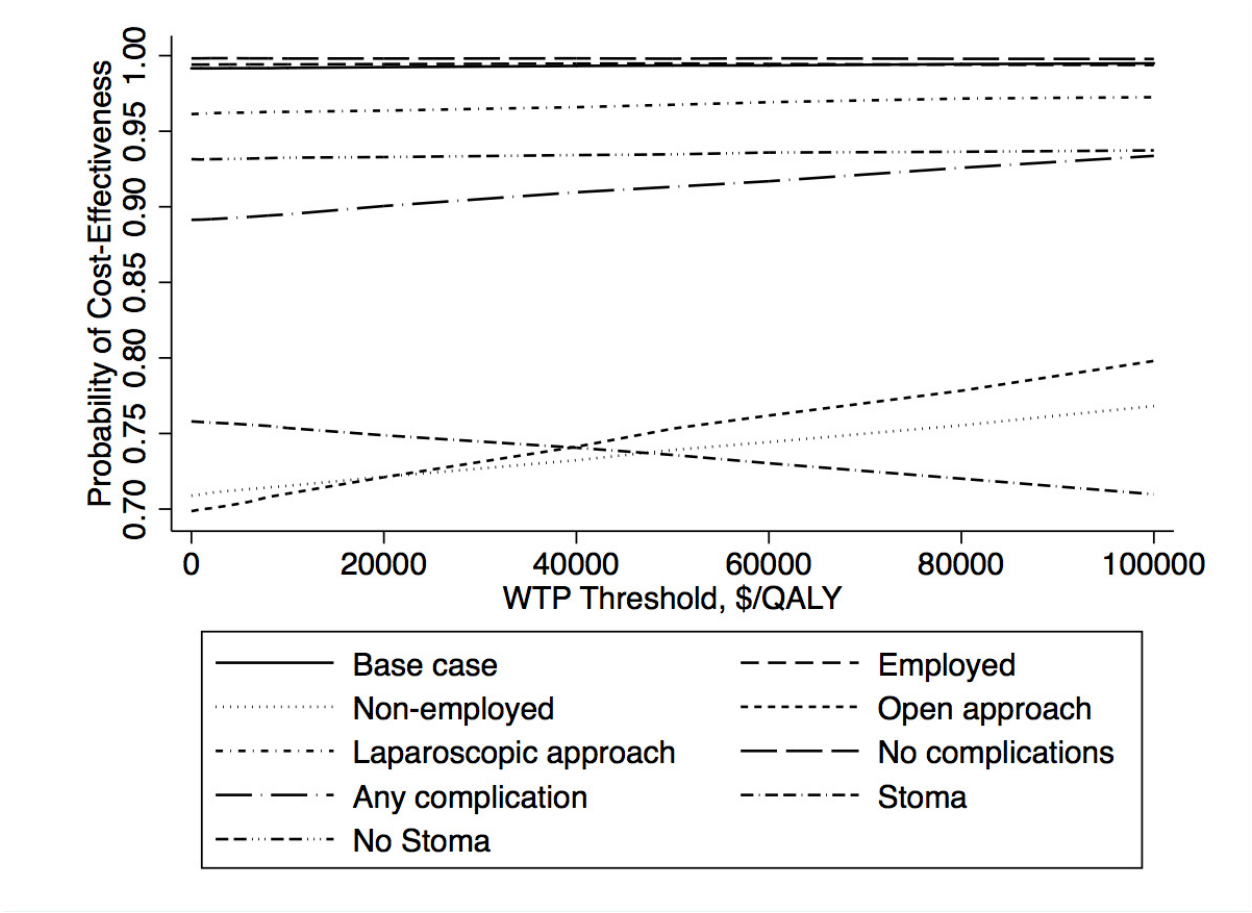


Figure 4 – Cost-effectiveness acceptability curve of the base case and sensitivity analysis results (from the societal perspective).



Supplemental Digital Content 4 – Cost-effectiveness acceptability curve of the base case and subgroup analysis results (from the societal perspective).



9. SUMMARY AND CONCLUSIONS

9.1 GENERAL FINDINGS

In this dissertation, I have demonstrated the cost-effectiveness of an ERAS program over conventional perioperative management in patients undergoing elective colorectal surgery. First, a systematic review (Chapter 2) was performed to identify and evaluate the existing economic evidence. The overall quality of the identified economic evaluations was poor, and major knowledge gaps were found regarding the societal impact of ERAS programs. Therefore the main objective of this dissertation was to determine the cost-effectiveness of ERAS, taking into account the evidence gaps identified in the systematic review. However, the design of the cost-effectiveness analysis required two important elements that had not been previously described: a utility instrument that accurately captured the construct of postoperative recovery, and a local setting-specific effect estimate so that sample size calculations could be performed.

Prior to psychometric evaluation of existing utility instruments, a clear definition of postoperative recovery was required. The next manuscript (Chapter 3) defined the construct of postoperative recovery as a complex multidimensional construct that followed an expected trajectory. Multi-attribute utility instruments were thought to be ideal for the measurement of recovery process given that they evaluate multiple different dimensions in a single summary measure. An examination of the contents of pre-existing instruments demonstrated that the SF-6D contained many of the domains that were thought to be important for recovery. The third manuscript (Chapter 5) in this dissertation validated the SF-6D as a measure of postoperative recovery. It was found to be responsive to the expected trajectory, discriminated between groups of patients with and without complications (which affects the recovery process), and correlated with other measures of the postoperative recovery construct. However, content examination of

other utility instruments also identified the EQ-5D as a potentially useful measure. The fourth manuscript (Chapter 6) compared the validity of the SF-6D and the EQ-5D as measures of postoperative recovery. It was shown that the SF-6D exhibited superior psychometric characteristics for the construct of recovery compared to the EQ-5D and the EQ-VAS.

The last step in the study design of the cost-effectiveness analysis was to determine a locally specific effect estimate so that an anticipated sample size could be calculated. The fifth manuscript (Chapter 7) performed a cost analysis comparing ERAS and conventional perioperative management for esophagectomy at one of the institutions participating in the formal cost-effectiveness analysis study. Despite no difference in the incidence of postoperative complications, the implementation of an ERAS program for esophagectomy was associated with cost savings of CAN$2666 per case. This estimate was in line with economic evaluations of other surgical interventions that were advocated to improve recovery in colorectal surgery.

The last manuscript in this dissertation (Chapter 8) describes the results of the cost-effectiveness analysis comparing ERAS and conventional perioperative management in patients undergoing elective colorectal surgery. The study was designed with an emphasis on the evaluation of post-discharge outcomes and socioeconomic impact, especially beyond 30-days after surgery, which had been lacking in previous studies. The SF-6D was used as the main outcome measure. An observer unaware of the study hypothesis prospectively recorded in-hospital outcomes. Questionnaires were used to measure outpatient outcomes at four- and eight-weeks after surgery. The questionnaires that were used in this study had been previously designed specifically to measure outpatient resource utilization and socioeconomic impact in the Canadian setting.[211,212] Recruitment took place over a one-year period, during which 190 patients (95 in each group) completed the study. Patient and operative characteristics were similar

between the two cohorts, except for a higher incidence of laparoscopic procedures in the ERAS group. Overall, ERAS was associated with decreased costs from a societal perspective, but not from the institutional or healthcare system perspectives. Furthermore, no difference in effectiveness, as measured by the SF-6D, was found between ERAS and conventional care. Nevertheless ERAS was associated with a high probability of cost-effectiveness at all willingness-to-pay thresholds. Several sensitivity and subgroup analyses were performed to account for several of the study limitations. ERAS remained cost-effectiveness in all analyses.

One of the main limitations of previous economic evaluations of ERAS pathways in colorectal surgery is the absence of program implementation and maintenance costs. Two previous studies have addressed this topic.[31,223] Roulin and colleagues estimated the attributable cost per patient for their ERAS program (for the first 50 patients recruited at their centre) was €1011 (= CAN$1301 in 2012), which included the costs of the carbohydrate drinks and nutritional supplements, logbook, "enhanced recovery implementation costs", travel and lunch expenses of multidisciplinary team training, and the cost of a full-time dedicated nurse for a 6 month period.[223] Similarly, Sammour and colleagues estimated that their ERAS program costs amounted to a total of NZD96391 (= CAN$78339 in 2007) for their first 50 patients, resulting in an average additional cost of NZD1927 per patient (= CAN$1566). Their costs included travel costs of their multidisciplinary group to Denmark (to learn under Dr. Henrik Kehlet), 15 month salary for their research fellow, patient booklet, and nutritional supplement costs. Comparatively, ERAS program costs at our institution amounted to CAN$153 per patient, which was due to the unique structure of our ERAS program. Our costs were mainly comprised of the yearly salary of a full-time dedicated ERAS nurse coordinator, opportunity costs of the multidisciplinary group members, and the operating costs of the medical informatics group, who create and continuously

modify the patient education booklet. While the total yearly cost of our program was similar to the other two studies (CAN$108770), the total cost at our institution was distributed amongst a larger number of patients, as our multidisciplinary ERAS program has implemented pathways for a variety of surgical procedures, not just colorectal surgery. The expertise and learning curve that was gained from implementing ERAS for colorectal surgery was then put to use to successfully develop and implement other pathways, such as esophagectomy[167] and prostatectomy (Abou-Haider H et al., unpublished data). Given this structure, it was difficult to ascertain the exact cost of the colorectal pathway specifically. Nevertheless, this approach allowed for the significant upfront investment and yearly costs to be spread out across a much larger number of patients, as well as improving outcomes and decreasing costs for other procedures.[217]

9.2 LIMITATIONS

The cost-effectiveness analysis in Chapter 7 was conducted according to commonly accepted economic methodology[52,59] and addressed many of the limitations of the existing economic evaluations that were identified in the literature review[204]. However, the results should be interpreted in light of several important limitations. Despite our best attempts at completing follow-ups for all participating patients, there was a small proportion of missing data (21/190 patients, all of whom missed the eight-week assessment). These missing data were handled by multiple imputation in the base-case results. Multiple imputation was incorporated into bootstrap uncertainty analysis by first creating the bootstrapped samples (10,000 iterations), then performing multiple imputation on each of the samples, as described by Shao and Sitter.[229] This method has the advantage of including the variability produced by multiple imputation but not

requiring the rules for combining multiply-imputed estimates to calculate confidence intervals (since the standard deviation relies solely on the distribution of the β estimates of the bootstrapped samples). The amount of missing data was also well within the range in which multiple imputation still provides non-biased estimates.[131] However, multiple imputation requires that missing data be missing-at-random. While this assumption can never be definitely proven, there was no difference in the incidence of postoperative complications, which was hypothesized to be one of the main reasons for potential dropout, between patients with and without missing data (50% vs. 43%, p=0.567). While it could be argued that these patients could have presented with complications at an outside institution, one of the inclusion criteria was residence proximity to the participating institutions, which in theory should have limited these events.

One of the other limitations was that the unit costs between the two participating institutions were different, with the unit costs at the ERAS institution generally higher compared to the conventional care institution. This could have partly explained why no difference in costs from the institutional perspective was found, although it is not clear whether the higher unit costs were due to ERAS or higher medical or administrative costs. Outcomes at the participating conventional care institution were in line with those reported in the literature, suggesting that the overall results were not due to an 'outlier effect'. Furthermore, this institution was in the process of adopting an ERAS program, although it was not yet in place during the conduct of our study. Nevertheless, several elements could have been implemented piecemeal. However, examination of the adherence to ERAS elements demonstrated that a formal ERAS program was nonetheless associated with improved processes of care. A sensitivity analysis was performed in which all uniform unit costs (those of the ERAS institution) were assigned to mimic what an institution

may experience after ERAS implementation. In this scenario, a significant difference in institution costs in favor of ERAS was found.

The cohort study design also resulted in several imbalances in patient characteristics between the two groups. There was a higher incidence of laparoscopic procedures in the ERAS group, which could have affected the cost-effectiveness results. Laparoscopy has higher operative costs, but is associated with improved clinical and patient-reported outcomes, even within a conventional perioperative care program.[224] More recent evidence has shown that the benefits of laparoscopy were further increased when performed within an ERAS program, and vice versa.[28,225] These findings were also reproduced within the present study, as patients undergoing laparoscopic procedures experienced less morbidity in the conventional care group and shorter hospitalization for both ERAS and conventional care (see Appendix 3). There was no difference in quality of life between patients undergoing laparoscopic and open surgery in either group. Subgroup analysis demonstrated that ERAS remained cost-effective even within a subgroup of patients undergoing laparoscopic surgery, as well as those undergoing open surgery. These findings suggest the increased operative costs of laparoscopy could be offset through an ERAS program.

There may also have been a clinically important imbalance in the proportion of patients that were employed pre-operatively, as well as in the distribution of annual salary, between the ERAS and conventional care groups. Given that productivity losses contribute a significant portion to societal costs, any difference may bias the overall result. It was also interesting to note that a higher proportion of patients in the ERAS were not willing to disclose their annual salary. It is unclear whether these patients refused because their salary was too low or too high (social desirability bias). All patients that did not disclose their salaries were assigned the annual

Canadian median salary in the analysis. The results did not differ in a sensitivity analysis in which all employed patients were assigned the annual Canadian median salary, or in subgroup analyses for employed and non-employed patients. Furthermore, the type of employment and insurance status were not included in the study questionnaires. These factors may have affected the base-case results, as patients with physically demanding occupations may have required a longer rehabilitation period compared to those with office-based work. The human capital approach was used to estimate productivity losses, despite the fact that the Canadian Agency for Drugs and Technology in Health Guidelines for the Economic Evaluation of Health Technologies recommends the friction method.[230] This may affect the results if patients with skilled occupations are harder to replace, leading to increased societal productivity losses. However, the friction method is not likely to make a difference since the friction period in Canada has been estimated to be 14.6 weeks[216] (i.e. longer than the time horizon of the present study). Certain authors have also argued that friction losses are only worth 80% of the true productivity loss , reflecting compensation mechanisms and certain redundancy.[215,231] However, this empirical value is not likely to be applicable to all occupation types, nor is it clear as to the meaning of this elasticity factor, leading to its general exclusion from practical use.[232] Nevertheless, a sensitivity analysis was performed using the friction valuation of productivity loss (i.e., applying an elasticity factor of 0.8). The results of this analysis (Chapter 8, Table 5) remained favorable for ERAS.

Lastly, the results did not demonstrate a difference in recovery between ERAS and conventional care, despite the use of a validated measure.[91] There may be several reasons for these findings. First, the ideal time points at which to measure recovery have not yet been clearly identified. The four- and eight-week time points that were used for the SF-6D validation and

cost-effectiveness studies were based on previous empirical work and defined to coincide with the first postoperative follow-up visit (usually at four weeks after surgery).[69,74,233] These previous studies demonstrated that the significant proportion of patients had already returned to baseline function by six- to eight-weeks after surgery. Furthermore, it was hypothesized that any measurement after this point was likely to be confounded by any adjuvant therapy (which is usually administered within eight-weeks after surgery). Increased frequency of measurement in the postoperative period may have picked up larger differences between the ERAS and conventional care groups. However, Jones and colleagues administered the EQ-5D on postoperative days 2, 3, 5, 7, 10, 14, and 28 in a trial comparing ERAS and conventional care in patients undergoing major liver resection and did not find any difference between the two groups (their study also used the area under the curve method).[12] The cost-effectiveness study was also not powered to detect small changes in quality of life. Power calculations using the minimally important difference for the SF-6D (0.041[117]) and the standard deviation of the four-week SF-6D of the study sample (0.132) would have required more than 150 patients per group to detect the minimally important difference at $\alpha=0.05$ and $\beta=0.8$. Lastly, the relationship between improved short-term biologic and physiologic changes and long-term outcomes has not yet been definitively demonstrated.[72] That is, functional outcomes that are important to patients in the late phase of recovery may not be affected by the immediate postoperative benefits of ERAS such as quicker return of gastrointestinal function and decreased surgical stress response.

9.3 FUTURE DIRECTIONS

There are several important questions that have arisen from the results of this dissertation, which should be the focus of future research. Chapter 7 has demonstrated that ERAS is

associated with decreased societal costs over a 60-day time horizon. However, costs and outcomes beyond this time frame are unknown. It can be hypothesized that the improved outcomes after ERAS may allow for more patients to receive adjuvant therapy with less delay. Over the long term, this may be associated with increased costs (due to the high cost of adjuvant therapies) and may even paradoxically decrease quality of life (due to the side effects of chemotherapy), although this will depend on the time horizon. Ultimately, it can be hypothesized that ERAS should increase survival (and therefore be associated with increased QALYs), and remain cost-effective over a lifetime time horizon. A prospective population-based study would be required to investigate this hypothesis.

Another question that remains unanswered is the proper valuation of having open beds as a result of ERAS (i.e., the opposite of the opportunity cost of the beds occupied as a result of not using ERAS). Multiple randomized trials have consistently demonstrated that ERAS reduces length of stay.[11] As a result, there may be increased numbers of available hospital beds that can be directly attributable to ERAS. However the downstream effects of this increased efficiency has not been clearly elucidated. On one hand, increased bed availability may increase operative capacity and decrease emergency department wait times. Reductions in operative wait time have been associated with improved long-term cancer survival and quality of life[234,235], although this has not been consistently demonstrated for all neoplasms[236]. However the true budgetary impact will be dependent on local factors, such as operational capacity and payment systems. Institutions that continuously function at full capacity will likely benefit the most from increased bed availability. Hospitals under prospective payment systems (i.e. a predetermined amount based on a diagnosis) or global budgets may experience increased overall costs if available beds are occupied by high-cost acute cases instead of elective patients at the tail end of their surgical

admission (which contribute minimally to hospital resource use and costs[20]). Conversely, revenue may also be increased due to higher capacity, as shorter length of stay may allow for higher turnover.[237] Future studies should account for all of these variables to determine the true costs and benefits of ERAS.

9.4 CONCLUSION

In conclusion, I have provided evidence to support the cost-effectiveness of ERAS over conventional care for patients undergoing elective colorectal surgery. In particular, this analysis addressed many of the limitations of previous economic evaluations that were identified from a systematic review, namely the inclusion of post-discharge recovery measures and description of the socioeconomic impact of ERAS. This dissertation also identified the SF-6D as a valid outcome to measure recovery. This instrument was validated specifically for the construct of postoperative recovery, conceptually defined as a multidimensional construct that followed an expected trajectory of deterioration and rehabilitation back to or surpassing the preoperative baseline. Finally, a future research agenda was suggested, which focused mainly on the costs and benefits of ERAS on a population level. Ultimately, this dissertation has demonstrated that ERAS leads to high value cost-effective healthcare by lowering costs without compromising outcomes.[238]

10. APPENDICES

10.1 APPENDIX 1 – Alternate sample size calculations

The sample size calculations for the main cost-effectiveness study (Chapter 8) were performed according to a pilot study (Chapter 7) that only incorporated cost data. The estimated effect size (difference in means between two groups divided by the standard deviation) from these pilot data resulted in required sample size of 100 patients per group at $\alpha = 0.05$ and $\beta = 0.2$. However limitations with this approach include the fact that the effect size formula assumes a normal distribution (cost data tend to be right-skewed), and that only costs or effects (and not both) can be considered. An alternate method to calculate require sample sizes for economic evaluation is to base it on the expected net monetary benefit (NMB), as is demonstrated in Equation 1.[203,239] This allows for the inclusion of the expected costs and benefits, as well as uncertainty around these point estimates, within a single calculation. This calculation estimates the sample size that is required to be 95% confident that the NMB will be greater than 0.

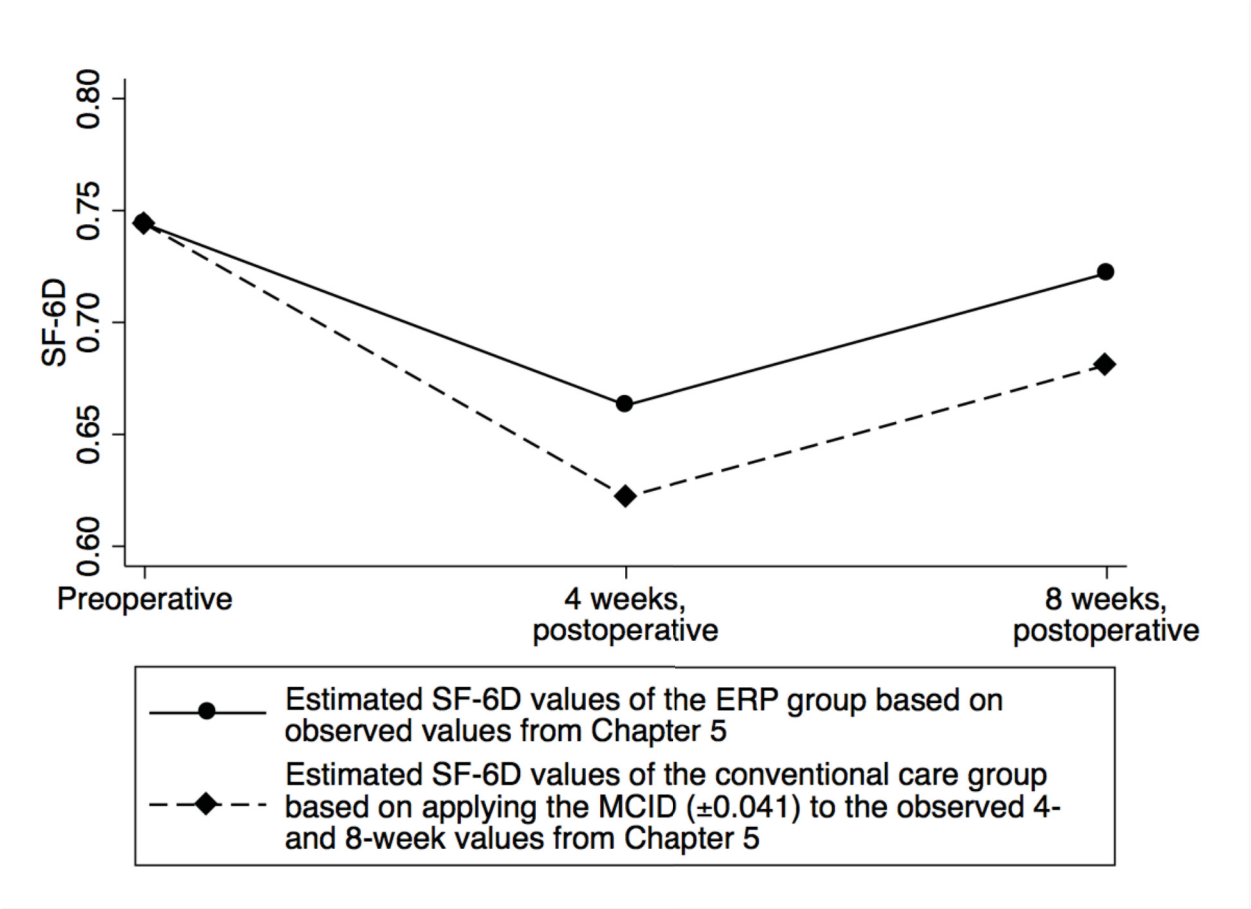
$$\textbf{Equation 1}: n = \frac{2(Z_\alpha Z_\beta)^2\left[sd_c{}^2+(Wsd_q)^2-(2W\rho sd_c sd_q)\right]}{(WQ-C)^2}$$

where n represents the expected sample size per group; Z_α, the z-statistic for the α error (Z_α =1.96 for α=0.05); Z_β, the z-statistic for the β error (Z_β=0.84 for β=0.2); sd_c, the standard deviation for the cost in each treatment group; sd_q, the standard deviation for the effect in each treatment group; W, the willingness to pay threshold; ρ, the correlation between difference in costs and effects; C, expected difference in costs; and Q, the expected difference in effects.

This calculation requires an estimate of the effect difference (represented by Q in Equation 1). Figure 10.1.1 demonstrates the estimated difference in QALYs, based on the assumption that there would be a clinically significant difference between ERP and conventional care (i.e. a difference of 0.041 at four- and eight-weeks, based on the minimal clinically important difference of the SF-6D[117]). Baseline values for the ERP group were obtained from Chapter 5.

The estimated difference in effectiveness was calculated to be 1.91 quality-adjusted days (=0.005 QALYs). The standard deviation of QALYs in the ERP group was calculated to be 0.003 based on these data from Chapter 5.

Figure 10.1.1 – Estimated difference in SF-6D between the ERP and conventional care groups based on applying the SF-6D MCID to observed values from Chapter 5



Furthermore, Equation 1 also requires a defined WTP threshold. There are several issues with the inclusion of the WTP, as the appropriate WTP threshold is not always clear, and the sample size does not linearly increase with the WTP threshold.[239] Therefore the estimated sample size should be calculated for a range of WTP thresholds. Table 10.1.1 displays the estimated sample sizes based on the estimates for cost (from Chapter 8) and effectiveness (see above) at $\alpha = 0.05$ and $\beta = 0.2$ at varying WTP thresholds. The correlation between difference in costs and

effects was observed to be moderate ($\rho = -0.405$). In addition, the sd_c is equal to \$6665 rather than \$5000 as previously assumed (given that the effect size (ES = $\Delta C/ sd_c$) used in the original sample size calculations was 0.4 with an estimated cost difference of \$2666). The estimated sample size at a WTP threshold of \$0/QALY is identical to the estimated sample size calculated in Chapter 10 (i.e. incorporating only costs).

Table 10.1.1 – Estimated sample sizes for varying WTP threshold

WTP threshold	Sample size required
\$0 per QALY gained	98
\$50000 per QALY gained	121
\$100000 per QALY gained	154

Z_α, the z-statistic for the α error (Z_α =1.96 for α=0.05); Z_β, the z-statistic for the β error (Z_β =0.84 for β=0.2); sd_c, the standard deviation for the cost in each treatment group (sd_c = \$6665); sd_q, the standard deviation for the effect in each treatment group (sd_q =0.003 QALYs); W, the willingness to pay threshold; ρ, the correlation between difference in costs and effects ($\rho = -0.405$); C, expected difference in costs (C = \$2666); and Q, the expected difference in effects (Q = 0.005 QALYs).

Alternatively, the power of the study can be calculated using Equation 2, where the power of the study is the area under the standard normal distribution to the left of Z_β.

$$\textbf{Equation 2}: Z_\beta = \sqrt{\frac{n(WQ-C)^2}{2(sd_c{}^2+(Wsd_q)^2-(2W\rho sd_c sd_q)}} - Z_\alpha$$

where n represents the expected sample size per group; Z_α, the z-statistic for the α error (Z_α =1.96 for α=0.05); Z_β, the z-statistic for the β error (Z_β =0.84 for β=0.2); sd_c, the standard deviation for the cost in each treatment group; sd_q, the standard deviation for the effect in each treatment group; W, the willingness to pay threshold; ρ, the correlation between difference in costs and effects; C, expected difference in costs; and Q, the expected difference in effects.

Table 10.1.2 reports the power of the study using $n = 95$ (the final sample size of each group) and the observed final values from Chapter 10 at varying WTP thresholds.

Table 10.1.2 – Study power with $n = 95$ per group at varying WTP thresholds

WTP threshold	Power
$0 per QALY gained	0.77
$50000 per QALY gained	0.73
$100000 per QALY gained	0.69

Z_α, the z-statistic for the α error (Z_α =1.96 for α=0.05); Z_β, the z-statistic for the β error (Z_β =0.84 for β=0.2); sd_c, the standard deviation for the cost in each treatment group (sd_c = $6770); sd_q, the standard deviation for the effect in each treatment group (sd_q =0.003 QALYs); W, the willingness to pay threshold; ρ, the correlation between difference in costs and effects ($\rho = -0.405$); C, expected difference in costs (C = $2987); and Q, the expected difference in effects (Q = 0.002 QALYs).

Based on these results (Tables 10.1.1 and 10.1.2), this study is adequately powered to detect the hypothesized differences in costs and effects at a WTP threshold of $0/QALY, but power decreases at higher thresholds. This is because uncertainty around the NMB rises with the WTP threshold, since the difference in effectiveness is multiplied by an increasingly large λ (as $NMB = \lambda\Delta E - \Delta C$). Therefore a disproportionately large sample is required to reduce the uncertainty around the NMB, but this degree of certainty (i.e. 95% confidence) may not always be necessary. A decision maker may be content to reduce costs without compromising outcomes, and therefore set the WTP threshold at $0 per QALY, as was the case in our study. The difference in effectiveness (measured in terms of QALYs) between the ERP and conventional care groups was minimal, yet the difference in costs was of much greater magnitude in favor of the ERP. This suffices to show that the ERP is cost-effective, even in the absence of statistical significance at higher WTP thresholds.

10.2 APPENDIX 2 – Results of net monetary benefit regression

In order to control for comorbid status (as well as other factors), an additional analysis using the net monetary benefit (NMB) framework was performed. NMB is calculated as $\lambda \times \Delta E - \Delta C$, where λ denotes the willingness to pay threshold, ΔE, the difference in QALYs, and ΔC, the difference in costs. A positive NMB implies that the additional value of a new treatment is more than the extra cost, and therefore is cost-effective. NMB regression combines cost-effectiveness methodology with regression techniques (in this case, multiple linear regression).[240] This analysis allowed for the calculation of the independent effect of the ERP (see Table 10.2.1), which still remained statistically and clinically significant. Since there are no universally accepted willingness to pay threshold values, three different λ were used for NMB regression: \$0, \$50,000, and \$100,000 per additional QALY gained.[241] Uncertainty around point estimates for NMB regression were calculated using the bootstrap method described above. The results of the NMB regression also demonstrate that the ERP had a net beneficial effect, independent of age, comorbidities, laparoscopy, presence of a stoma, complications, and employement status (Chapter 8, Table 5). The number of comorbidities, preoperative employment, and complications all had a negative NMB. None of the interaction terms between the ERP variable and other covariates in the model were significant, suggesting that the ERP was cost-effective amongst all subgroups. This analysis corroborated the results of the subgroup analyses in Chapter 8.

Table 10.2.1 – Results of net monetary benefit (NMB) regression. Note that NMB > 0 denotes a beneficial effect. λ = willingness to pay threshold. 95% confidence intervals (CI) are derived from 2.5th and 97.5th percentiles of 10,000 bootstrap replications.

	NMB with λ = $0 (95% CI)	**NMB with λ = $50,000 (95% CI)**	**NMB with λ = $100,000 (95% CI)**
ERP (vs. CC)	2062 (378, 5055)	2128 (432, 5406)	2162 (491, 5275)
Age, per year increase	22 (-45, 103)	37 (-38, 116)	47 (-36, 128)
Charlson Comorbidity Index, per point increase	-821 (-1392, -204)	-849 (-1494, -194)	-880 (-1509, -154)
Employed (vs. non-employed)	-5981 (-8088, -3506)	-5731 (-8006, -3390)	-5571 (-7884, -3009)
Laparoscopy (vs. open)	69 (-2201, 2129)	156 (-2256, 1863)	298 (-1650, 2337)
Stoma (vs. no stoma)	-2867 (-5693, 958)	-2843 (-6496, 803)	-2822 (-6194, 899)
Comprehensive Complication Index, per point increase	-344 (-582, -180)	-355 (-592, -187)	-367 (-584, -196)

10.2 APPENDIX 3 – Clinical outcomes stratified by surgical approach

Table 10.3.1 – Clinical outcomes stratified by surgical approach (laparoscopy versus open) within the ERP and conventional care groups

	Laparoscopy	**Open**	***p*-value**
Mean total hospitalization, days (SD)			
ERAS group	6.0 (5.1)	9.6 (9.5)	0.042
Conventional care group	6.6 (5.1)	12.6 (15.5)	0.019
60-day emergency room visits			
ERAS group	20% (14/71)	21% (5/24)	0.906
Conventional care group	18% (8/45)	18% (9/50)	0.977
60-day readmissions			
ERAS group	13% (9/71)	13% (3/24)	0.982
Conventional care group	11% (5/45)	10% (5/50)	0.860
60-day complications			
ERAS group	41% (29/71)	38% (9/24)	0.772
Conventional care group	29% (13/45)	56% (28/50)	0.008

11. REFERENCES

1. Schilling PL, Dimick JB, Birkmeyer JD. Prioritizing quality improvement in general surgery. J Am Coll Surg 2008;207:698-704.
2. Kehlet H. Future perspectives and research initiatives in fast-track surgery. Langenbecks Arch Surg 2006;391:495-8.
3. Fearon KC, Ljungqvist O, Von Meyenfeldt M, et al. Enhanced recovery after surgery: a consensus review of clinical care for patients undergoing colonic resection. Clin Nutr 2005;24:466-77.
4. Moiniche S, Bulow S, Hesselfeldt P, Hestbaek A, Kehlet H. Convalescence and hospital stay after colonic surgery with balanced analgesia, early oral feeding, and enforced mobilisation. Eur J Surg 1995;161:283-8.
5. Bardram L, Funch-Jensen P, Jensen P, Crawford ME, Kehlet H. Recovery after laparoscopic colonic surgery with epidural analgesia, and early oral nutrition and mobilisation. Lancet 1995;345:763-4.
6. Basse L, Hjort Jakobsen D, Billesbolle P, Werner M, Kehlet H. A clinical pathway to accelerate recovery after colonic resection. Ann Surg 2000;232:51-7.
7. Bardram L, Funch-Jensen P, Kehlet H. Rapid rehabilitation in elderly patients after laparoscopic colonic resection. Br J Surg 2000;87:1540-5.
8. Gustafsson UO, Scott MJ, Schwenk W, et al. Guidelines for perioperative care in elective colonic surgery: Enhanced Recovery After Surgery (ERAS(R)) Society recommendations. Clin Nutr 2012;31:783-800.
9. Feldman LS, Baldini G, Lee L, Carli F. Enhanced Recovery Pathways: Organization of evidence-based, fast-track perioperative care. ACS Surgery: Principles and Practices Hamilton, ON: Decker Intellectual Properties; 2013.
10. Spanjersberg WR, Reurings J, Keus F, van Laarhoven CJ. Fast track surgery versus conventional recovery strategies for colorectal surgery. Cochrane Database Syst Rev 2011:CD007635.
11. Zhuang CL, Ye XZ, Zhang XD, Chen BC, Yu Z. Enhanced recovery after surgery programs versus traditional care for colorectal surgery: a meta-analysis of randomized controlled trials. Dis Colon Rectum 2013;56:667-78.
12. Jones C, Kelliher L, Dickinson M, et al. Randomized clinical trial on enhanced recovery versus standard care following open liver resection. Br J Surg 2013;100:1015-24.
13. Lemanu DP, Singh PP, Berridge K, et al. Randomized clinical trial of enhanced recovery versus standard care after laparoscopic sleeve gastrectomy. Br J Surg 2013;100:482-9.
14. Kehlet H, Buchler MW, Beart RW, Jr., Billingham RP, Williamson R. Care after colonic operation--is it evidence-based? Results from a multinational survey in Europe and the United States. J Am Coll Surg 2006;202:45-54.
15. Carter F, Kennedy RH. Setting up an Enhanced Recovery Programme. In: Francis N, Kennedy RH, Ljungqvist O, Mythen MG, eds. Manual of Fast Track Recovery for Colorectal Surgery. London: Springer; 2012.
16. Kehlet H, Wilmore DW. Multimodal strategies to improve surgical outcome. Am J Surg 2002;183:630-41.
17. Emanuel EJ. Where are the health care cost savings? Jama 2012;307:39-40.
18. Health Expenditure, total (% of GDP). The World Bank, 2014. (Accessed July 14, 2014, at http://data.worldbank.org/indicator/SH.XPD.TOTL.ZS.)

19. Owens DK, Qaseem A, Chou R, Shekelle P, Clinical Guidelines Committee of the American College of P. High-value, cost-conscious health care: concepts for clinicians to evaluate the benefits, harms, and costs of medical interventions. Ann Intern Med 2011;154:174-80.
20. Taheri PA, Butz DA, Greenfield LJ. Length of stay has minimal impact on the cost of hospital admission. J Am Coll Surg 2000;191:123-30.
21. Dimick JB, Pronovost PJ, Cowan JA, Lipsett PA. Complications and costs after high-risk surgery: where should we focus quality improvement initiatives? J Am Coll Surg 2003;196:671-8.
22. Lee L, Li C, Landry T, et al. A Systematic Review of Economic Evaluations of Enhanced Recovery Pathways for Colorectal Surgery. Ann Surg 2013.
23. Mowatt G, Houston G, Hernandez R, et al. Systematic review of the clinical effectiveness and cost-effectiveness of oesophageal Doppler monitoring in critically ill and high-risk surgical patients. Health Technol Assess 2009;13:iii-iv, ix-xii, 1-95.
24. Neville A, Lee L, Antonescu I, et al. Systematic review of outcomes used to evaluate enhanced recovery after surgery. Br J Surg 2013;In press.
25. Kehlet H, Wilmore DW. Evidence-based surgical care and the evolution of fast-track surgery. Ann Surg 2008;248:189-98.
26. Lassen K, Soop M, Nygren J, et al. Consensus review of optimal perioperative care in colorectal surgery: Enhanced Recovery After Surgery (ERAS) Group recommendations. Arch Surg 2009;144:961-9.
27. Varadhan KK, Neal KR, Dejong CH, Fearon KC, Ljungqvist O, Lobo DN. The enhanced recovery after surgery (ERAS) pathway for patients undergoing major elective open colorectal surgery: a meta-analysis of randomized controlled trials. Clin Nutr 2010;29:434-40.
28. Vlug MS, Wind J, Hollmann MW, et al. Laparoscopy in combination with fast track multimodal management is the best perioperative strategy in patients undergoing colonic surgery: a randomized clinical trial (LAFA-study). Ann Surg 2011;254:868-75.
29. Ren L, Zhu D, Wei Y, et al. Enhanced Recovery After Surgery (ERAS) program attenuates stress and accelerates recovery in patients after radical resection for colorectal cancer: a prospective randomized controlled trial. World J Surg 2012;36:407-14.
30. Carey K, Stefos T, Shibei Z, Borzecki AM, Rosen AK. Excess costs attributable to postoperative complications. Medical care research and review : MCRR 2011;68:490-503.
31. Sammour T, Zargar-Shoshtari K, Bhat A, Kahokehr A, Hill AG. A programme of Enhanced Recovery After Surgery (ERAS) is a cost-effective intervention in elective colonic surgery. The New Zealand medical journal 2010;123:61-70.
32. Kehlet H. Multimodal approach to control postoperative pathophysiology and rehabilitation. Br J Anaesth 1997;78:606-17.
33. Evers S, Goossens M, de Vet H, van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. International journal of technology assessment in health care 2005;21:240-5.
34. Chiou CF, Hay JW, Wallace JF, et al. Development and validation of a grading system for the quality of cost-effectiveness studies. Medical care 2003;41:32-44.
35. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. Bmj 1996;313:275-83.

36. Economic and Clinical Evaluation of Enhanced Recovery in Colorectal Surgery; NCT01610726. ClinicalTrials.gov, 2012. (Accessed October 16, 2012, at http://clinicaltrials.gov/ct2/show/NCT01610726.)
37. Fast-track Perioperative Program for Laparoscopic Colorectal Surgery; NCT01341366. ClinicalTrials.gov, 2011. (Accessed October 15, 2012, at http://clinicaltrials.gov/ct2/show/NCT01341366.)
38. Study Comparing Conventional and Fast Track Multi-Discipline Treatment Interventions for Colorectal Cancer; NCT01080547. ClinicalTrials.gov, 2010. (Accessed October 15, 2012, at http://clinicaltrials.gov/ct2/show/NCT01080547.)
39. Reurings JC, Spanjersberg WR, Oostvogel HJ, et al. A prospective cohort study to investigate cost-minimisation, of Traditional open, open fAst track recovery and laParoscopic fASt track multimodal management, for surgical patients with colon carcinomas (TAPAS study). BMC surgery 2010;10:18.
40. Kennedy RH, Francis A, Dutton S, et al. EnROL: A multicentre randomised trial of conventional versus laparoscopic surgery for colorectal cancer within an enhanced recovery programme. BMC cancer 2012;12:181.
41. Archibald LH, Ott MJ, Gale CM, Zhang J, Peters MS, Stroud GK. Enhanced recovery after colon surgery in a community hospital system. Dis Colon Rectum 2011;54:840-5.
42. Kariv Y, Delaney CP, Senagore AJ, et al. Clinical outcomes and cost analysis of a "fast track" postoperative care pathway for ileal pouch-anal anastomosis: a case control study. Dis Colon Rectum 2007;50:137-46.
43. King PM, Blazeby JM, Ewings P, et al. The influence of an enhanced recovery programme on clinical outcomes, costs and quality of life after surgery for colorectal cancer. Colorectal Dis 2006;8:506-13.
44. Bosio RM, Smith BM, Aybar PS, Senagore AJ. Implementation of laparoscopic colectomy with fast-track care in an academic medical center: benefits of a fully ascended learning curve and specialty expertise. Am J Surg 2007;193:413-5; discussion 5-6.
45. Health technology assessment of fast tracking colorectal surgery. DACEHTA, Danish Centre for Evaluation and Health Technology Assessment, 2005. (Accessed August 17, 2012, at http://www.sst.dk/publ/Publ2005/CEMTV/Acc_kolonkirurgi/Acc_kolonkir_patientforloeb.pdf.)
46. Jurowich CF, Reibetanz J, Krajinovic K, et al. [Cost analysis of the fast track concept in elective colonic surgery]. Zentralbl Chir 2011;136:256-63.
47. Stephen AE, Berger DL. Shortened length of stay and hospital cost reduction with implementation of an accelerated clinical care pathway after elective colon resection. Surgery 2003;133:277-82.
48. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Economic evaluation using patient-level data. Methods for the Economic Evaluation of Health Care Programmes. 3rd edition ed. New York: Oxford University Press; 2005.
49. Ahmed J, Khan S, Gatt M, Kallam R, MacFie J. Compliance with enhanced recovery programmes in elective colorectal surgery. Br J Surg 2010;97:754-8.
50. Maessen J, Dejong CH, Hausel J, et al. A protocol is not enough to implement an enhanced recovery programme for colorectal resection. Br J Surg 2007;94:224-31.
51. Mullins CD. Double counting and the reporting of cost per event avoided. Clinical therapeutics 2006;28:602-3.
52. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-effectiveness in Health and Medicine. Jama 1996;276:1253-8.

53. Neumann PJ. Costing and perspective in published cost-effectiveness analysis. Medical care 2009;47:S28-32.
54. Riewpaiboon A, Malaroje S, Kongsawatt S. Effect of costing methods on unit cost of hospital medical services. Tropical medicine & international health : TM & IH 2007;12:554-63.
55. Barnett PG. An improved set of standards for finding cost for cost-effectiveness analysis. Medical care 2009;47:S82-8.
56. Briggs A, Gray A. The distribution of health care costs and their statistical analysis for economic evaluation. Journal of health services research & policy 1998;3:233-45.
57. Gerkens S, Crott R, Cleemput I, et al. Comparison of three instruments assessing the quality of economic evaluations: a practical exercise on economic evaluations of the surgical treatment of obesity. International journal of technology assessment in health care 2008;24:318-25.
58. Jefferson T, Demicheli V, Vale L. Quality of systematic reviews of economic evaluations in health care. Jama 2002;287:2809-12.
59. Drummond MF, Sculpher MJ, Torrance G, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes. 3rd ed. ed. New York: Oxford University Press; 2005.
60. Fast track inpatient and ambulatory surgery. 2008. (Accessed February 3, 2013, at http://acssurgery.com/.)
61. Perelman J, Closon MC. Impact of socioeconomic factors on in-patient length of stay and their consequences in per case hospital payment systems. Journal of health services research & policy 2011;16:197-202.
62. Bruce J, Russell EM, Mollison J, Krukowski ZH. The measurement and monitoring of surgical adverse events. Health Technol Assess 2001;5:1-194.
63. Kleinbeck SV, Hoffart N. Outpatient recovery after laparoscopic cholecystectomy. AORN journal 1994;60:394, 7-8, 401-2.
64. Allvin R, Berg K, Idvall E, Nilsson U. Postoperative recovery: a concept analysis. Journal of advanced nursing 2007;57:552-8.
65. Moore FD. Getting well: the biology of surgical convalescence. Annals of the New York Academy of Sciences 1958;73:387-400.
66. Ljungqvist O. Jonathan E. Rhoads lecture 2011: Insulin resistance and enhanced recovery after surgery. JPEN Journal of parenteral and enteral nutrition 2012;36:389-98.
67. Veenhof AA, Vlug MS, van der Pas MH, et al. Surgical stress response and postoperative immune function after laparoscopy or open surgery with fast track or standard perioperative care: a randomized trial. Ann Surg 2012;255:216-21.
68. Hall GM, Peerbhoy D, Shenkin A, Parker CJ, Salmon P. Relationship of the functional recovery after hip arthroplasty to the neuroendocrine and inflammatory responses. Br J Anaesth 2001;87:537-42.
69. Lawrence VA, Hazuda HP, Cornell JE, et al. Functional independence after major abdominal surgery in the elderly. J Am Coll Surg 2004;199:762-72.
70. Mayo NE, Feldman L, Scott S, et al. Impact of preoperative change in physical function on postoperative recovery: argument supporting prehabilitation for colorectal surgery. Surgery 2011;150:505-14.
71. Feldman LS, Kaneva P, Demyttenaere S, Carli F, Fried GM, Mayo NE. Validation of a physical activity questionnaire (CHAMPS) as an indicator of postoperative recovery after laparoscopic cholecystectomy. Surgery 2009;146:31-9.

72. Carli F, Mayo N. Measuring the outcome of surgical procedures: what are the challenges? Br J Anaesth 2001;87:531-3.
73. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002;166:111-7.
74. Moriello C, Mayo NE, Feldman L, Carli F. Validating the six-minute walk test as a measure of recovery after elective colon resection surgery. Arch Phys Med Rehabil 2008;89:1083-9.
75. Stewart AL, Mills KM, King AC, Haskell WL, Gillis D, Ritter PL. CHAMPS physical activity questionnaire for older adults: outcomes for interventions. Medicine and science in sports and exercise 2001;33:1126-41.
76. Quintana JM, Cabriada J, Arostegui I, et al. Health-related quality of life and appropriateness of cholecystectomy. Ann Surg 2005;241:110-8.
77. Kluivers KB, Riphagen I, Vierhout ME, Brolmann HA, de Vet HC. Systematic review on recovery specific quality-of-life instruments. Surgery 2008;143:206-15.
78. Bartels SA, Vlug MS, Ubbink DT, Bemelman WA. Quality of life after laparoscopic and open colorectal surgery: a systematic review. World J Gastroenterol 2010;16:5035-41.
79. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. Jama 1995;273:59-65.
80. Barclay-Goddard R, Epstein JD, Mayo NE. Response shift: a brief overview and proposed research priorities. Qual Life Res 2009;18:335-46.
81. Paddison JS, Sammour T, Kahokehr A, Zargar-Shoshtari K, Hill AG. Development and validation of the Surgical Recovery Scale (SRS). J Surg Res 2011;167:e85-91.
82. Urbach DR, Harnish JL, McIlroy JH, Streiner DL. A measure of quality of life after abdominal surgery. Qual Life Res 2006;15:1053-61.
83. Bergman S, Feldman LS, Mayo NE, et al. Measuring surgical recovery: the study of laparoscopic live donor nephrectomy. Am J Transplant 2005;5:2489-95.
84. Bartels SA, Vlug MS, Ubbink DT, Bemelman WA. Quality of life after laparoscopic and open colorectal surgery: a systematic review. World J Gastroenterol 2010;16:5035-41.
85. Korolija D, Sauerland S, Wood-Dauphinee S, et al. Evaluation of quality of life after laparoscopic surgery: evidence-based guidelines of the European Association for Endoscopic Surgery. Surg Endosc 2004;18:879-97.
86. Mayo NE, Hum S, Kuspinar A. Methods and measures: what's new for MS? Multiple sclerosis 2013;19:709-13.
87. Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. Arch Phys Med Rehabil 1989;70:857-60.
88. Prieto L, Alonso J, Lamarca R. Classical Test Theory versus Rasch analysis for quality of life questionnaire reduction. Health and quality of life outcomes 2003;1:27.
89. Aldrete JA, Kroulik D. A postanesthetic recovery score. Anesth Analg 1970;49:924-34.
90. Myles PS, Hunt JO, Nightingale CE, et al. Development and psychometric testing of a quality of recovery score after general anesthesia and surgery in adults. Anesth Analg 1999;88:83-90.
91. Lee L, Elfassy N, Li C, et al. Valuing postoperative recovery: validation of the SF-6D health-state utility. J Surg Res 2013;184:108-14.
92. Bergman S, Feldman LS, Barkun JS. Evaluating surgical outcomes. Surg Clin North Am 2006;86:129-49, x.

93. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. Journal of health economics 2002;21:271-92.
94. Russell LB, Gold MR, Siegel JE, Daniels N, Weinstein MC. The role of cost-effectiveness analysis in health and medicine. Panel on Cost-Effectiveness in Health and Medicine. Jama 1996;276:1172-7.
95. Harrison MJ, Ahmad Y, Haque S, et al. Construct and criterion validity of the short form-6D utility measure in patients with systemic lupus erythematosus. J Rheumatol 2012;39:735-42.
96. Dolan P. Modeling valuations for EuroQol health states. Medical care 1997;35:1095-108.
97. Horsman J, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI): concepts, measurement properties and applications. Health and quality of life outcomes 2003;1:54.
98. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Medical care 1992;30:473-83.
99. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993;85:365-76.
100. Brazier J, Czoski-Murray C, Roberts J, Brown M, Symonds T, Kelleher C. Estimation of a preference-based index from a condition-specific measure: the King's Health Questionnaire. Medical decision making : an international journal of the Society for Medical Decision Making 2008;28:113-26.
101. Soler ZM, Wittenberg E, Schlosser RJ, Mace JC, Smith TL. Health state utility values in patients undergoing endoscopic sinus surgery. The Laryngoscope 2011;121:2672-8.
102. Atroshi I, Gummesson C, McCabe SJ, Ornstein E. The SF-6D health utility index in carpal tunnel syndrome. The Journal of hand surgery, European volume 2007;32:198-202.
103. Khanna D, Furst DE, Wong WK, et al. Reliability, validity, and minimally important differences of the SF-6D in systemic sclerosis. Qual Life Res 2007;16:1083-92.
104. Gaujoux-Viala C, Rat AC, Guillemin F, et al. Responsiveness of EQ-5D and SF-6D in patients with early arthritis: results from the ESPOIR cohort. Annals of the rheumatic diseases 2012.
105. Li C, Carli F, Lee L, et al. Impact of a trimodal prehabilitation program on functional recovery after colorectal cancer surgery: a pilot study. Surg Endosc 2013;27:1072-82.
106. Carli F, Charlebois P, Stein B, et al. Randomized clinical trial of prehabilitation in colorectal surgery. Br J Surg 2010;97:1187-97.
107. Wongyingsinn M, Baldini G, Charlebois P, Liberman S, Stein B, Carli F. Intravenous lidocaine versus thoracic epidural analgesia: a randomized controlled trial in patients undergoing laparoscopic colorectal surgery using an enhanced recovery program. Reg Anesth Pain Med 2011;36:241-8.
108. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg 2004;240:205-13.
109. Ware JE, Kosinski M, Dewey JE. How to score version two of the SF-36 Health Survey. Lincoln, RI: QualityMetric, Incorporated; 2000.
110. Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. Journal of clinical epidemiology 2006;59:1087-91.
111. Moons KG, Donders RA, Stijnen T, Harrell FE, Jr. Using the outcome for imputation of missing predictor values was preferred. Journal of clinical epidemiology 2006;59:1092-101.

112. Rubin D. Multiple imputation for non response in surveys. New York: Wiley; 1987.
113. Cohen J. Statistical power for the behavioral sciences. New York: Academic Press; 1977.
114. Fryback DG, Dunham NC, Palta M, et al. US norms for six generic health-related quality-of-life indexes from the National Health Measurement study. Medical care 2007;45:1162-70.
115. Marra CA, Woolcott JC, Kopec JA, et al. A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. Soc Sci Med 2005;60:1571-82.
116. Carli F, Zavorsky GS. Optimizing functional exercise capacity in the elderly surgical population. Current opinion in clinical nutrition and metabolic care 2005;8:23-32.
117. Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. Qual Life Res 2005;14:1523-32.
118. Maartense S, Dunker MS, Slors JF, et al. Laparoscopic-assisted versus open ileocolic resection for Crohn's disease: a randomized trial. Ann Surg 2006;243:143-9; discussion 50-3.
119. Maartense S, Dunker MS, Slors JF, et al. Hand-assisted laparoscopic versus open restorative proctocolectomy with ileal pouch anal anastomosis: a randomized trial. Ann Surg 2004;240:984-91; discussion 91-2.
120. Klarenbeek BR, Veenhof AA, Bergamaschi R, et al. Laparoscopic sigmoid resection for diverticulitis decreases major morbidity rates: a randomized control trial: short-term results of the Sigma Trial. Ann Surg 2009;249:39-44.
121. Janson M, Lindholm E, Anderberg B, Haglind E. Randomized trial of health-related quality of life after open and laparoscopic surgery for colon cancer. Surg Endosc 2007;21:747-53.
122. King PM, Blazeby JM, Ewings P, et al. Randomized clinical trial comparing laparoscopic and open surgery for colorectal cancer within an enhanced recovery programme. Br J Surg 2006;93:300-8.
123. Eshuis EJ, Slors JF, Stokkers PC, et al. Long-term outcomes following laparoscopically assisted versus open ileocolic resection for Crohn's disease. Br J Surg 2010;97:563-8.
124. Arthur HM, Daniels C, McKelvie R, Hirsh J, Rush B. Effect of a preoperative intervention on preoperative and postoperative outcomes in low-risk patients awaiting elective coronary artery bypass graft surgery. A randomized, controlled trial. Ann Intern Med 2000;133:253-62.
125. Beattie AH, Prach AT, Baxter JP, Pennington CR. A randomised controlled trial evaluating the use of enteral nutritional supplements postoperatively in malnourished surgical patients. Gut 2000;46:813-8.
126. Carli F, Mayo N, Klubien K, Schricker T, Trudel J, Belliveau P. Epidural analgesia enhances functional exercise capacity and health-related quality of life after colonic surgery: results of a randomized trial. Anesthesiology 2002;97:540-9.
127. Paddison JS, Booth RJ, Hill AG, Cameron LD. Comprehensive assessment of peri-operative fatigue: development of the Identity-Consequence Fatigue Scale. J Psychosom Res 2006;60:615-22.
128. Hofhuis JG, Spronk PE, van Stel HF, Schrijvers GJ, Rommes JH, Bakker J. The impact of critical illness on perceived health-related quality of life during ICU treatment, hospital stay, and after hospital discharge: a long-term follow-up study. Chest 2008;133:377-85.

129. Wood AM, White IR, Hillsdon M, Carpenter J. Comparison of imputation and modelling methods in the analysis of a physical activity trial with missing outcomes. International journal of epidemiology 2005;34:89-99.
130. Fielding S, Fayers P, Ramsay C. Predicting missing quality of life data that were later recovered: an empirical comparison of approaches. Clin Trials 2010;7:333-42.
131. Shrive FM, Stuart H, Quan H, Ghali WA. Dealing with missing data in a multi-question depression scale: a comparison of imputation methods. BMC medical research methodology 2006;6:57.
132. Boyle MH, Furlong W, Feeny D, Torrance GW, Hatcher J. Reliability of the Health Utilities Index--Mark III used in the 1991 cycle 6 Canadian General Social Survey Health Questionnaire. Qual Life Res 1995;4:249-57.
133. Kaplan RM, Sieber WJ, Ganiats TG. The Quality of Well-Being Scale: comparison of the interviewer-administered version with a self-administered questionnaire. Psych Health 1997;12:783-91.
134. Hawthorne G, Richardson J, Osborne R. The Assessment of Quality of Life (AQoL) instrument: a psychometric measure of health-related quality of life. Qual Life Res 1999;8:209-24.
135. Whitehurst DG, Bryan S, Lewis M. Systematic review and empirical comparison of contemporaneous EQ-5D and SF-6D group mean scores. Medical decision making : an international journal of the Society for Medical Decision Making 2011;31:E34-44.
136. Brazier J, Deverill M. A checklist for judging preference-based measures of health related quality of life: learning from psychometrics. Health economics 1999;8:41-51.
137. Lee L, Tran T, Mayo NE, Carli F, Feldman LS. What does it really mean to "recover" from an operation? Surgery 2013.
138. Rasanen P, Roine E, Sintonen H, Semberg-Konttinen V, Ryynanen OP, Roine R. Use of quality-adjusted life years for the estimation of effectiveness of health care: A systematic literature review. International journal of technology assessment in health care 2006;22:235-41.
139. Joore M, Brunenberg D, Nelemans P, et al. The impact of differences in EQ-5D and SF-6D utility scores on the acceptability of cost-utility ratios: results across five trial-based cost-utility studies. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 2010;13:222-9.
140. Sinha S, Sinha S, Ashby E, Jayaram R, Grocott MP. Quality of reporting in randomized trials published in high-quality surgical journals. J Am Coll Surg 2009;209:565-71 e1.
141. Martin RC, 2nd, Brennan MF, Jaques DP. Quality of complication reporting in the surgical literature. Ann Surg 2002;235:803-13.
142. Sach TH, Barton GR, Jenkinson C, Doherty M, Avery AJ, Muir KR. Comparing cost-utility estimates: does the choice of EQ-5D or SF-6D matter? Medical care 2009;47:889-94.
143. Barton GR, Sach TH, Avery AJ, Doherty M, Jenkinson C, Muir KR. Comparing the performance of the EQ-5D and SF-6D when measuring the benefits of alleviating knee pain. Cost effectiveness and resource allocation : C/E 2009;7:12.
144. Marra CA, Rashidi AA, Guh D, et al. Are indirect utility measures reliable and responsive in rheumatoid arthritis patients? Qual Life Res 2005;14:1333-44.
145. Charlson M, Szatrowski TP, Peterson J, Gold J. Validation of a combined comorbidity index. Journal of clinical epidemiology 1994;47:1245-51.
146. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-Item Health Survey 1.0. Health economics 1993;2:217-27.

147. EuroQol G. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199-208.
148. Hanmer J, Lawrence WF, Anderson JP, Kaplan RM, Fryback DG. Report of nationally representative values for the noninstitutionalized US adult population for 7 health-related quality-of-life scores. Medical decision making : an international journal of the Society for Medical Decision Making 2006;26:391-400.
149. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-10.
150. Bland JM, Altman DG. Comparing methods of measurement: why plotting difference against standard method is misleading. Lancet 1995;346:1085-7.
151. Lee L, Tran T, Mayo NE, Carli F, Feldman LS. What does it really mean to "recover" from an operation? Surgery 2013.
152. Janssen MF, Pickard AS, Golicki D, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. Qual Life Res 2013;22:1717-27.
153. Bryan S, Longworth L. Measuring health-related utility: why the disparity between EQ-5D and SF-6D? The European journal of health economics : HEPAC : health economics in prevention and care 2005;6:253-60.
154. Dworkin RH, Turk DC, Wyrwich KW, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. The journal of pain : official journal of the American Pain Society 2008;9:105-21.
155. Cano SJ, Hobart JC. The problem with health measurement. Patient preference and adherence 2011;5:279-90.
156. Davis JC, Liu-Ambrose T, Khan KM, Robertson MC, Marra CA. SF-6D and EQ-5D result in widely divergent incremental cost-effectiveness ratios in a clinical trial of older women: implications for health policy decisions. Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA 2012;23:1849-57.
157. Wilson TR, Alexander DJ, Kind P. Measurement of health-related quality of life in the early follow-up of colon and rectal cancer. Dis Colon Rectum 2006;49:1692-702.
158. Johnson JA, Luo N, Shaw JW, Kind P, Coons SJ. Valuations of EQ-5D health states: are the United States and United Kingdom different? Medical care 2005;43:221-8.
159. Husted H. Fast-track hip and knee arthroplasty: clinical and organizational aspects. Acta orthopaedica Supplementum 2012;83:1-39.
160. Arumainayagam N, McGrath J, Jefferson KP, Gillatt DA. Introduction of an enhanced recovery protocol for radical cystectomy. BJU Int 2008;101:698-701.
161. Ottesen M, Sorensen M, Rasmussen Y, Smidt-Jensen S, Kehlet H, Ottesen B. Fast track vaginal surgery. Acta obstetricia et gynecologica Scandinavica 2002;81:138-46.
162. Findlay JM, Gillies RS, Millo J, Sgromo B, Marshall RE, Maynard ND. Enhanced Recovery for Esophagectomy: A Systematic Review and Evidence-Based Guidelines. Ann Surg 2013.
163. Dorcaratto D, Grande L, Pera M. Enhanced Recovery in Gastrointestinal Surgery: Upper Gastrointestinal Surgery. Dig Surg 2013;30:70-8.
164. Coolsen MM, van Dam RM, van der Wilt AA, Slim K, Lassen K, Dejong CH. Systematic review and meta-analysis of enhanced recovery after pancreatic surgery with particular emphasis on pancreaticoduodenectomies. World J Surg 2013;37:1909-18.

165. Coolsen MM, Wong-Lun-Hing EM, van Dam RM, et al. A systematic review of outcomes in patients undergoing liver surgery in an enhanced recovery after surgery pathways. HPB : the official journal of the International Hepato Pancreato Biliary Association 2013;15:245-51.
166. Vanounou T, Pratt W, Fischer JE, Vollmer CM, Jr., Callery MP. Deviation-based cost modeling: a novel model to evaluate the clinical and economic impact of clinical pathways. J Am Coll Surg 2007;204:570-9.
167. Li C, Ferri LE, Mulder DS, et al. An enhanced recovery pathway decreases duration of stay after esophagectomy. Surgery 2012;152:606-14; discussion 14-6.
168. Cancer trends progress report - 2011/2012 update. National Cancer Institute. (Accessed November 5, 2012, at http://progressreport.cancer.gov/doc_detail.asp?pid=1&did=2009&chid=95&coid=926&mid.)
169. Finks JF, Osborne NH, Birkmeyer JD. Trends in hospital volume and operative mortality for high-risk surgery. N Engl J Med 2011;364:2128-37.
170. Allareddy V, Ward MM, Allareddy V, Konety BR. Effect of meeting Leapfrog volume thresholds on complication rates following complex surgical procedures. Ann Surg 2010;251:377-83.
171. Varghese TK, Jr., Wood DE, Farjah F, et al. Variation in esophagectomy outcomes in hospitals meeting Leapfrog volume outcome standards. Ann Thorac Surg 2011;91:1003-9; discussion 9-10.
172. Davenport DL, Henderson WG, Khuri SF, Mentzer RM, Jr. Preoperative risk factors and surgical complexity are more predictive of costs than postoperative complications: a case study using the National Surgical Quality Improvement Program (NSQIP) database. Ann Surg 2005;242:463-8; discussion 8-71.
173. Hammermeister KE, Shroyer AL, Sethi GK, Grover FL. Why it is important to demonstrate linkages between outcomes of care and processes and structures of care. Medical care 1995;33:S5-16.
174. Anderson CI, Nelson CS, Graham CF, et al. Disorganized care: the findings of an iterative, in-depth analysis of surgical morbidity and mortality. J Surg Res 2012;177:43-8.
175. Kehlet H. Fast-track colorectal surgery. Lancet 2008;371:791-3.
176. Husted H, Holm G. Fast track in total hip and knee arthroplasty--experiences from Hvidovre University Hospital, Denmark. Injury 2006;37 Suppl 5:S31-5.
177. Low DE, Kunz S, Schembre D, et al. Esophagectomy--it's not just about mortality anymore: standardized perioperative clinical pathways improve outcomes in patients with esophageal cancer. J Gastrointest Surg 2007;11:1395-402; discussion 402.
178. Preston SR, Markar SR, Baker CR, Soon Y, Singh S, Low DE. Impact of a multidisciplinary standardized clinical pathway on perioperative outcomes in patients with oesophageal cancer. Br J Surg 2013;100:105-12.
179. Cao S, Zhao G, Cui J, et al. Fast-track rehabilitation program and conventional care after esophagectomy: a retrospective controlled cohort study. Supportive care in cancer : official journal of the Multinational Association of Supportive Care in Cancer 2012.
180. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. Lancet 2003;362:1225-30.
181. Seely AJ, Ivanovic J, Threader J, et al. Systematic classification of morbidity and mortality after thoracic surgery. Ann Thorac Surg 2010;90:936-42; discussion 42.

182. Statistiques interactives. Health and Social Services Agency (ASSS) of Montreal. (Accessed October 15, 2012, at http://cmisstats.mtl.rtss.qc.ca/tabloclip/index.php?resource_s=com.ressources.m30.p900.)
183. Health Care Trend Rate. Canadian Institute of Actuaries. (Accessed January 23, 2012, at http://www.actuaries.ca/members/publications/2012/212031e.pdf.)
184. Current and Historical Rate Tables. XE. at http://www.xe.com/currencytables/?from=EUR&date=2011-12-31.)
185. Vanounou T, Pratt W, Fischer JE, Vollmer CM, Jr., Callery MP. Deviation-based cost modeling: a novel model to evaluate the clinical and economic impact of clinical pathways. J Am Coll Surg 2007;204:570-9.
186. Adamina M, Kehlet H, Tomlinson GA, Senagore AJ, Delaney CP. Enhanced recovery pathways optimize health outcomes and resource utilization: a meta-analysis of randomized controlled trials in colorectal surgery. Surgery 2011;149:830-40.
187. Jaeger JM, Collins SR, Blank RS. Anesthetic management for esophageal resection. Anesthesiology clinics 2012;30:731-47.
188. Cense HA, Lagarde SM, de Jong K, et al. Association of no epidural analgesia with postoperative morbidity and mortality after transthoracic esophageal cancer resection. J Am Coll Surg 2006;202:395-400.
189. Popping DM, Elia N, Marret E, Remy C, Tramer MR. Protective effects of epidural analgesia on pulmonary complications after abdominal and thoracic surgery: a meta-analysis. Arch Surg 2008;143:990-9; discussion 1000.
190. Neal JM, Wilcox RT, Allen HW, Low DE. Near-total esophagectomy: the influence of standardized multimodal management and intraoperative fluid restriction. Reg Anesth Pain Med 2003;28:328-34.
191. Lassen K, Coolsen MM, Slim K, et al. Guidelines for perioperative care for pancreaticoduodenectomy: Enhanced Recovery After Surgery (ERAS(R)) Society recommendations. Clin Nutr 2012;31:817-30.
192. Nygren J, Thacker J, Carli F, et al. Guidelines for perioperative care in elective rectal/pelvic surgery: Enhanced Recovery After Surgery (ERAS(R)) Society recommendations. Clin Nutr 2012;31:801-16.
193. Lassen K, Kjaeve J, Fetveit T, et al. Allowing normal food at will after major upper gastrointestinal surgery does not increase morbidity: a randomized multicenter trial. Ann Surg 2008;247:721-9.
194. Wheble GA, Benson RA, Khan OA. Is routine postoperative enteral feeding after oesophagectomy worthwhile? Interactive cardiovascular and thoracic surgery 2012;15:709-12.
195. Andersen HK, Lewis SJ, Thomas S. Early enteral nutrition within 24h of colorectal surgery versus later commencement of feeding for postoperative complications. Cochrane Database Syst Rev 2006:CD004080.
196. Michelet P, D'Journo XB, Roch A, et al. Protective ventilation influences systemic inflammation after esophagectomy: a randomized controlled study. Anesthesiology 2006;105:911-9.
197. Yuill KA, Richardson RA, Davidson HI, Garden OJ, Parks RW. The administration of an oral carbohydrate-containing fluid prior to major elective upper-gastrointestinal surgery preserves skeletal muscle mass postoperatively--a randomised clinical trial. Clin Nutr 2005;24:32-7.

198. Zehr KJ, Dawson PB, Yang SC, Heitmiller RF. Standardized clinical care pathways for major thoracic cases reduce hospital costs. Ann Thorac Surg 1998;66:914-9.
199. Statistics on Hospital Stays. Agency for Healthcare Research and Quality, 2013. (Accessed February 26, 2013, at http://hcupnet.ahrq.gov/HCUPnet.jsp.)
200. Schwarzbach M, Rossner E, Schattenberg T, Post S, Hohenberger P, Ronellenfitsch U. Effects of a clinical pathway of pulmonary lobectomy and bilobectomy on quality and cost of care. Langenbecks Arch Surg 2010;395:1139-46.
201. Dimick JB, Chen SL, Taheri PA, Henderson WG, Khuri SF, Campbell DA, Jr. Hospital costs associated with surgical complications: a report from the private-sector National Surgical Quality Improvement Program. J Am Coll Surg 2004;199:531-7.
202. Darer J, Pronovost P, Bass EB. Use and evaluation of critical pathways in hospitals. Effective clinical practice : ECP 2002;5:114-9.
203. Glick HA. Sample size and power for cost-effectiveness analysis (part 1). PharmacoEconomics 2011;29:189-98.
204. Lee L, Li C, Landry T, et al. A systematic review of economic evaluations of enhanced recovery pathways for colorectal surgery. Ann Surg 2014;259:670-6.
205. Slankamenac K, Graf R, Barkun J, Puhan MA, Clavien PA. The comprehensive complication index: a novel continuous scale to measure surgical morbidity. Ann Surg 2013;258:1-7.
206. Lee L, Mata J, Augustin BR, et al. A comparison of the validity of two indirect utility instruments as measures of postoperative recovery. J Surg Res 2014;190:79-86.
207. Billingham LJ, Abrams KR, Jones DR. Methods for the analysis of quality-of-life and survival data in health technology assessment. Health Technol Assess 1999;3:1-152.
208. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. Health economics 2005;14:487-96.
209. Normes et pratiques de gestion pour les centres hospitaliers. Santé et Services sociaux (Québec). (Accessed October 22, 2013, at http://msssa4.msss.gouv.qc.ca/fr/document/d26ngest.nsf/lcaCH?OpenView.)
210. Manuel de facturation. Régie de l'assurance maladie - Québec, 2013. (Accessed May 1, 2013, at http://www.ramq.gouv.qc.ca/fr/professionnels/medecins-specialistes/manuels/Pages/facturation.aspx.)
211. Longo CJ, Deber R, Fitch M, Williams AP, D'Souza D. An examination of cancer patients' monthly 'out-of-pocket' costs in Ontario, Canada. Eur J Cancer Care (Engl) 2007;16:500-7.
212. Longo CJ, Fitch M, Deber RB, Williams AP. Financial and family burden associated with cancer treatment in Ontario, Canada. Supportive care in cancer : official journal of the Multinational Association of Supportive Care in Cancer 2006;14:1077-85.
213. Individuals by total income level, by province and territory (Canada). Statistics Canada, 2013. 2014, at http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/famil105a-eng.htm.)
214. PPPs and exchange rates. Organization for Economic Co-operation and Development, 2013. (Accessed April 8, 2014, at http://stats.oecd.org/Index.aspx?DataSetCode=CPL.)
215. Koopmanschap MA, Rutten FF, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. Journal of health economics 1995;14:171-89.
216. Hopkins RB, Goeree R, Longo CJ. Estimating the national wage loss from cancer in Canada. Curr Oncol 2010;17:40-9.

217. Lee L, Li C, Robert N, et al. Economic impact of an enhanced recovery pathway for oesophagectomy. Br J Surg 2013;100:1326-34.
218. Janson M, Bjorholt I, Carlsson P, et al. Randomized clinical trial of the costs of open and laparoscopic surgery for colonic cancer. Br J Surg 2004;91:409-17.
219. Neville A, Lee L, Antonescu I, et al. Systematic review of outcomes used to evaluate enhanced recovery after surgery. Br J Surg 2014;101:159-71.
220. Aasa A, Hovback M, Bertero CM. The importance of preoperative information for patient participation in colorectal surgery care. Journal of clinical nursing 2013;22:1604-12.
221. Blazeby JM, Soulsby M, Winstone K, King PM, Bulley S, Kennedy RH. A qualitative evaluation of patients' experiences of an enhanced recovery programme for colorectal cancer. Colorectal Dis 2010;12:e236-42.
222. Hospital adjusted expenses per inpatient day. The Henry J. Kaiser Family Foundation. (Accessed April 27, 2014, at http://kff.org/other/state-indicator/expenses-per-inpatient-day/.)
223. Roulin D, Donadini A, Gander S, et al. Cost-effectiveness of the implementation of an enhanced recovery protocol for colorectal surgery. Br J Surg 2013;100:1108-14.
224. Schwenk W, Haase O, Neudecker J, Muller JM. Short term benefits for laparoscopic colorectal resection. Cochrane Database Syst Rev 2005:CD003145.
225. Kennedy RH, Francis EA, Wharton R, et al. Multicenter Randomized Controlled Trial of Conventional Versus Laparoscopic Surgery for Colorectal Cancer Within an Enhanced Recovery Programme: EnROL. J Clin Oncol 2014;32:1804-11.
226. Smith DM, Sherriff RL, Damschroder L, Loewenstein G, Ubel PA. Misremembering colostomies? Former patients give lower utility ratings than do current patients. Health psychology : official journal of the Division of Health Psychology, American Psychological Association 2006;25:688-95.
227. Delaney CP, Senagore AJ, Gerkin TM, et al. Association of surgical care practices with length of stay and use of clinical protocols after elective bowel resection: results of a national survey. Am J Surg 2010;199:299-304; discussion
228. Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. Br J Surg 1991;78:355-60.
229. Shao J, Sitter RR. Bootstrap for imputed survery data. J Am Stat Assoc 1996;91:1278-88.
230. Guidelines for the economic evaluation of health technologies: Canada. Canadian Agency for Drugs and Technologies in Health, 2006. (Accessed May 9, 2013, at http://www.cadth.ca/media/pdf/186_EconomicGuidelines_e.pdf.)
231. Tranmer JE, Guerriere DN, Ungar WJ, Coyte PC. Valuing patient and caregiver time: a review of the literature. PharmacoEconomics 2005;23:449-59.
232. Krol M, Brouwer W. How to estimate productivity costs in economic evaluations. PharmacoEconomics 2014;32:335-44.
233. Li C, Carli F, Lee L, et al. Impact of a trimodal prehabilitation program on functional recovery after colorectal cancer surgery: a pilot study. Surg Endosc 2012.
234. Smith EC, Ziogas A, Anton-Culver H. Delay in surgical treatment and survival after breast cancer diagnosis in young women by race/ethnicity. JAMA surgery 2013;148:516-23.
235. Bourgade V, Drouin SJ, Yates DR, et al. Impact of the length of time between diagnosis and surgical removal of urologic neoplasms on survival. World journal of urology 2014;32:475-9.
236. Amri R, Bordeianou LG, Sylla P, Berger DL. Treatment Delay in Surgically-Treated Colon Cancer: Does It Affect Outcomes? Ann Surg Oncol 2014.

237. Kahn KL, Keeler EB, Sherwood MJ, et al. Comparing outcomes of care before and after implementation of the DRG-based prospective payment system. Jama 1990;264:1984-8.
238. Porter ME. What is value in health care? N Engl J Med 2010;363:2477-81.
239. Glick HA. Sample size and power for cost-effectiveness analysis (Part 2): the effect of maximum willingness to pay. PharmacoEconomics 2011;29:287-96.
240. Hoch JS. Improving efficiency and value in palliative care with net benefit regression: an introduction to a simple method for cost-effectiveness analysis with person-level data. Journal of pain and symptom management 2009;38:54-61.
241. Laupacis A, Feeny D, Detsky AS, Tugwell PX. How attractive does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluations. Cmaj 1992;146:473-81.