Assessments

ARE POLICY DECISIONS ON SURGICAL PROCEDURES INFORMED BY ROBUST ECONOMIC EVIDENCE? A SYSTEMATIC REVIEW

Roberta Ara
ScHARR, Senior Research Fellow, School of Health and Related Research (ScHARR), Regent Court, 30 Regent Street, Sheffield, S1 4DA

Hasan Basarir
ScHARR, Research Fellow, The University of Sheffield, School of Health and Related Research (ScHARR), Regent Court, 30 Regent Street, Sheffield, S1 4DA

Anju D. Keetharuth
ScHARR, Research Associate, The University of Sheffield, School of Health and Related Research (ScHARR), Regent Court, 30 Regent Street, Sheffield, S1 4DA

Marco Barbieri
Centre for Health Economics (CHE), University of York, Heslington, York YO10 5DD, UK

Helen L.A. Weatherly
Senior Research Fellow, Centre for Health Economics (CHE), University of York, Heslington, York YO10 5DD, UK

Mark J.S. Sculpher
Professor of Health Economics, Deputy Director of the Policy Research Unit in Economic Evaluation of Health and Care Interventions (EEPRU), Centre for Health Economics (CHE), University of York, Heslington, York YO10 5DD, UK

Hashim Ahmed
MRC Clinician Scientist and Honorary Consultant Urological Surgeon, University College Hospital, 235 Euston Road, London, NW1 2BU

Steven Brown
Consultant Colorectal Surgeon, Sheffield Teaching Hospitals, NHS Foundation Trust, Northern General Hospital, Herries Road, Sheffield, S5 7AU

Objectives: The aim of this study was to examine the empirical and methodological cost-effectiveness evidence of surgical interventions for breast, colorectal, or prostate cancer.
Methods: A systematic search of seven databases including MEDLINE, EMBASE, and NHSEED, research registers, the NICE Web site and conference proceedings was conducted in April 2012. Study quality was assessed in terms of meeting essential, preferred and UK NICE specific requirements for economic evaluations.
Results: The seventeen (breast = 3, colorectal = 7, prostate = 7) included studies covered a broad range of settings (nine European; eight non-European) and six were published over 10 years ago. The populations, interventions and comparators were generally well defined. Very few studies were informed by literature reviews and few used synthesized clinical evidence. Although the interventions had potential differential effects on recurrence and mortality rates, some studies used relatively short time horizons. Univariate sensitivity analyses were reported in all studies but less than a third characterized all uncertainty with a probabilistic sensitivity analysis. Although a third of studies incorporated patients' health-related quality of life data, only four studies used social tariff values.
Conclusions: There is a dearth of recent robust evidence describing the cost-effectiveness of surgical interventions in the management of breast, colorectal and prostate cancers. Many of the recent publications did not satisfy essential methodological requirements such as using clinical evidence informed by a systematic review and synthesis. Given the ratio of potential benefit and harms associated with cancer surgery and the volume of resources consumed by these, there is an urgent need to increase economic evaluations of these technologies.

Key words: Economic evaluation, Technology assessment, Review, Surgical procedures, Neoplasms

Health technology assessments (HTA) provide a connection between the research and decision-making worlds and over the last 20 years have become central to informing healthcare policy de-

cision making in many settings worldwide (1). The HTA process involves a systematic evaluation of the direct (intended) and indirect (unintended) effects and impacts of healthcare technologies. Organizations such as the Pharmaceutical Benefits Advisory Committee (PBAC) in Australia, the Heath Care Insurance Board (CVZ) in the Netherlands, and the National Institute for Health and Care Excellence (NICE) in the United Kingdom now publish guidelines to ensure consistency and transparency in the HTA decision-making and reimbursement process (2–4).

Once the therapeutic value of novel interventions is established, evidence of value for money and efficiency is required before allocation of scarce healthcare resources in many settings. To ensure consistency and increase the quality of evidence submitted, NICE provides detail of a "reference case" for economic evidence used to support submissions (4). While some of NICE's requirements (such as discount rates and perspectives on costs) differ from those stipulated for non-UK settings, many are essential requirements for any economic evaluation in healthcare. These are clearly defined, relevant comparators; systematic literature reviews for health effects; synthesis of clinical evidence where applicable; suitable time horizons; health related quality of life (HRQoL) data reported by patients; and exploration of uncertainty (5–7). An often recommended criterion, which is not standard across the board, is that results are reported in terms of incremental cost per quality-adjusted life-year (QALY) whereby the HRQoL weights for the QALY are valued by the general population (4;8).

The last decade has seen a rapid growth in the volume of literature describing economic evaluations of interventions in healthcare. In particular the number of articles describing cost-effectiveness evaluations in cancer has increased substantially as new pharmaceuticals are brought to market. Surgery, radiotherapy, and chemotherapy are the main treatments for cancer (9), and considerable proportions of healthcare budgets are consumed by these interventions. In 2010–2011, cancer was the third largest category of the overall UK healthcare budget, accounting for 5.4 percent of the overall cost at £5.8 bn (EUR 7.24 bn) (10), while total economic costs attributed to colorectal, breast, and prostate cancers were estimated to be £1.6 bn (EUR 2.0 bn), £1.5 bn (EUR 1.87 bn) and £0.6 bn (EUR 0.75 bn), respectively (11). Although the clinical and cost-effectiveness of novel pharmaceutical interventions for cancer is well documented due to the submission processes for reimbursement in many settings, it is unclear whether existing and novel surgical and radiotherapy interventions are treated with the same rigorous scrutiny.

The objective of the study was to examine the methodological and empirical cost-effectiveness evidence specifically in relation to surgical interventions in breast, colorectal and prostate cancer. Relevant literature was identified through systematic searches and methodological standards were assessed using quality criteria, both essential and preferred, and informed by the NICE reference case.

METHODS

Search Strategy

Systematic searches were undertaken (April 2012) in the following databases: Medline, EMBASE, CDSR, NHSEED, HTA, DARE, and EconLit. The CEA registry, the NICE website, recent conference proceedings, and reference lists of any included studies and existing reviews were also searched. No language, publication, or date restrictions were applied to the searches. The full search strategy is available online (12).

Study Selection Strategy

Studies were selected for inclusion in two stages. An initial sieve was conducted on title and abstract and full articles were retrieved for any potentially relevant studies. Studies were included in the reviews if they assessed the cost-effectiveness of a surgical technique in individuals with either breast, colorectal, or prostate cancer, and presented results in the form of an incremental cost per QALY or an incremental cost per life-year [LY]). Data were extracted using a customized template designed to capture relevant information (Table 1–2). Studies were quality appraised (Figure 2) in terms of meeting essential (items 1–7: patient group and indication clearly defined, comparators clearly defined, effectiveness evidence based on a systematic review, appropriate time horizon, HRQoL data reported directly by patients or carers, probabilistic sensitivity analysis used to quantify uncertainty), preferred (items 8–9: cost-effectiveness analysis using QALYs, HRQoL preference data valuation by representative sample of the public), and UK NICE specific (items 10–14: 3.5 percent discount rate, comparators used in NHS, UK setting, UK NHS and personal social services costs, cost per QALY below the NICE threshold value, assumed to be £30 k per QALY for evaluations in cancer (i.e., society is willing to pay £30k [EUR 37.4 k] for one additional QALY) requirements for economic evaluations in healthcare (4). Many of the criteria that were applied are relevant to economic models irrespective of setting and the preferred items overlap with those in published quality checklists (8;13). The evidence was reviewed by a single researcher and any issues were discussed within the review team until a consensus was achieved in cases of uncertainty or lack of clarity.

RESULTS

The searches identified 1,133, 2,408, and 722 unique references for the breast, colorectal and prostate cancer reviews, respectively (Figure 1). Seventy-eight full papers were retrieved for detailed inspection. Of these, sixteen studies met the inclusion criteria (Table 1). One additional study published after completion of the searches was identified and has been included in this article.

Three studies were in breast cancer (publication dates 1997 to 2003) (14–16), seven were in colorectal cancer (publication dates 2000 to 2011) (17–23) and seven were in prostate cancer (publication dates 2000 to 2012) (24–30). Two of the three breast cancer studies were set in the United States (14;16), and one was in Norway (15). Two of the seven colorectal studies were set in the United Kingdom (18;23), two in the United States (21;22), one in Canada (19), one in New Zealand (20), and one in Spain (17). Three of the seven prostate studies were

Economic evidence of surgical procedures in cancer

Table 1. Characteristics of the Studies Included

Author, year	Setting	Population	Stage	Intervention	Comparator	Time horizon	Clinical evidence
Breast cancer							
Norum, 1997(15)	Norway	Women diagnosed with early stage breast cancer	Stage I and II	BCS – lumpectomy	MRM	10 years	Several clinical studies
Malin, 2002(14)	US	8 hypothetical women classified based on their estrogen-receptor type (positive or negative), lymph node type (positive or negative), and age (45 or 60 years of age)	Stage I,II and III	a) Adjuvant therapy strategies after surgery, b) radiation after BCS, c) reconstruction after surgery	Mastectomy alone	Lifetime	Several RCTs
Polsky, 2003(16)	US	Women 65 years of age or older, who were diagnosed early stage breast cancer between 1992 and 1994	Stage I and II	BCSRT	Mastectomy	5 years	National survey cohort
Colorectal cancer							
Miller, 2000(22)	US	Locally recurrent rectal cancer patients (n = 68), 62% were males, median age was 58, median time from diagnosis to surgery was 24 months and median time from diagnosis to reoccurrence was 14 months	Locally recurrent rectal cancer	Surgical resection	Diagnostic or palliative surgery	3.5 years	Patients' records for survival and resource use
Govindarajan, 2006(19)	Canada	70-year-old male with a complete emergent malignant left colonic obstruction secondary to a left-sided colon cancer with no evidence of metastatic disease.	Stages II, III and IV	Emergency colonic stenting as a bridge to definitive surgery	Emergency surgery	Lifetime	Meta-analysis
Murray, 2006(23)	UK	A cohort of 65-year-old colorectal cancer patients	Stages I, II and III	LS	OS	25 years	Mixture of RCTs, a meta-analysis, and assumptions
deVerteuil, 2007(18)	UK	Colorectal cancer patients over 50 years of age	Different stages have been considered as subgroup analysis (results are not reported)	LS	OS	25 years	Meta-analysis
Hayes, 2007(20)	New Zealand	Colorectal cancer patients from meta-analysis of several trials	No reference to stage	Laparoscopic-assisted colectomy (LAC)	Conventional open colectomy (OC)	Not stated	Two clinical studies
Callejo, 2011(17)	Spain	Colorectal cancer	No reference to stage	LS	OS	No information	Meta-analysis
Jensen, 2011 (21)	US	Colon or rectal cancer	No reference to stage	LS	OS	No information	No information

Table 1. Continued.

Author, year	Setting	Population	Stage	Intervention	Comparator	Time horizon	Clinical evidence
Prostate cancer							
Jager, 2000(26)	US	A cohort of 65 year-old men with prostate cancer eligible for RP	No reference to stage	RP based on MR imaging (extra capsular disease contraindicates RP)	RP based on clinical staging	Lifetime	Clinical studies
Calvert, 2003(24)	UK	60-year-old men diagnosed with moderately differentiated (Gleason sum score 5 − 7) prostate cancer	Early stage localised prostate cancers	DNA-ploid test followed by RP for those with nondiploid test	a) RP for all, b) WW and monitoring	Lifetime	Limited detail
O'Malley, 2007(29)	Australia	Prostate cancer patients eligible for surgery	Late stages as the patients are undergoing surgery	LRARP	Open RP	1 year	Cohort
NICE, 2008(ref)	UK	Men with clinically localised PC	Stages I to IV	RP	WW	20 years	Clinical trial
Hohwu, 2011(25)	Denmark	Retrospective cohort study a total of 231 men (50 to 69 years of age) and with clinically localised prostate cancer underwent RP	PC stages cT1-T2	RALP	RRP	1 year	Retrospective cohort
Lyth, 2012(27)	Sweden	695 patients enrolled in SPCG-4 trial between October 1989 and February 1999. Eligible patients were aged less than 75 years, life expectancy was greater than 10 years and no other malignancy was present. The prostate-specific antigenvalue had to be below 50 ng/ml.	Patients with newly diagnosed tumours that are localized to the prostate	RP	WW	Lieftime	RCT
Ramsey 2012(30)	UK	Men with localised prostate cancer undergoing radical prostatectomy at designated pelvic cancer surgical treatment centres within the UK NHS	Stage I to IV	Laparoscopic prostatectomy and robotic prostatectomy	OS	10 years	Systematic review

BCS, breast conserving surgery; BSCRT, breast conserving surgery with radiation; CEA, cost-effectiveness analysis; LRARP, laparoscopic remotely assisted radical prostatectomy; LS, laparoscopic surgery; LY, life-year; MRM, modified radical mastectomy; MR, magnetic resonance; OS, open surgery; QALY, quality adjusted life year; RP, radical prostatectomy; RALP, Robot-assisted laparoscopic prostatectomy; RRP, retropubic radical prostatectomy; WW, watchful waiting.
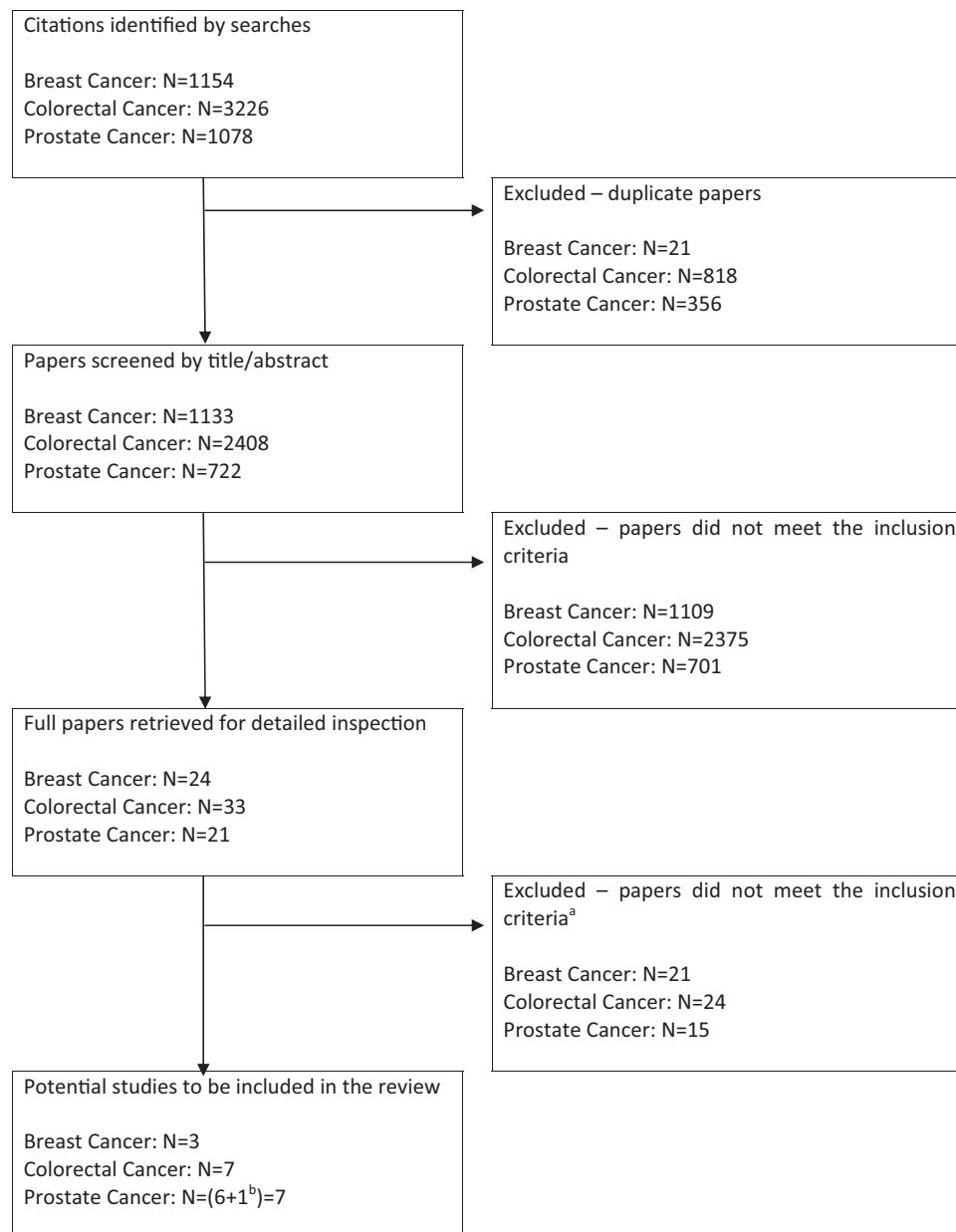
Table 2. Methods and Study Conclusions

Author, year	Synopsis of methods	Study conclusion
Breast cancer		
Norum, 1997(15)	CEA with a focus on the cost of surgery, radiotherapy, adjuvant therapy, follow up, and the cost of relapse	BCS is cost-effective and should not be displaced by mastectomy. BCS following MRM may be considered but there was little evidence to support this treatment option.
Malin, 2002(14)	CEA that combines short-term gains from surviving the first 10 years with gains from reduced rates of breast cancer recurrence over the rest of the woman's life	Adjuvant therapy after surgery, BCS with radiation and reconstruction after mastectomy tend to be more cost-effective than surgery alone.
Polsky, 2003(16)	Retrospective cohort study with two CEAs — a traditional analysis that compared mastectomy as the standard of care with BCSRT as the new intervention and an alternative CEA where the standard of care is assumed to be an open regimen in which both BCSRT and mastectomy are widely available	Providing a choice between mastectomy and breast conservation surgery is economically attractive when the economic analysis includes the benefit of patient choice of treatment.
Colorectal cancer		
Miller, 2000(22)	Economic evaluation using institution records	Patients with recurrent rectal carcinoma assess the discomfort due to surgery and morbidity to be less severe than health care providers. Surgical resection may be a cost-effective intervention, particularly when cost-effectiveness is calculated using patient preferences.
Govindarajan, 2006(19)	Markov Model	Colonic stenting as a bridge to surgery is more effective and less costly than surgery in the treatment of emergent, malignant left colonic obstruction.
Murray, 2006(23)	Markov model that allows a cohort of patients to progress from their initial operation to convalescence to "disease free" state or death	Laparoscopic resection is associated with a quicker recovery period with no evidence of a change in the mortality rate for a modest additional cost (between £250 and £300 per patient), although operation times were longer. Assuming equivalence of long-term outcomes, a judgement is required as to whether the benefits associated with earlier recovery are worth this extra cost.
deVerteuil, 2007(18)	Markov model for a cohort of patients from their initial operation to convalescence to "disease free" state or death	LS is likely to be associated with short-term quality of life benefits, similar long-term outcomes, and an additional £300 per patient. A judgment is required as to whether the short-term benefits are worth this extra cost.
Hayes, 2007(20)	Costing exercise with hospital cost data & RCT recovery data	LAC for cancer appears to be cost-effective relative to OC for the lower of the average cost estimates but there is a great deal of uncertainty around the costs. Potential future reductions in operating times, conversion rates and postoperative stays are likely to further improve cost-effectiveness.
Callejo, 2011(17)	Markov model	Although the cost for LS was higher than for OS, LS appeared to accelerate recovery time
Jensen, 2011 (21)	Model based	Laparoscopic resection for colon and rectal cancer results in both decreased costs and slight improvements in quality of life, making it the preferred approach in suitable patients

Table 2. Continued.

Author, year	Synopsis of methods	Study conclusion
Prostate cancer		
Jager, 2000(26)	Decision tree model with two assumptions: i) MR staging was performed in addition to other staging methods in patients considered candidates for radical prostatectomy and ii) extra-capsular disease on MR images contraindicates surgery.	It is not yet conclusively determined whether preoperative MR staging is appropriate, but results of decision analysis suggest that MR staging is cost-effective for men with moderate or high prior probability of extracapsular disease.
Calvert, 2003(24)	Markov model	A treatment selection policy based on DNA-ploidyanalysis would be cost-effective if the test can achieve specificity levels at or above 80%.
O'Malley, 2007(29)	Economic evaluation with data from a cohort study	The incremental cost-effectiveness for LRARP compared with open surgery is well below the accepted Australian pharmaceutical threshold.
NICE, 2008(ref)	Markov Model which took into account the progression of the patients underlying prostate cancer and the side effects due to individual treatments.	The cost-effectiveness of radical prostatectomy is dependent on the health outcomes included. If only patient survival is considered, then radical prostatectomy is arguably cost-effective. When quality-of-life considerations for prostate cancer and treatment-related side effects are included, watchful waiting becomes the dominant option.
Hohwu, 2011(25)	Economic evaluation alongside a retrospective cohort study	RALP was more effective and more costly. A way to improve the cost effectiveness may be to perform RALP at fewer high volume urology centres and utilise the full potential of each robot.
Lyth, 2012(27)	A semi-Markov model that use symptomatic disease instead of hormonally controlled metastasis and use refractory disease instead of refractory metastasis, with an assumption that the patients cannot die of PC unless they have developed a refractory disease.	The cost-effectiveness ratio for RP varies with age, Gleason score and prostate-specific antigen values. RP was not perceived to be cost-effective in men aged 75 years with low Gleason and low PSA. Higher threshold values for patients with clinically localized prostate cancer could be discussed.
Ramsey 2012(30)	An individual-based event model constructed by discrete-event simulation approach	Robotic prostatectomy had lower perioperative morbidity and a reduced risk of a positive surgical margin compared with laparoscopic prostatectomy, although there was considerable uncertainty.

BCS, breast conserving surgery; BSCRT, breast conserving surgery with radiation; CEA, cost-effectiveness analysis; LRARP, laparoscopic remotely assisted radical prostatectomy; LS, laparoscopic surgery; LY, life year; MRM, modified radical mastectomy; MR, magnetic resonance; OS, open surgery; QALY, quality-adjusted life-year; RP, radical prostatectomy; RALP, Robot-assisted laparoscopic prostatectomy; RRP, retropubic radical prostatectomy; WW, watchful waiting.

Figure 1. Flow diagram for study inclusion.

set in the United Kingdom (24;28;30), with the remaining four in Australia (29), Denmark (25), Sweden (27), or the United States (26). The studies scan a broad publication range with six published before 2004 (older publications) and five published after 2010 (recent publications). Of particular interest, there were no publications for breast cancer after 2003.

Essential Basic Requirements

Defined Research Question. A basic requirement of any economic evaluation in health care is a clear and well defined decision problem which states both the exact clinical definition of the indication of interest, and detailed information on the interventions being compared (4). This enables reviewers and readers to assess if the analyses and results are relevant to their own settings and practice. The patient populations were reasonably well defined in the majority of studies (Figure 2), but there were exceptions. In the colorectal and prostate studies, the exact patient population or stage of cancer was either not clear or was not provided in some articles (17;20;21;26).

The interventions and comparators were generally clearly defined in all the studies. The three breast cancer studies compared breast conserving surgery with mastectomy with one also comparing reconstruction after surgery (14). Six of the

	Breast cancer			Colorectal cancer							Prostate cancer						
	Norum 1997(15)	Malin 2002(14)	Polsky 2003(16)	Miller 2000(22)	Govindarajan 2006 (19)	Murray 2006(23)	De Verteuil 2007(18)	Hayes 2007(20)	Callejo 2011(17)	Jensen 2011(21)	Jager 2000(26)	Calvert 2003(24)	O'Malley 2007(29)	NICE 2008	Hohwu 2011(25)	Lyth 2012(27)	Ramsay 2012(30)
Essential requirements																	
Patient group/indication clearly defined	+	+	+	+	+	+	+/-	+/-	+/-	+	+	+	+/-	+	+	+	+
Comparators clearly defined	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Effectiveness evidence based on a systematic review	-	-	-	-	+	+	+	+	+	-	-	-	-	-	-	-	+
Appropriate time horizon	-	+	+	-	+	+	+	-	-	-	+	+	+	+	+	+	-
HRQoL data reported directly by patients &/or carers	-	-	+	+	+	-	+	+	-	-	-	-	-	?	+	+	+
Probabilistic sensitivity analysis used to quantify uncertainty	-	-	-	-	+	+	-	-	+	-	-	-	-	-	-	+	+
Relevant one-way sensitivity analyses	-	+	+	-	+	+	+	+	-	+	+	+	-	+	+	+	+
Preferred requirements																	
Cost-effectiveness analysis using QALYs	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
HRQoL preference data valuation by representative sample of the public	-	-	-	-	-	-	+	+	-	-	?	+	-	-	+	-	?
UK/NICE specific requirements																	
Discount rate, 3.5%	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+
Comparators used in the NHS	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+
UK Setting	-	-	-	+	-	+	+	-	-	-	-	+	-	+	-	-	+
UK NHS and personal social services costs	+	+	+	-	-	+	+	-	-	-	-	+	-	+	-	-	+
Cost/QALY < NICE threshold value	+**	+**	-	-	+**	-	-	-	+[a]	+[b]	+**	+**	+**	+/-	-	+/-	+**

** Cost/QALY < GBP 20k (EUR 25k) NICE threshold value, * Cost/QALY < GBP 30k (EUR 37.4k) NICE threshold value. [a]Cost-effective under EUR 36.7k. [b]probability of cost-effectiveness 60% considering a willingness-to-pay between EUR 25k/QALY and EUR 37.4k/QALY

E=essential basic requirements of any economic evaluation, P=preferred requirement, UK=requirement specific to UK NICE reference case

\+ Satisfies requirements   - Does not satisfy requirements   +/- Partially satisfies requirements   ? Unclear or information not provided

Figure 2. Quality of included studies compared to the NICE reference case.

colorectal studies compared laparoscopic surgery with open resection (17;18,20;21;23), one compared surgical resection with diagnostic or palliative surgery (22), and one compared emergency colonic stenting as a bridge to definitive surgery with emergency surgery (19). Two of the prostate cancer studies compared radical prostatectomy with watchful waiting (27;28), one compared radical prostatectomy for those with non-diploid test with either radical prostatectomy for all or, watchful waiting (24), one compared radical prostatectomy based on magnetic resonance (MR) imaging with radical prostatectomy based on clinical staging (26), one compared laparoscopic remotely assisted radical prostatectomy with open radical prostatectomy (29), one compared laparoscopic prostatectomy and robotic prostatectomy with open surgery (30), and the final study compared robot-assisted laparoscopic prostatectomy with retropubic radical prostatectomy (25).

Clinical Evidence. Another essential requirement relates to the clinical data used to quantify potential benefits of the interventions under evaluation. Evidence of a full literature search and, where appropriate, a synthesis of evidence on health effects informed by the systematic review, are mandatory to ensure the data used represent all the evidence available at the time of the study and the associated uncertainty surrounding the point estimates (2–4;6). The effectiveness rates used in the majority of the studies included in this study were not informed by systematic literature reviews or a synthesis of all available data. In breast cancer, none of the studies used effectiveness evidence informed by a systematic literature review or a synthesis of data. In colorectal cancer, five of the seven studies used the results of meta-analyses (17–20;23). One study used event rates observed in a single hospital as per their research question (22), and one provided insufficient detail to determine the source of the effectiveness evidence used (21). In prostate cancer, one study used the results of a meta-analysis (30), two used data from single randomized control trials (RCTs) with no evidence to suggest alternative evidence was used in sensitivity analyses (27;28), two used data from retrospective cohort studies (25;29), one study was informed by a mixture of evidence from RCTs and observational data obtained from a systematic search supplemented by expert opinions (26), and one provided insufficient detail to determine the source (24).

Time Horizon. Many of the interventions will have a differential effect on recurrence rates and/or mortality risk hence the most appropriate time horizon for the costs and benefits would be a lifetime (4). However, for individuals with cancer, a lifetime horizon is potentially considerably shorter than that of the general population due to the reduced life expectancy associated with the condition. Over half (10/17) of the studies took account of longer-term costs and benefits associated with the interventions using a minimum of 10 year horizons (14;18–19;23;24;26–28). The remaining studies arguably underestimated both the costs and benefits of the interventions. One study used the aver-

age weighted survival duration (42 months for resected patients) (22). One restricted costs and QALYs to the first post-operative year (25), while one used 5-year retrospective data without extrapolating either costs or benefits (16). Two studies, which limited the perspectives to that of single hospitals, presented incremental benefits over short horizons without extrapolating to account for downstream benefits or costs. One restricted the time horizon to length of stay, estimating the incremental difference in recovery time (20), while the second study estimated the reduction in months (range, 3.79 to 8.67 months) of incontinence or erectile dysfunction (29).

HRQoL. HRQoL data should be collected directly from patients with the conditions described in the economic model and preferably from patients in receipt of the interventions under evaluation (4;7). When it is not possible for patients with the conditions to provide these values, then proxy values may be obtained from their principle carer. However, values obtained from healthcare professionals are not appropriate and have been shown to differ substantially from those obtained from patients with the condition being valued (31).

The description of sources of HRQoL data was poor in many of the studies. While over a third of the studies (7/17) incorporated HRQoL data which were collected directly from individuals with the condition of interest, three studies used HRQoL data obtained from clinicians (14;15,22;23), or nurses (15;23), and one study assumed HRQoL values based on the New Zealand social tariff (20). There was insufficient evidence to determine the source of the HRQoL data in six of the studies (15;17;21;24;26;29).

Uncertainty. All economic evaluations should accurately characterize the decision uncertainty associated with the interventions under evaluation. At the very least, univariate sensitivity analyses should be conducted to illustrate both structural uncertainty (for example assumptions relating to the description of the clinical pathway, or the extrapolation of data) and uncertainty around key parameter values (for example if different sources for HRQoL or cost data are available) (4). Ideally, a full probabilistic sensitivity analysis using Monte Carlo simulations would be performed to demonstrate the uncertainty around the mean point estimates used (8), and is encouraged by several decision-making bodies (3;4;6).

The majority of studies (13/17) performed univariate sensitivity analyses to illustrate the effect on results associated with variations in key parameter inputs. Many of the authors reported cost-effectiveness results were sensitive to changes in parameter values and particularly for recurrence rates, hernia rates, HRQoL values, and side-effects of treatments. The full range of uncertainty was poorly captured with less than one third (5/17) of the studies presenting the results of a full probabilistic sensitivity analysis (17;19;23;27;30). It is perhaps surprising that some of the newer publications did not report these results as this is now considered to be standard practice in health economics.

Preferred Requirements

Although the preferred outcome from economic evaluations varies depending on the reimbursement body, the cost per QALY, where the valuation of the HRQoL is based on public preferences, is now a pivotal requirement for submissions to NICE (4). The QALY describes both survival and HRQoL weights in a single metric and is the recommended gold standard to facilitate comparison across disparate conditions and equity in policy decisions (8). Although all studies reported results in terms of a cost per QALY (or cost per quality adjusted month) (19), few (4/17) used general population weights for all utilities (18;20;24;25). One study mapped from HRQoL values onto a scale assuming 0 and 100 were equal to death and full health, respectively (15), while two studies used patient values from the patients' completed Euroqol VAS, again assuming 0 and 100 were equal to death and full health, respectively (16;27). Three studies elicited weights from clinicians using standard gamble (14;22;23), while a fourth used time-trade off to derive values from patients. Six studies provided insufficient detail to determine if general population weights were used (14;15;17;21;26;29).

UK Specific Requirements

Discount rates are both jurisdiction and time dependent and are used to ensure results reflect the present value of the cost and benefits accrued over the duration of the analyses (4). Not surprisingly, due to the differences in settings and the publication dates, only 3/17 of the studies used the current UK recommended discount rates (3.5 percent per annum) for costs and benefits (27;28;30). However, all studies that used horizons over one year did discount costs and effects.

While some reimbursement bodies restrict the perspective to the direct healthcare resource costs and savings directly attributed to a health service, others take a broader societal view and require indirect costs such as productivity losses or caregiver's time (5). Two-thirds (13/17) of the studies limited costs to direct healthcare costs (14–16;18–20;22–24;26–30), and two of these restricted costs to those directly incurred by hospitals (20–22). The remainder took a broader perspective and costs such as productivity losses (21;25), and caregiver's time (21), were included.

Applying the NICE threshold value of GBP 20–30 k (EUR 25–37.4 k) per QALY, irrespective of the time of publication, the setting or adherence to the NICE methods guide, breast conserving surgery plus axillary mode dissection compared with modified radical mastectomy would be considered cost-effective while breast conservation surgery with radiation therapy compared with mastectomy would not be considered cost-effective (14). There are multiple reasons for this including cosmesis, patient preferences, and the added cost of reconstruction which is used after mastectomy in over 20 percent of all cases and recurrence rates requiring

further interventions (32;33). It should be noted that the latter comparator is no longer appropriate in the current clinical context as it is now current practice, with rare exceptions, for radiation to be administered after conservation therapy.

Applying the GBP 20–30 k (EUR 25–37.4 k) per QALY threshold to the studies on colorectal cancer, the probability that laparoscopic surgery would be considered cost-effective compared with open surgery ranged between 30 percent (18) and 40 percent (23) and univariate analyses produced results where laparoscopic surgery either dominated (higher costs and lower benefits) open surgery or was dominated (lower costs and larger benefits) by open surgery suggesting considerable uncertainty (18;20;23). The evaluation of emergency colonic stenting as a bridge to definitive surgery dominated emergency surgery in individuals with emergent, malignant left colonic obstruction (19), and non-surgical treatment dominated both surgical resection and diagnostic/palliative surgery in individuals with recurrent colorectal cancer (22).

Applying the GBP 20–30 k (EUR 25–37.4 k) per QALY threshold to the studies for prostate cancer, radical prostatectomy treatment is unlikely to be considered cost-effective compared with watchful waiting due to considerable uncertainty around the incremental cost-effectiveness ratio (ICER). Again, the cost per QALY results ranged from below GBP 20 k (EUR 25 k) per QALY (except for men aged 75 and over) when the decision was based on patient survival only (27), to dominated when side effects of the surgical technique were taken into account (28). While a marker-based (DNA-ploidy) treatment selection policy would be considered cost-effective compared with watchful waiting using these thresholds, both policies dominated the "prostatectomy for all" option (24). However, the ICER was highly sensitive to factors such as sensitivity and specificity (24), and in one study the ICER ranged from approximately GBP 2 k (EUR 2.5 k) per QALY to GBP 84 k (EUR 105 k) per QALY (27). Robotic prostatectomy was likely to be cost-effective at GBP 30 k (EUR 37.4 k) per QALY in comparison to laparoscopic prostatectomy when the number of procedures performed per year with the robotic system was over 150 (30), whereas radical prostatectomy was cost-effective at GBP 20 k (EUR 25 k) per QALY threshold compared with open surgery (29).

DISCUSSION

The limited numbers of studies included in the reviews highlight the scarcity of economic evidence for surgical procedures in these indications. The searches identified just three studies for breast cancer surgery, all were published before 2003 (14–16). Of the seven colorectal cancer studies, only two were published after 2010;(17;21) and of the seven prostate cancer studies, three were published after 2010 (25;27;30).

Although some UK specific requirements were included in the methodological checklist (Figure 1), the vast majority of

criteria used to assess the scientific rigor of the studies are fundamental requirements for any economic evaluation in health care and are, therefore, generalizable to other settings and decision-making bodies. Scientific rigor is improving in economic evaluations generally, and older publications may not conform to current practice. The studies included in the reviews cover a broad publication range with six (6/17) published before 2004 (older publications) and five (5/17) published after 2010 (recent publications). Assessing the studies chronologically (Figure 2), there is a marked difference in the proportion that satisfies the essential criteria. For example, although only 5/17 of all the studies reported results of full probabilistic sensitivity analyses, none (0/6) were older publications and three (3/5) were recent publications. Similarly, although only 6/17 used a systematic review and meta-analysis to inform the clinical evidence used, none (0/6) were older publications and three (3/5) were recent publications. Similar rates were observed for the use of HRQoL data reported directly by the patients or carers with none (0/6) of the older publications using these data compared with two (2/5) of the recent publications. While the majority of the older evaluations were not undertaken using the current requirements for sources of evidence (clinical data informed by a systematic literature review/synthesis, HRQoL data from patients with the condition) and uncertainty (probabilistic sensitivity analysis) the recent publications would be expected to conform to these basic standards.

There were several essential requirements which were not satisfied in many of the publications. First, much of the clinical evidence was not supported by a systematic review of the literature or a synthesis of the relevant evidence, and very few authors presented the results of a full probabilistic sensitivity analysis. Relying on the effectiveness evidence from a single clinical study or hospital setting in isolation, with no reference to the relevant published evidence, can be misleading as the effectiveness evidence from a single study may differ substantially from other evidence in the area. This may hinder and even mislead policy decision makers in some cases. While it is not always possible to conduct a full systematic literature review and evidence synthesis within the resource allocations of a project, it is generally possible to compare the effectiveness evidence used with the literature. If the data used are comparable to the published literature, this will increase confidence in the results presented. If the data differ substantially from the published literature, then the results obtained using the different values should be compared (preferably using a probabilistic sensitivity analysis) to determine if the evaluation is sensitive to variations in these parameters. Exploring the uncertainty around the point estimates is particularly relevant if the evidence comes from a novel intervention where there is no existing evidence.

The second and closely related area where many of the studies failed to meet requirements was in the choice of time horizon used to sum the benefits and costs accrued from the alternative interventions. It is widely accepted that if an intervention has a differential effect on survival, the most appropriate horizon would be a lifetime. As the patients in the evaluations had cancer, then a shorter horizon (say 10 to 20 years) may be considered sufficient. However, many of the studies restricted the horizon to less than 5 years, or did not provide details of the length of horizon used. It is unlikely that a time horizon of 1 year will accurately reflect the full benefits of an intervention in patients with cancer. While some analysts may be reluctant to rely on assumptions to extrapolate beyond the duration of the clinical evidence, the effect of longer horizons should be explored in sensitivity analyses to determine what if any difference this makes to the results generated.

The final area where many of the studies failed to satisfy the essential requirements relates to the HRQoL data used within the analyses. Over half of the studies did not use data collected from patients (or their carers) with the particular conditions. Reporting standards for both the source of evidence and the methodology used to collect these data were particularly low. Many policy makers require that the preference-weights used to determine QALYs are obtained from the general population. There is an argument that in some health conditions, generic measures such as the EQ-5D may not be sensitive to small changes in dimensions of health or features of a particular condition. In these cases, a condition specific measure of health, such as the Functional Assessment of Cancer Therapy (FACT) for patients with cancer, may be more relevant (34). At the moment the FACT measure does not have a corresponding preference-weighted index which could be used to determine the QALYs in an economic evaluation. While several of the studies (4/17) conducted analyses to elicit preference weights for defined health states in the evaluations, in many of the studies it was impossible to determine what preference-weights, if any, had been used.

To our knowledge, little or no research has been conducted to date which has compared the equivalence of reporting standards for economic evaluation for surgical procedures and those for pharmaceutical products. Comparing methods and standards in economic evaluation of medical devices and drugs, authors highlighted key challenges for evaluating medical devices. These may generalize to surgical interventions including the practical difficulties in conducting randomized controlled trials (35–36), the dependence of outcome on the surgeon's expertise (35). and the wider organizational impacts associated with introducing new devices which may differ by location (35–36). In some countries (for example Australia and Canada), HTAs and cost-effectiveness analyses focus on pharmaceuticals rather than the full range of health-care technologies, which may go some way to explaining why there is a dearth of evidence on surgical interventions.

Given the scarcity of literature and relatively dated clinical evidence for some of the indications and procedures, it is reasonable to question what economic evidence is used to

support clinical guidelines in these areas. In the United Kingdom, the number of HTA and NICE guidance issued can be used as an indicator of the priority given to cancer health care (37;38). Since its inception, NICE has issued guidance relating to twenty-eight (11;13) pharmaceutical interventions and seven (13) intervention procedures for breast (colorectal, prostate) cancer, and there have been forty-five (25;32) HTA submissions.

The most recent clinical guideline for early (and locally advanced) breast cancer has very limited de-novo evidence in terms of alternative surgical techniques (39;40), The cost-effectiveness literature was not reviewed under the following topics: "Surgery to the breast" (including Paget's disease), "evaluation and management of a positive lymph node" and "breast reconstruction" as "the Guideline Development Group (GDG) did not consider this topic as a health economic priority". The only independent evaluation conducted to inform the GDG's recommendations examined alternative strategies used for staging of the cancer before surgery.

For colorectal cancer, the recommendations for laparoscopic surgery in the current UK clinical guidelines (41) were taken directly from NICE technology appraisal guidance on "Laparoscopic surgery for colorectal cancer TA105" (42), which used cost-effectiveness evidence reported in an article included in the current review (23). Although this evidence suggested laparoscopic surgery was dominated by open surgery, the laparoscopic surgery was recommended over open surgery as the NICE committee informed by TA105 were "persuaded there were important differences" in length of stay and return to normal activities for laparoscopic surgery, and even though costs were greater and there was little direct evidence of quality of life benefits, "it was likely that such benefits exist . . . . and these would be sufficient to make laparoscopic procedures cost-effective". For prostate cancer, the existing cost-effectiveness evidence was considered to be limited as none included evidence from a recently published RCT, hence, a de novo model (included in this review) was constructed to incorporate these data (see 28).

Authors of a recent report expressed concern with regard to the future of academic surgery in the United Kingdom and highlighted the need for additional research activities in cancer surgery and in particular multidisciplinary research (40). Surgery is often associated with the greatest potential benefit (cure/not cure) and potential harm (risk of morbidity/mortality) when comparing modalities in the cancer treatment pathway (40). It is, therefore, likely that new surgical devices and changes in techniques and timing will be cost-effective and even potentially cost-saving compared with pharmaceutical interventions. However, such a proposition needs to be evidence-based and the scarcity of economic evidence on surgical techniques suggests that this is one area where the surgical community and policy-decision makers could benefit from a wider multi-disciplinary approach to research.

CONCLUSION

There is a dearth of recent robust evidence describing the cost-effectiveness of surgical interventions in these indications. Many of the recent publications did not satisfy essential requirements such as exploring the full uncertainty associated with the evidence or using clinical evidence informed by a systematic review and synthesis. Policy decisions for surgical techniques in these indications do not appear to be informed by robust economic evidence in all cases and due to the ratio of potential benefit and harm associated with cancer surgery there is an urgent need to increase economic activities in this area.

CONTACT INFORMATION

Roberta Ara, MSc(hons), ScHARR, Senior Research Fellow, School of Health and Related Research (ScHARR), Regent Court, 30 Regent Street, Sheffield, S1 4DA

Hasan Basarir, PhD, ScHARR, Research Fellow, The University of Sheffield, School of Health and Related Research (ScHARR), Regent Court, 30 Regent Street, Sheffield, S1 4DA

Anju D. Keetharuth, Phd, ScHARR, Research Associate, The University of Sheffield, School of Health and Related Research (ScHARR), Regent Court, 30 Regent Street, Sheffield, S1 4DA

Marco Barbieri, MSc(hons), Centre for Health Economics (CHE), University of York, Heslington, York YO10 5DD, UK

Helen L.A. Weatherly, MSc(hons), Senior Research Fellow, Centre for Health Economics (CHE), University of York, Heslington, York YO10 5DD, UK

Mark J.S. Sculpher, Phd, Professor of Health Economics, Deputy Director of the Policy Research Unit in Economic Evaluation of Health and Care Interventions (EEPRU), Centre for Health Economics (CHE), University of York, Heslington, York YO10 5DD, UK

Hashim Ahmed, FRCS(Urol), PhD, BM, BCh (Oxon), BA(Hons), MRC Clinician Scientist and Honorary Consultant Urological Surgeon, University College Hospital, 235 Euston Road, London, NW1 2BU

Steven Brown, MBChB FRCS MD, Consultant Colorectal Surgeon, Sheffield Teaching Hospitals, NHS Foundation Trust, Northern General Hospital, Herries Road, Sheffield, S5 7AU

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

References

1. Sorenson C, Drummond MF, Kanavos P. *Ensuring value for money in health care: The role of health technology assessment in the European Union*. Observatory Studies Report N.11. Copenhagen: European Observatory on Health Systems and Policies; 2008.
2. Pharmaceutical Benefits Advisory Committee. *Guidelines for Preparing Submissions to the Pharmaceutical Benefits Advisory Committee* (version 4.3). Australian Government: Department of Health and Ageing; 2008.

3. Dutch Healthcare Insurance Board. *Guidelines for pharmacoeconomic research* (updated version. Valid as from April 1st 2006). Dieman: Dutch Health Care Insurance Board; 2006.

4. NICE. *Guide to methods of technology appraisal*. London UK; 2013. [updated January 1, 2013]. http://publications.nice.org.uk/pmg9 (accessed March 31, 2014).

5. Johannesson M, Jonsson B, Jonsson L, et al. *Why should economic evaluation medical innovations have a societal perspective*. No 51. London: Office for Health Economic Briefings; 2009.

6. Canadian Agency for Drugs and Technologies in Health. *Guidelines for the economic evaluation of health technologies: Canada*, 3rd ed. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006.

7. Lopez-Bastida J, Oliva J, Antonanzas F, et al. Spanish recommendations on economic evaluation of health technologies. *Eur J Health Econ*. 2010;11:513-520.

8. Gold MR, Siegel JE, Russell LB, et al. *Cost-effectiveness in health and medicine*. New York: Oxford University; 1996.

9. National Audit Office. 2010. [Internet] [updated January 1, 2013]. http://www.nao.org.uk (accessed March 6, 2013).

10. Harker R. House of Commons (2012). *NHS funding and expenditure*. Standard Note SN/SG/724. www.nhshistory.net/parlymoney.pdf (accessed February 20, 2013).

11. Leal J. 2012. *Pharmatimes online*. [Internet] [updated January 1, 2013]. http://www.pharmatimes.com/article/12-11-08/Cancer_costing_UK_%C2%A315_8_billion_a_year.aspx (accessed March 6, 2013).

12. Ara R, Barbieri M, Weatherly H, et al. *A rapid review of the cost-effectiveness of interventions across the cancer patient pathways: Radiotherapy and surgery*. Report for the Department of Health. www.eepru.org.uk (accessed March 6, 2014).

13. Drummond MF, Sculpher MJ, Torrance GW, et al. Methods for the economic evaluation of health care programs. 3rd ed. Oxford: Oxford University Press; 2005.

14. Malin JL, Keeler E, Wang C, et al. Using cost-effectiveness analysis to define a breast cancer benefits package for the uninsured. *Breast Cancer Res Treat*. 2002;74:143-153.

15. Norum J, Olsen JA, Wist EA. Lumpectomy or mastectomy? Is breast conserving surgery too expensive? *Breast Cancer Res Treat*. 1997;45:7-14.

16. Polsky D, Mandelblatt JS, Weeks JC, et al. Economic evaluation of breast cancer treatment: Considering the value of patient choice. *J Clin Oncol*. 2003 ;21:1139-1146.

17. Callejo D, Guerra M, Reza M, et al. Cost-utility analysis of laparoscopic versus open surgery for colorectal cancer. *Value Health*. 2011;14:A263.

18. de Verteuil RM, Hernandez RA, Vale L, Aberdeen Health Technology Assessment Group. Economic evaluation of laparoscopic surgery for colorectal cancer. *Int J Technol Assess Health Care*. 2007;23:464-472.

19. Govindarajan A, Naimark D, Coburn NG, et al. Use of colonic stents in emergent malignant left colonic obstruction: A Markov chain Monte Carlo decision analysis. *Dis Colon Rectum*. 2007;50:1811-1824.

20. Hayes JL, Hansen P. Is laparoscopic colectomy for cancer cost-effective relative to open colectomy? *Anz J Surg*. 2007;77:782-786.

21. Jensen C, Abcarian H. Cost-effectiveness of laparoscopic versus open resection for colon and rectal cancer. *Dis Colon Rectum*. 2011;54:E186-E187.

22. Miller AR, Cantor SB, Peoples GE, et al. Quality of life and cost effectiveness analysis of therapy for locally recurrent rectal cancer. *Dis Colon Rectum*. 2000;43:1695-1701.

23. Murray A, Lourenco T, de Verteuil R, et al. Clinical effectiveness and cost-effectiveness of laparoscopic surgery for colorectal cancer: Systematic reviews and economic evaluation. *Health Technol Assess*. 2006;10:1-141.

24. Calvert NW, Morgan AB, Catto JW, et al. Effectiveness and cost-effectiveness of prognostic markers in prostate cancer. *Br J Cancer*. 2003;88:31-35.

25. Hohwu L, Borre M, Ehlers L, et al. A short-term cost-effectiveness study comparing robot-assisted laparoscopic and open retropubic radical prostatectomy. *J Med Econ*. 2011;14:403-409.

26. Jager GJ, Severens JL, Thornbury JR, et al. Prostate cancer staging: Should MR imaging be used?–A decision analytic approach. *Radiology*. 2000;215:445-451.

27. Lyth J, Andersson SO, Andren O, et al. A decision support model for cost-effectiveness of radical prostatectomy in localized prostate cancer. *Scan J Urol Nephrol*. 2012;46:19-25.

28. The Guideline Development Group. *Prostate cancer: Diagnosis and treatment*. Clinical Guideline CG58 2008. http://www.nice.org.uk/CG58 (accessed March 6, 2014).

29. O'Malley SP, Jordan E, O'Malley SP, et al. Review of a decision by the Medical Services Advisory Committee based on health technology assessment of an emerging technology: The case for remotely assisted radical prostatectomy. *Int J Technol Assess Health Care*. 2007;23: 286-91.

30. Ramsay C, Pickard R, Robertson C, et al. Systematic review and economic modelling of the relative clinical benefit and cost-effectiveness of laparoscopic surgery and robotic surgery for removal of the prostate in men with localised prostate cancer. *Health Technol Assess*. 2012;16:1-313.

31. Ashby J, O'Hanlon M, Buxton MJ. The time trade-off technique: How do the valuations of breast cancer patients compare to those of other groups? *Qual Life Res*. 1994:3;257-265.

32. Clinical Outcomes of Surgical Therapy Group. COST trial: Clinical outcomes of surgical therapy group. A comparison of laparoscopically assisted and open colectomy for colon cancer. *N Engl J Med*. 2004;350:2050-2059.

33. Hazebroek EJ; The COLOR Study Group,. COLOR: A randomized clinical trial comparing laparascopic and open resection for colon cancer. *Surg Endosc*. 2002;16:949-953.

34. Cella DF, Tulsky DS, Gray G, et al. The functional assessment of cancer therapy scale: Development and validation of the general measure. *J Clin Oncol*. 1993;11:570-579.

35. Drummond M, Griffin A, Tarricone R. Economic evaluation of medical devices and drugs - same or different. *Value Health*. 2009;12:402-403.

36. Sorenson C, Tarricone R, Siebert M, et al. Applying health economics for policy decision making: Do devices differ from drugs? *Eurospace*. 2011;13(Suppl 2):ii54-ii58.

37. nice.org.uk [Internet] [updated February 1, 2013]. http://www.nice.org.uk (accessed February 13, 2013).

38. hta.ac.uk [Internet] [updated February 1, 2013]. http://www.hta.ac.uk (accessed February 13, 2013).

39. The Guideline Development Group. *Early and locally advanced breast cancer: Diagnosis and treatment*. Clinical Guideline CG80 2009 London 2009. http://www.guidancenice.org.uk/CG80 (accessed February 13, 2013).

40. NCRI Secretariat. National Cancer Research Institute. *Challenges and opportunities in surgical cancer research in the UK*. http://www.ncri.org.uk (accessed April 2, 2013).

41. The Guideline Development Group. *Colorectal cancer: The diagnosis and management of colorectal cancer*. Clinical Guideline CG131 2011 London 2011. http://www.nice.org.uk/CG131 (accessed February 13, 2013).

42. NICE. *Laparoscopic surgery for the treatment of colorectal cancer* (review). NICE technology appraisal guidance 105 (2006). London 2006. http//www.nice.org.uk/TA105 (accessed February 13, 2013).