NIHR | National Institute for Health Research

Journals Library

Check for updates

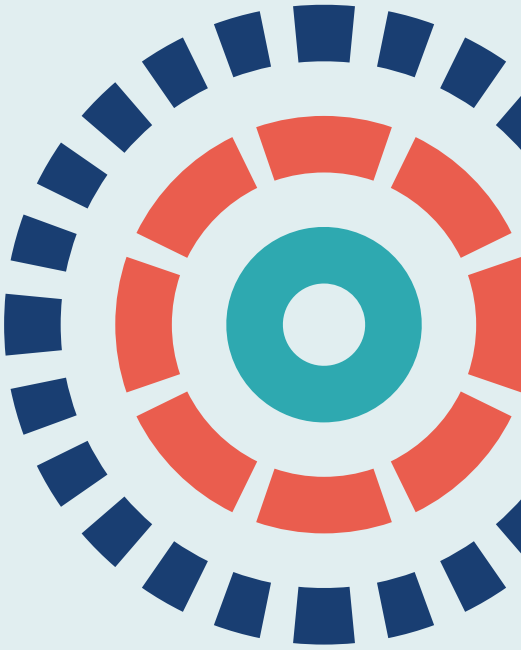
Efficacy and Mechanism Evaluation

Volume 6 • Issue 10 • September 2019

ISSN 2050-4365

Robotic-assisted surgery compared with laparoscopic resection surgery for rectal cancer: the ROLARR RCT

David Jayne, Alessio Pigazzi, Helen Marshall, Julie Croft, Neil Corrigan, Joanne Copeland, Philip Quirke, Nicholas West, Richard Edlin, Claire Hulme and Julia Brown



DOI 10.3310/eme06100

Robotic-assisted surgery compared with laparoscopic resection surgery for rectal cancer: the ROLARR RCT

David Jayne,[1]* Alessio Pigazzi,[2] Helen Marshall,[3] Julie Croft,[3] Neil Corrigan,[3] Joanne Copeland,[3] Philip Quirke,[4] Nicholas West,[4] Richard Edlin,[5] Claire Hulme[6] and Julia Brown[3]

[1]Academic Surgery, Leeds Institute of Biological and Clinical Sciences, University of Leeds, Leeds, UK
[2]Department of Surgery, University of California, Irvine, CA, USA
[3]Clinical Trials Research Unit, Leeds Institute of Clinical Trials Research, University of Leeds, Leeds, UK
[4]Pathology and Tumour Biology, Leeds Institute of Cancer and Pathology, University of Leeds, Leeds, UK
[5]Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand
[6]Academic Unit of Health Economics, University of Leeds, Leeds, UK

*Corresponding author

Declared competing interests of authors: Alessio Pigazzi is a consultant and proctor for Intuitive Surgical Inc. (Sunnyvale, CA, USA) and receives personal fees from Covidien plc (Medtronic plc; Dublin, Ireland) and Ethicon, Inc. (Somerville, NJ, USA) outside the submitted work. David Jayne is a proctor for Intuitive Surgical Inc. and was formerly a member of the Efficacy and Mechanism Evaluation (EME) Strategy Group and the EME Prioritisation Group, and was previously involved in an EME Intraoperative Imaging Review. Claire Hulme was formerly a member of the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) Commissioning Board. Julia Brown is a member of the HTA Clinical Evaluation and Trials Board, HTA Funding Board Policy Group, HTA Mental, Psychological and Occupational Health Methods Group, HTA Post-Board Funding Teleconference Group and NIHR Standing Advisory Committee.

Published September 2019
DOI: 10.3310/eme06100

This report should be referenced as follows:

Jayne D, Pigazzi A, Marshall H, Croft J, Corrigan N, Copeland J, *et al*. Robotic-assisted surgery compared with laparoscopic resection surgery for rectal cancer: the ROLARR RCT. *Efficacy Mech Eval* 2019;**6**(10).

Efficacy and Mechanism Evaluation

ISSN 2050-4365 (Print)

ISSN 2050-4373 (Online)

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full EME archive is freely available to view online at www.journalslibrary.nihr.ac.uk/eme. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Efficacy and Mechanism Evaluation* journal
Reports are published in *Efficacy and Mechanism Evaluation* (EME) if (1) they have resulted from work for the EME programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

EME programme

The Efficacy and Mechanism Evaluation (EME) programme funds ambitious studies evaluating interventions that have the potential to make a step-change in the promotion of health, treatment of disease and improvement of rehabilitation or long-term care. Within these studies, EME supports research to improve the understanding of the mechanisms of both diseases and treatments.

The programme support translational research into a wide range of new or repurposed interventions. These may include diagnostic or prognostic tests and decision-making tools, therapeutics or psychological treatments, medical devices, and public health initiatives delivered in the NHS.

The EME programme supports clinical trials and studies with other robust designs, which test the efficacy of interventions, and which may use clinical or well-validated surrogate outcomes. It only supports studies in man and where there is adequate proof of concept. The programme encourages hypothesis-driven mechanistic studies, integrated within the efficacy study, that explore the mechanisms of action of the intervention or the disease, the cause of differing responses, or improve the understanding of adverse effects. It funds similar mechanistic studies linked to studies funded by any NIHR programme.

The EME programme is funded by the Medical Research Council (MRC) and the National Institute for Health Research (NIHR), with contributions from the Chief Scientist Office (CSO) in Scotland and National Institute for Social Care and Health Research (NISCHR) in Wales and the Health and Social Care Research and Development (HSC R&D), Public Health Agency in Northern Ireland.

This report

The research reported in this issue of the journal was funded by the EME programme as project number 08/52/01. The contractual start date was in March 2010. The final report began editorial review in March 2018 and was accepted for publication in March 2019. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The EME editors and production house have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research. The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the MRC, NETSCC, the EME programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the EME programme or the Department of Health and Social Care.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Abstract

Robotic-assisted surgery compared with laparoscopic resection surgery for rectal cancer: the ROLARR RCT

David Jayne,[1]* Alessio Pigazzi,[2] Helen Marshall,[3] Julie Croft,[3] Neil Corrigan,[3] Joanne Copeland,[3] Philip Quirke,[4] Nicholas West,[4] Richard Edlin,[5] Claire Hulme[6] and Julia Brown[3]

[1]Academic Surgery, Leeds Institute of Biological and Clinical Sciences, University of Leeds, Leeds, UK
[2]Department of Surgery, University of California, Irvine, CA, USA
[3]Clinical Trials Research Unit, Leeds Institute of Clinical Trials Research, University of Leeds, Leeds, UK
[4]Pathology and Tumour Biology, Leeds Institute of Cancer and Pathology, University of Leeds, Leeds, UK
[5]Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand
[6]Academic Unit of Health Economics, University of Leeds, Leeds, UK

*Corresponding author d.g.jayne@leeds.ac.uk

Background: Robotic rectal cancer surgery is gaining popularity, but there are limited data about its safety and efficacy.

Objective: To undertake an evaluation of robotic compared with laparoscopic rectal cancer surgery to determine its safety, efficacy and cost-effectiveness.

Design: This was a multicentre, randomised trial comparing robotic with laparoscopic rectal resection in patients with rectal adenocarcinoma.

Setting: The study was conducted at 26 sites across 10 countries and involved 40 surgeons.

Participants: The study involved 471 patients with rectal adenocarcinoma. Recruitment took place from 7 January 2011 to 30 September 2014 with final follow-up on 16 June 2015.

Interventions: Robotic and laparoscopic rectal cancer resections were performed by high anterior resection, low anterior resection or abdominoperineal resection. There were 237 patients randomised to robotic and 234 to laparoscopic surgery. Follow-up was at 30 days, at 6 months and annually until 3 years after surgery.

Main outcome measures: The primary outcome was conversion to laparotomy. Secondary end points included intra- and postoperative complications, pathological outcomes, quality of life (QoL) [measured using the Short Form questionnaire-36 items version 2 (SF-36v2) and the Multidimensional Fatigue Inventory-20 (MFI-20)], bladder and sexual dysfunction [measured using the International Prostatic Symptom Score (I-PSS), the International Index of Erectile Function (IIEF) and the Female Sexual Function Index (FSFI)], and oncological outcomes. An economic evaluation considered the costs of robotic and laparoscopic surgery, including primary and secondary care costs up to 6 months post operation.

Results: Among 471 randomised patients [mean age 64.9 years, standard deviation (SD) 11.0 years; 320 (67.9%) men], 466 (98.9%) patients completed the study. Data were analysed on an intention-to-treat basis. The overall rate of conversion to laparotomy was 10.1% and occurred in 19 (8.1%) patients in the robotic-assisted group and in 28 (12.2%) patients in the conventional laparoscopic group {unadjusted risk

difference 4.12% [95% confidence interval (CI) –1.35% to 9.59%], adjusted odds ratio 0.61 [95% CI 0.31 to –1.21]; $p = 0.16$}. Of the nine prespecified secondary end points, including circumferential resection margin positivity, intraoperative complications, postoperative complications, plane of surgery, 30-day mortality and bladder and sexual dysfunction, none showed a statistically significant difference between the groups. No difference between the treatment groups was observed for longer-term outcomes, disease-free and overall survival (OS). Males were at a greater risk of local recurrence than females and had worse OS rates. The costs of robotic and laparoscopic surgery, excluding capital costs, were £11,853 (SD £2940) and £10,874 (SD £2676) respectively.

Conclusions: There is insufficient evidence to conclude that robotic rectal surgery compared with laparoscopic rectal surgery reduces the risk of conversion to laparotomy. There were no statistically significant differences in resection margin positivity, complication rates or QoL at 6 months between the treatment groups. Robotic rectal cancer surgery was on average £980 more expensive than laparoscopic surgery, even when the acquisition and maintenance costs for the robot were excluded.

Future work: The lower rate of conversion to laparotomy in males undergoing robotic rectal cancer surgery deserves further investigation. The introduction of new robotic systems into the market may alter the cost-effectiveness of robotic rectal cancer surgery.

Trial registration: Current Controlled Trials ISRCTN80500123.

Funding: This project was funded by the Efficacy and Mechanism Evaluation (EME) programme, a Medical Research Council and National Institute for Health Research (NIHR) partnership, with contributions from the Chief Scientist Office, Scottish Government Health and Social Care Directorate, the Health and Care Research Wales and the Health and Social Care Research and Development Division, Public Health Agency in Northern Ireland. The funders of the study had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; and preparation, review or approval of the manuscript or the decision to submit for publication. The project will be published in full in *Efficacy and Mechanism Evaluation*; Vol. 6, No. 10. See the NIHR Journals Library website for further project information. Philip Quirke and Nicholas West were supported by Yorkshire Cancer Research Campaign and the MRC Bioinformatics initiative. David Jayne was supported by a NIHR Research Professorship.

Contents

List of tables

List of figures

List of abbreviations

AIC	Akaike information criterion
APR	abdominoperineal resection
ASA	American Society of Anesthesiologists
BMI	body mass index
CI	confidence interval
CRF	case report form
CRM	circumferential resection margin
CRM+	circumferential resection margin positivity
CT	computed tomography
CTRU	Clinical Trials Research Unit
DFS	disease-free survival
DMEC	Data Monitoring and Ethics Committee
EBE	empirical Bayes' estimate
EQ-5D	EuroQol-5 Dimensions
EQ-5D-3L	EuroQol-5 Dimensions, three-level version
FSFI	Female Sexual Function Index
GP	general practitioner
HAR	high anterior resection
HCHS	Health and Community Health Service
HR	hazard ratio
HRQoL	health-related quality of life
ICC	intracluster correlation coefficient
ICER	incremental cost-effectiveness ratio
ID	identifier
IIEF	International Index of Erectile Function
I-PSS	International Prostatic Symptom Score
IQR	interquartile range
IRB	institutional review board
ISRCTN	International Standard Randomised Controlled Trial Number
LAR	low anterior resection
MCS	mental component score
MFI-20	Multidimensional Fatigue Inventory-20
OR	odds ratio
OS	overall survival
PCS	physical component score
PSSRU	Personal Social Services Research Unit
QALY	quality-adjusted life-year
QoL	quality of life
SF-6D	Short Form questionnaire-6 Dimensions
SF-36v2	Short Form questionnaire-36 items version 2
TME	total mesorectal excision
TSC	Trial Steering Committee
WHO	World Health Organization

Plain English summary

Robotic systems are being used to remove cancers of the rectum (back passage), but there is little evidence that they produce better results than standard laparoscopic (keyhole) surgery. The aim of the ROLARR study was to perform a thorough investigation of the benefits of robotic rectal cancer surgery, comparing it with laparoscopic rectal cancer surgery.

A total of 471 patients with rectal cancer, from 26 hospitals in 10 countries, were allocated at random to undergo either robotic or laparoscopic surgery. Data were collected at 30 days, at 6 months and annually until 3 years following surgery.

There was no significant difference in the numbers of patients who required conversion to an open operation, involving a large cut on the abdomen, to complete their surgery between the robotic (8.1%) and laparoscopic (12.2%) treatments. Male patients undergoing robotic surgery were less likely to need an open operation. Similarly, there were no differences in surgical complications, bladder and sexual function, and quality of life between the robotic and laparoscopic surgery. Robotic surgery produced similar results to laparoscopic surgery in treating rectal cancer. Overall, males were more at risk of the cancer coming back. Robotic operations were £980 more expensive than laparoscopic operations because the surgery took longer and the robotic instruments were more expensive.

We conclude that robotic surgery does not reduce the need to perform open surgery in a small number of patients with rectal cancer. Robotic surgery is more expensive than laparoscopic surgery, with no obvious benefits for patients in the short or long term.

Scientific summary

Background

Total mesorectal excision is the standard of care in rectal cancer surgery, involving complete removal of the tumour along with the draining lymphatics within an intact mesorectal envelope. The feasibility and safety of laparoscopic surgery have been established for colon cancer, but the case for rectal cancer is less clear. At the time of the study's design in 2010, the MRC CLASICC trial [Guillou PJ, Quirke P, Thorpe H, Walker J, Jayne DG, Smith AM, *et al.* Short-term endpoints of conventional versus laparoscopic-assisted surgery in patients with colorectal cancer (MRC CLASICC trial): multicentre, randomised controlled trial. *Lancet* 2005;**365**:1718–26] was the only randomised study, to our knowledge, to include an evaluation of laparoscopic compared with open rectal cancer surgery. Concern was expressed about the higher rate of circumferential resection margin (CRM) involvement in the laparoscopic group (12.4%) than in the open group (6.3%) for patients undergoing anterior resection. This, however, did not translate into a difference in local recurrence at either 3-year follow-up or 5-year follow-up. The difference in CRM involvement was felt to reflect the increased technical difficulties associated with the laparoscopic technique in the rectal cancer group.

Since completion of the CLASICC trial, the COLOR II [van der Pas MH, Haglind E, Cuesta MA, Fürst A, Lacy AM, Hop WC, Bonjer HJ, COlorectal cancer Laparoscopic or Open Resection II (COLOR II) Study Group. Laparoscopic versus open surgery for rectal cancer (COLOR II): short-term outcomes of a randomised, phase 3 trial. *Lancet Oncol* 2013;**14**:210–18] and COREAN studies [Kang SB, Park JW, Jeong SY, Nam BH, Choi HS, Kim DW, *et al.* Open versus laparoscopic surgery for mid or low rectal cancer after neoadjuvant chemoradiotherapy (COREAN trial): short-term outcomes of an open-label randomised controlled trial. *Lancet Oncol* 2010;**11**:637–45] compared laparoscopic with open surgery for rectal cancer. Both studies reported better short-term outcomes following laparoscopic rectal cancer resection than open surgery, and similar pathological outcomes compared with open surgery. In contrast, there have been two large randomised trials, AlaCarte (Stevenson AR, Solomon MJ, Lumley JW, Hewett P, Clouston AD, Gebski VJ, *et al.* Effect of laparoscopic-assisted resection vs open resection on pathological outcomes in rectal cancer: the ALaCaRT randomized clinical trial. *JAMA* 2015;**314**:1356–63) and ACOSOG (Fleshman J, Branda M, Sargent DJ, Boller AM, George V, Abbas M, *et al.* Effect of laparoscopic-assisted resection vs open resection of stage II or III Rectal Cancer on Pathologic Outcomes: the ACOSOG Z6051 randomized clinical trial. *JAMA* 2015;**314**:1346–55), that have cast doubt on the benefits of laparoscopic compared with open rectal cancer surgery.

Robotic-assisted laparoscopic surgery was introduced with the promise to eliminate many of the technical difficulties inherent in laparoscopic surgery, providing intuitive manipulation of the laparoscopic instruments with 7 degrees of freedom of movement, a three-dimensional field of view, a stable camera platform with zoom magnification, dexterity enhancement and an ergonomic operating environment.

There have been numerous reports from single centres, analyses of national databases and several systematic reviews and meta-analyses, but no large randomised comparison with laparoscopic or open rectal cancer surgery. Results from the meta-analyses tell a broadly similar story, with no clear advantage of robotic over laparoscopic surgery in terms of short-term outcomes, with the exception of lower conversion rates and a suggestion of improved postoperative bladder and sexual function. The main disadvantage of robotic, compared with laparoscopic, surgery is the increased hospital costs.

Objectives

The purpose of the trial was to perform a rigorous evaluation of robotic-assisted rectal cancer surgery compared with laparoscopic rectal cancer surgery by means of a randomised controlled trial. The key short-term outcomes included assessment of technical ease of the operation, as determined by the clinical indicator of low conversion rate to open operation, and clear pathological resection margins as an indicator of surgical accuracy and improved oncological outcome. In addition, quality-of-life (QoL) assessment and analysis of cost-effectiveness were performed. Longer-term outcomes concentrated on oncological aspects of the surgery with analysis of disease-free survival (DFS) and overall survival (OS) and local recurrence rates at the 3-year follow-up.

Methods

The ROLARR trial was an international, multicentre, prospective, randomised, controlled, unblinded, parallel-group trial comparing robotic-assisted versus laparoscopic surgery for the curative treatment of rectal cancer. Participating surgeons had to have previously performed a minimum of 30 minimally invasive (laparoscopic or robotic) rectal cancer resections (at least 10 laparoscopic and at least 10 robotic). The trial received national ethics approval in the UK and either ethics committee approval or institutional review board (IRB) approval as required at the location of each of the international centres; all participants gave written informed consent.

Patients were eligible if they were aged $\geq$ 18 years with a diagnosis of rectal adenocarcinoma amenable to curative surgery by low anterior resection, high anterior resection (HAR) or abdominoperineal resection (APR). Patients had to be suitable and fit for robotic-assisted or standard laparoscopic rectal resection. Exclusion criteria included locally advanced cancers not amenable to curative surgery or requiring en bloc multivisceral resection, synchronous colorectal tumours, coexistent inflammatory bowel disease, malignancy within the past 5 years, or pregnancy.

Preoperative investigation and preparation was as per institutional protocol. Laparoscopic mesorectal resection was performed in accordance with each surgeon's usual practice. Robotic surgery involved either a totally robotic approach or a hybrid approach; the only absolute requirement was that the robot had to be used for mesorectal resection. The specifics of each operation were at the discretion of the operating surgeon, as was the decision to convert to an open operation. Detailed guidance was provided to ensure consistent histopathological analysis and reporting of the rectal dissection specimens according to internationally agreed criteria. Digital photographs of the specimen and sequential cross-sectional views were collected to allow blinded assessment of the quality of the plane of surgery. To enable a central pathology review, the tissue slides (or high-quality digital slide images) were submitted.

Postoperative care was as per institutional protocol; however, the protocol required that patients underwent a clinical assessment at 30 days and at 6 months post operation. Follow-up data were collected on an annual basis until the last participant reached 3 years post randomisation.

Participants completed questionnaires prior to randomisation (baseline) and at 30 days and at 6 months postoperatively. General QoL [Short Form questionnaire-36 items version 2 (SF-36v2)] and fatigue [Multidimensional Fatigue Inventory-20 (MFI-20)] data were collected at baseline and at 30 days and at 6 months postoperatively. In addition, patient-reported bladder and sexual function questionnaires [International Prostatic Symptom Score (I-PSS) and International Index of Erectile Function/ Female Sexual Function Index (IIEF/FSFI)] were completed at baseline and at 6 months post operation. Participants in the UK and USA also completed the EuroQol-5 Dimensions (EQ-5D) at baseline, at 30 days and at 6 months post operation, and a resource utilisation questionnaire at 30 days and at 6 months post operation for the health economic component of the trial.

Results

Between 7 January 2011 and 30 September 2014, 1276 patients were assessed for eligibility by 40 surgeons from 26 sites across 10 countries (i.e. UK, Italy, Denmark, USA, Finland, South Korea, Germany, France, Australia and Singapore). The numbers of patients recruited in each country (together with the number of sites in the country) were as follows: UK, $n = 131$ (6); Italy, $n = 105$ (5); Denmark, $n = 92$ (3); USA, $n = 59$ (9); Finland, $n = 35$ (1); South Korea, $n = 18$ (1); Germany, $n = 16$ (1); France, $n = 11$ (1); Australia, $n = 2$ (1); and Singapore, $n = 2$ (1). Four hundred and seventy-one (36.9%) of these patients were randomised: 234 to laparoscopic and 237 to robotic surgery. From this group, 466 patients underwent an operation, with 456 (97.9%) undergoing the allocated treatment. The final follow-up date was 16 June 2015.

The two treatment groups were well balanced with respect to baseline characteristics and operative procedures. On average, patients received an operation performed by a surgeon with experience of around a median of 91 [interquartile range (IQR) 45–180] previous laparoscopic and a median of 50 (IQR 30–101) previous robotic operations.

The rate of conversion to open surgery was 47 out of 466 (10.1%) patients overall, 28 out of 230 (12.2%) in the laparoscopic group and 19 out of 236 (8.1%) in the robotic group (unadjusted difference in proportions 4.12%, 95% CI –1.35% to 9.59%). There was no statistically significant difference between robotic surgery and conventional laparoscopic surgery with respect to odds of conversion [adjusted odds ratio (OR) 0.614, 95% CI 0.311 to 1.211; $p = 0.16$].

Of the 466 patients who had an operation, 459 (98.5%) had complete pathology data available. Furthermore, 26 out of 459 (5.7%) patients had a positive CRM: 14 out of 224 (6.25%) in the laparoscopic group and 12 out of 235 (5.11%) in the robotic group (unadjusted difference in proportions 1.14%, 95% CI –3.10% to 5.38%). There was no statistically significant difference in the odds of CRM positivity between the groups (adjusted OR 0.785, 95% CI 0.350 to 1.762; $p = 0.56$).

There were 70 out of 466 (15.0%) patients who had an intraoperative complication, 34 out of 230 (14.8%) in the laparoscopic group and 36 out of 236 (15.3%) in the robotic group (unadjusted risk difference –0.5%, 95% CI –6.0% to 7.0%). There was no significant difference between the groups (adjusted OR 1.020, 95% CI 0.599 to 1.736; $p = 0.94$). There was a significant difference in the odds of having an intraoperative complication between males and females (adjusted OR 3.083, 95% CI 1.543 to 6.158; $p = 0.0015$).

There were 151 out of 466 (32.4%) patients who had a postoperative complication within 30 days of their operation, 73 out of 230 (31.7%) in the laparoscopic group and 78 out of 236 (33.1%) in the robotic group (unadjusted risk difference –1.3%, 95% CI –9.8% to 7.2%). There was no significant difference between the treatment groups (adjusted OR 1.043, 95% CI 0.689 to 1.581; $p = 0.84$). There was a significant difference in the odds of having a postoperative complication within 30 days of operation between males and females (adjusted OR 3.083, 95% CI 1.573 to 4.183; $p = 0.0002$).

There were 72 out of 466 (15.5%) patients who had a postoperative complication after 30 days and within 6 months of their operation, 38 out of 230 (16.5%) in the laparoscopic group and 34 out of 236 (14.4%) in the robotic group (unadjusted risk difference 2.1%, 95% CI –4.5% to 8.7%). There was no significant difference between the groups (adjusted OR 0.719, 95% CI 0.411 to 1.258; $p = 0.25$).

Bladder function scores, as measured by the I-PSS, were similar between the groups at baseline and at 6 months. The adjusted estimated difference in mean I-PSS (robotic minus standard) was –0.7426 (95% CI –2.0722 to 0.5870; $p = 0.2726$). The estimated difference in mean I-PSS between patients with a difference in baseline score of 10 points, all else being equal, was 4.20 (95% CI 3.23 to 5.17; $p < 0.0001$).

The distribution of sexual function scores, as measured by the IIEF, was very similar between the treatment groups at baseline and at 6 months. Median IIEF scores at 6 months were notably lower than at baseline in both groups; the estimated difference in mean IIEF (robotic minus standard) was –0.8020 (95% CI –5.7039 to 4.1000; $p = 0.7468$).

The female sexual function score, as measured by the FSFI, at baseline was marginally lower in the robotic group. The distribution of scores was very similar between the treatment groups at 6 months; the estimated difference in mean FSFI (robotic minus standard) was –1.2309 (95% CI –6.0030 to 3.5413; $p = 0.6010$).

Patient-reported generic health was measured using the SF-36v2, providing a physical component score (PCS) and a mental component score (MCS). The baseline PCS and MCS were similar in the two treatment groups at all time points. At the 6-month follow-up, the adjusted estimated difference in mean PCS between the groups (robotic vs. laparoscopic) was –0.1220 (95% CI –1.6281 to 1.3840; $p = 0.8737$). The adjusted estimated difference in MCS between the groups (robotic vs. laparoscopic) was –0.4875 (95% CI –2.6008 to 1.6258; $p = 0.6508$).

The Multidimensional Fatigue Inventory is a self-report instrument consisting of five scales of fatigue: general fatigue, physical fatigue, reduced activity, reduced motivation and mental fatigue. The distribution of scores was similar between the two treatment groups at all time points for all five scales. At the 6-month follow-up, the estimated adjusted difference in mean general fatigue between the groups (robotic vs. laparoscopic) was –0.2517 (95% CI –0.5965 to 1.0999; $p = 0.5603$), the difference in physical fatigue was 0.3964 (95% CI –0.4404 to 1.2332; $p = 0.3527$), the difference in reduced activity was –0.1634 (95% CI –0.9777 to 0.6510; $p = 0.6938$), the difference in reduced motivation was –0.03917 (95% CI –0.7324 to 0.6540; $p = 0.9117$) and the difference in mental fatigue was 0.1374 (95% CI –0.6626 to 0.9374; $p = 0.7360$).

A total of 351 out of 456 (77.0%) patients' specimens were graded by pathological assessment of the plane of surgery. There were 178 out of 233 (76.4%) in the laparoscopic group and 173 out of 223 (77.6%) in the robotic group who had best-quality surgery (mesorectal plane) (unadjusted risk difference 1.2%, 95% CI –6.5% to 8.9%). There was no significant difference of the odds of a mesorectal plane surgery between the groups (adjusted OR 0.943, 95% CI 0.565 to 1.572; $p = 0.821$).

Local recurrence was observed in 30 out of 471 (6.4%) patients, 14 out of 234 (6.0%) in the laparoscopic group and 16 out of 237 (6.8%) in the robotic group. There was no difference between the treatment groups in local recurrence rates at the 3-year follow-up; the estimated difference in cumulative incidence of local recurrence was 0.002 (95% CI –0.041 to 0.046). There was a difference in the probability of local recurrence between males and females, with males being more likely to experience local recurrence [adjusted hazard ratio (HR) 3.184, 95% CI 1.109 to 9.174; $p = 0.031$].

No difference was observed between the treatment groups in DFS at the 3-year follow-up, estimated adjusted HR (robotic vs. laparoscopic) of 1.030 (95% CI 0.713 to 1.489; $p = 0.874$). Disease recurrence was more common following APR and least common following HAR.

Death was observed for 46 out of 471 (9.8%) patients, 23 out of 234 (9.8%) in the laparoscopic group and 23 out of 237 (9.7%) in the robotic group, estimated HR (robotic vs. laparoscopic) 0.945 (95% CI 0.530 to 1.686; $p = 0.848$). Males were 2.187 (95% CI 1.017 to 4.700; $p = 0.045$) times more likely to die than females at 3 years' follow-up.

Quality-of-life scores were very similar between the treatment groups, with a difference in favour of robotic surgery of 0.013 quality-adjusted life-years (QALYs) at 6 months' follow-up. The overall cost difference was £980, with higher costs associated with robotic surgery, driven by longer operating times and higher instrument costs. The estimated incremental cost-effectiveness ratio (ICER) for robotic surgery was £69,837 per QALY, which is well in excess of the standard threshold of £20,000–30,000 per QALY.

Conclusions

Robotic rectal cancer surgery results in comparable outcomes to laparoscopic surgery. There is no statistical benefit in terms of conversion to open surgery, bladder or sexual function, pathological outcomes, or DFS and OS. The observed trend to reduced conversion in male patients requires further confirmation. Robotic rectal cancer surgery is not cost-effective compared with laparoscopic rectal cancer surgery because the increased costs far outweigh any marginal benefit in QoL.

Trial registration

This trial is registered as ISRCTN80500123.

Funding

This project was funded by the Efficacy and Mechanism Evaluation (EME) programme, a Medical Research Council and National Institute for Health Research (NIHR) partnership, with contributions from the Chief Scientist Office, Scottish Government Health and Social Care Directorate, the Health and Care Research Wales and the Health and Social Care Research and Development Division, Public Health Agency in Northern Ireland. The funders of the study had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; and preparation, review or approval of the manuscript or the decision to submit for publication. Philip Quirke and Nicholas West were supported by Yorkshire Cancer Research Campaign and the MRC Bioinformatics initiative. David Jayne was supported by a NIHR Research Professorship.

Chapter 1 Introduction

Total mesorectal excision (TME) is the standard of care in rectal cancer surgery, involving complete removal of the tumour along with the draining lymphatics within an intact mesorectal envelope.[1] The feasibility and safety of laparoscopic surgery has been established for colon cancer.[2–4] The case for rectal cancer is less clear, and, of the reported multicentre trials at the time of study design in 2010, only the MRC CLASICC trial included an evaluation of laparoscopic rectal cancer surgery compared with open rectal cancer surgery.[5] Although both laparoscopic and open rectal cancer resection were associated with similar lymph node yields, concern was expressed at the higher rate of circumferential resection margin (CRM) involvement in the laparoscopic group (12.4%) than in the open group (6.3%) for patients undergoing anterior resection (AR). This, however, did not translate into a difference in local recurrence at either 3-year[2] or 5-year follow-up.[6] The difference in CRM involvement was felt to reflect the increased technical difficulties associated with the laparoscopic technique in the rectal cancer subgroup. This was supported by the higher conversion rate in the laparoscopic rectal subgroup (34%) than the laparoscopic colon subgroup (25%).[5] Analysis of CLASICC data revealed higher morbidity and mortality rates associated with laparoscopic cases converted to open operation. Some of this increased morbidity may be related to more advanced cancers requiring conversion, but a proportion of it will inevitably have resulted from the increased operative time, increased technical difficulty and the need for a laparotomy wound in converted cases.

Since completion of the CLASICC trial, there have been several other large studies comparing laparoscopic with open surgery for rectal cancer. A large European randomised controlled trial, COLOR II, recruited 1103 participants to a non-inferiority study involving 30 centres in eight countries.[7] Laparoscopic surgery was reported to be advantageous in terms of short-term outcomes (quicker return of bowel function, reduced hospital stay), with similar morbidity and pathological outcomes to open surgery. The 3-year results from the same study were reported in 2015 and showed similar rates of locoregional recurrence and disease-free survival (DFS) and overall survival (OS) in both the laparoscopic and the open groups.[8] These findings were echoed by the results of the COREAN trial, which again reported better short-term outcomes following laparoscopic rectal cancer resection and similar pathological outcomes compared with open surgery.[9]

In contrast, there have been two large randomised trials, ALaCaRT[10] and ACOSOG,[11] that have cast doubt on the benefits of laparoscopic rectal cancer surgery compared with open rectal cancer surgery. Both were non-inferiority studies and both used a novel composite primary outcome combining rates of negative circumferential and distal cancer margins with completeness of mesorectal excision as a measure of oncological clearance. Both studies failed to demonstrate the non-inferiority of the laparoscopic compared with the open surgery approach, concluding that the evidence was not sufficient to support the routine use of the laparoscopic technique.

Robotic-assisted laparoscopic surgery was introduced into clinical practice in the early 1990s with the promise to eliminate many of the technical difficulties inherent in laparoscopic surgery. The technical advantages associated with robotic-assisted surgery include intuitive manipulation of the laparoscopic instruments with 7 degrees of freedom of movement, a three-dimensional field of view, a stable camera platform with zoom magnification, dexterity enhancement and an ergonomic operating environment.

The feasibility of robotics for TME rectal cancer resection was established by Pigazzi *et al.* in a series of six low rectal cancers.[12] A subsequent follow-up study of 39 rectal cancers treated prospectively by robotic-assisted resection reported a zero rate of conversion with a mortality of 0% and morbidity of 12.8%.[13] The only randomised trial at the time of design of the ROLARR study compared 18 patients assigned to robotic-assisted resection with 18 patients assigned to standard laparoscopic resection.[14] No difference was observed in the operative times, the conversion rates (two laparoscopic, zero robotic), or the quality of mesorectal resection. The only difference was the length of hospital stay, which was significantly shorter following robotic-assisted laparoscopic surgery (robotic assisted: 6.9 ± 1.3 days; standard laparoscopic: 8.7 ± 1.3 days; $p < 0.001$) and

attributed by the authors to a reduction in surgical trauma. Since the commencement of the ROLARR trial, there have been numerous reports from single centres, analyses of national databases,[15] and several systematic reviews and meta-analyses,[16–18] but no large randomised comparison with laparoscopic or open rectal cancer surgery. Results from the meta-analyses tell a broadly similar story, with no clear advantage for robotic over laparoscopic surgery in terms of short-term outcomes, with the exception of lower conversion rates and a suggestion of improved postoperative bladder and sexual function.[19] The disadvantage of robotic surgery, compared with laparoscopic surgery, appears to be the longer operating times and perhaps an increase in operative blood loss. Importantly, the hospital costs associated with the use of the robot are higher, which has fuelled the ongoing debate about whether or not robotic-assisted rectal cancer surgery can be justified in the absence of clear patient benefits and considering its higher hospital costs.[15,20,21]

The ROLARR trial was designed with the above concerns in mind and with the primary objective to evaluate the short-term safety and efficacy of robotic-assisted surgery compared with laparoscopic surgery for rectal cancer resection. The primary end point chosen was conversion to open surgery, on the basis that if the robot offered a technical advantage over laparoscopic surgery it should be reflected in a reduced conversion rate. Secondary end points were chosen to reflect the oncological nature of the investigation and the compelling need for rigorous patient-reported outcomes and cost-effectiveness evaluation.

Chapter 2 Methods

Objectives

The purpose of the trial was to perform a rigorous evaluation of robotic-assisted rectal cancer surgery by means of a randomised controlled trial. The chosen comparator was standard laparoscopic rectal cancer resection, which is essentially the same procedure but without the use of the robotic device. The two operative interventions were evaluated for short- and longer-term outcomes. The key short-term outcomes included assessment of technical ease of the operation, as determined by the clinical indicator of low conversion rate to open operation, and clear pathological resection margins as an indicator of surgical accuracy and improved oncological outcome. In addition, quality-of-life (QoL) assessment and analysis of cost-effectiveness were performed to aid evidence-based knowledge to inform the NHS and other service providers and decision-makers. The short-term outcomes were analysed after the last randomised patient had had 6 months of follow-up, to provide a timely assessment of the new technology, and were made available to the public, clinicians and health-care providers to inform health-care decision-making. Longer-term outcomes concentrated on oncological aspects of the disease and its surgical treatment with analysis of DFS, OS and local recurrence rates at 3 years' follow-up.

Trial design

The ROLARR trial was an international, multicentre, prospective, unblinded, parallel-group randomised controlled trial[22] comparing robotic-assisted with laparoscopic surgery for the curative treatment of rectal cancer (defined as an adenocarcinoma whose distal extent was situated at or within 15 cm of the anal margin) by low anterior resection (LAR), high anterior resection (HAR) or abdominoperineal resection (APR). The trial design required that each participating surgeon had performed a minimum of 30 minimally invasive (laparoscopic or robotic) rectal cancer resections (at least 10 laparoscopic and at least 10 robotic). The trial received national ethics approval in the UK and either ethics committee or institutional review board (IRB) approval as was required at the location of each of the international centres; all participants gave written informed consent. The trial conduct was overseen by an independent Trial Steering Committee (TSC) and Data Monitoring and Ethics Committee (DMEC). The trial was registered on the International Standard Randomised Controlled Trial Number (ISRCTN) register (ISRCTN80500123).

Participants

The inclusion criteria were:

1. Aged $\geq$ 18 years.
2. Able to provide written informed consent.
3. Diagnosis of rectal cancer (defined as an adenocarcinoma for which distal extent is situated at or within 15 cm of the anal margin, as assessed by endoscopic examination or radiological contrast study) amenable to curative surgery by LAR, HAR or APR, for example, staged T1–3, N0–2, M0 by imaging as per local practice. Although not mandated, computed tomography (CT) imaging with either additional magnetic resonance imaging (MRI) or transrectal ultrasound is recommended to assess distant and local disease.
4. Rectal cancer suitable for resection by either standard laparoscopic procedure or robotic-assisted laparoscopic procedure.
5. Fit for robotic-assisted or standard laparoscopic rectal resection.

6. An American Society of Anesthesiologists (ASA) physical status of ≤ 3.
7. Capable of completing required questionnaires at time of consent (provided questionnaires were available in a language spoken fluently by the participant).

The exclusion criteria were:

1. benign lesions of the rectum
2. benign or malignant diseases of the anal canal
3. locally advanced cancers not amenable to curative surgery
4. locally advanced cancers requiring en bloc multivisceral resection
5. synchronous colorectal tumours requiring multisegment surgical resection (a benign lesion within the resection field in addition to the main cancer would not exclude a patient)
6. coexistent inflammatory bowel disease
7. clinical or radiological evidence of metastatic spread
8. concurrent or previous diagnosis of invasive cancer within 5 years that could confuse diagnosis (non-melanomatous skin cancer or superficial bladder cancer treated with curative intent were acceptable; other cases were individually discussed with the chief investigator)
9. history of psychiatric or addictive disorder or other medical condition that in the opinion of the investigator would preclude the patient from meeting the trial requirements
10. pregnancy
11. participation in another rectal cancer clinical trial relating to surgical technique.

Preoperative investigation and preparation was as per institutional protocol. Laparoscopic mesorectal resection was performed in accordance with each surgeon's usual practice. Robotic surgery involved either a totally robotic approach or a hybrid approach; the only absolute requirement was that the robot had to be used for mesorectal resection. For the purposes of the trial, a totally robotic operation was defined as a resection of the entire surgical specimen with the use of robotic assistance. A hybrid operation was defined as use of laparoscopic techniques to mobilise the proximal colon, with robotic assistance employed to perform the rectal mesorectal dissection. It was permissible to perform a partial mesorectal excision with a suitable distal margin, rather than a TME.

The specifics of each operation were at the discretion of the operating surgeon (e.g. port-site placement, mobilisation of the splenic flexure, inferior mesenteric artery/vein division, high vs. low vascular division, etc.), as was the decision to convert to an open operation. Detailed guidance was provided to ensure consistent histopathological analysis and reporting of the rectal dissection specimens in accordance with internationally agreed criteria.[23] Digital photographs of the anterior and posterior of the specimen and sequential cross-sectional views of the surgical specimen, as well as close-ups of the front and back of the levator/anal sphincter (if appropriate), were collected (prior to dissection) to allow blinded assessment of the quality of the plane of surgery. To enable a central pathology review, the tissue slides (or high-quality digital slide images) were submitted.

Postoperative care was as per institutional protocol; however, the protocol required that patients underwent a clinical assessment at 30 days and at 6 months post operation. Any further visits were in accordance with local standard clinical practice. Follow-up data were collected on an annual basis until the last participant reached 3 years post randomisation.

Participants completed questionnaires prior to randomisation (baseline), and at 30 days and 6 months postoperatively. General QoL [Short Form questionnaire-36 items version 2 (SF-36v2)] and fatigue [Multidimensional Fatigue Inventory-20 (MFI-20)] data were collected at baseline and at the 30-day and 6-month postoperative visits. In addition, bladder and sexual function questionnaires [International Prostatic Symptom Score (I-PSS) and International Index of Erectile Function/Female Sexual Function Index (IIEF/FSFI)] were completed by patients at baseline and at 6 months post operation. Participants in the UK and USA also completed the EuroQol-5 Dimensions (EQ-5D) at baseline and at 30 days and at 6 months

post operation, and a resource utilisation questionnaire at 30 days and at 6 months post operation for the health economic component of the trial.

The SF-36v2,[24] a well-validated, multipurpose standard health-related QoL evaluation questionnaire, was used to assess generic QoL. It generates an eight-scale profile of functional health and well-being scores, as well as summary measures of physical and mental health. This information related to the previous 4-week time period.

The MFI-20 was used to assess fatigue;[25] it is a 20-item self-report validated instrument designed to measure current fatigue. It creates a global score as well as individual scale scores that cover the following dimensions: general fatigue, physical fatigue, reduced activity, reduced motivation and mental fatigue.

The I-PSS[26] was used to assess bladder function. This questionnaire includes seven questions relating to lower urinary tract function, which form an overall symptom score that can be used to classify bladder dysfunction as mild, moderate or severe. To assess sexual function, the IIEF[27] and FSFI[28] were used. Both are brief male-/female-specific questionnaires developed to assess various domains of sexual function. The IIEF, FSFI and I-PSS questionnaires obtained information relating to the patient's functioning over the previous 4 weeks.

For the health economic analysis, the EQ-5D questionnaire was used to assess self-reported utility. This is a standardised non-disease-specific instrument that describes and values health-related QoL and provides a single index value for a number of different health states. In addition, the resource utilisation questionnaire collected information on community-based medical resource usage [e.g. general practitioners (GPs), nurses, physiotherapists/occupational therapists, outpatients and medications]. Please refer to *Appendix 12* for a summary of protocol changes.

End points

Primary end point

Rate of intraoperative conversion to open surgery

Conversion to open surgery was defined as the use of a laparotomy wound for any part of the mesorectal dissection. The use of a small abdominal wound to facilitate a low, stapled anastomosis and/or specimen extraction was permissible and not considered as a conversion to open surgery. The decision to convert to an open operation was at the discretion of the operating surgeon. Details relating to the planned and actual operation were collected on the baseline and operative case report forms (CRFs).

Key secondary end points

Pathological circumferential resection margin positivity

Pathological circumferential resection margin positivity (CRM+) was defined as a distance of ≤ 1 mm of the cancer from the CRM as recorded on the local histopathology review.

Three-year local recurrence

Local recurrence was defined as evidence of locoregional disease within the surgical field. Time to local recurrence was calculated from the date of randomisation to the date of local recurrence, defined as the date of the relevant assessment (i.e. clinical, radiological and pathological) that first detected the local recurrence.

Further secondary end points

Intraoperative complications
Defined as adverse events occurring during surgery related to the surgical or related procedures (e.g. anaesthetic).

Thirty-day postoperative complications
Defined as an adverse event occurring during the first 30 days postoperatively and related to surgery or related procedures (e.g. anaesthetic).

Six-month postoperative complications (after 30 days)
Defined as an adverse event occurring during the first 6 months (after 30 days) postoperatively and related to surgery or related procedures (e.g. anaesthetic).

Thirty-day postoperative mortality
Defined as death from any cause within 30 days postoperatively.

Patient self-reported bladder function
Assessed by the patient self-reported I-PSS.

Patient self-reported sexual function
Assessed in males by the patient self-reported IIEF questionnaire and in females by the patient self-reported FSFI questionnaire.

Patient self-reported generic health
Assessed by the patient self-reported SF-36v2 questionnaire.

Patient self-reported fatigue
Assessed by the patient self-reported MFI-20 questionnaire.

Quality of the plane of surgery
Defined by the grading criteria using the local histological review. For an AR there was only one criterion: the quality of the mesorectum. For APR, the quality of the plane of surgery was assessed by the grade for the mesorectum and a second grade for the anorectal dissection below the levators. The quality of resection of the mesorectum was assessed as muscularis propria plane (worst), intramesorectal plane (intermediate) and mesorectal plane (best). The quality of surgery of the anorectum below the levators was assessed as intrasphincteric/submucosal plane (worst), sphincteric plane (intermediate) and levator plane (best).

Three-year disease-free survival
Disease-free survival time was defined as the time from date of randomisation to date of death from any cause, recurrent disease (locoregional or distant recurrence) or occurrence of a second primary cancer.

Three-year overall survival
Overall survival time was calculated from the date of randomisation to the date of death from any cause.

- Health economics evaluation (see *Health economic evaluation*).

Pathology central review

Local pathology data were used to carry out the analyses. A central blinded review of the local pathology data for all assessable patients was carried out. The agreement of local pathology and central pathology with respect to factors feeding into the analyses (e.g. T-staging) was assessed via summaries.

Sample size

Original sample size calculation and justification

The sample size calculation was based on ensuring that sufficient numbers of patients were recruited to address the primary end point of conversion to open rectal resection. A relative reduction of at least 50% (in absolute terms, 25% to 12.5% in the robotic-assisted laparoscopic group) was strongly believed to be achievable and also represented an extremely clinically important difference, not only in terms of outcomes for health-care providers but also in terms of patient-related outcomes, as it had been shown that patients who convert during surgery have worse outcomes. Therefore, using a conversion rate of 25% for standard laparoscopic surgery and a 50% relative reduction to be clinically relevant, with 80% power and a 5% (two-sided) significance level, 336 patients were required using a two-group continuity corrected chi-squared test of equal proportions (nQuery Advisor 6.01, Statistical Solutions, Saugus, MA, USA). Therefore, it was planned to recruit 400 patients (200 per group) to allow for early withdrawals, cross-over, protocol violations (e.g. benign tumours) and missing follow-up data.

Updated sample size

Recruitment to the original target sample size of 400 patients was completed 5 months earlier than planned and was under budget. Note that the original sample size of 400 patients aimed to achieve 80% power. Although this is conventionally considered to be sufficient, it is also commonly argued that 90% power is preferable. Given this, coupled with the fact that there was the opportunity to continue recruitment as a result of reaching the target of 400 patients early and under budget, we proposed to continue to recruit to the ROLARR trial until the date that was originally set to end recruitment. The aim of recruitment during this period was to recruit as many additional patients as possible to maximise power, up to a maximum of 520 patients (which, under the original sample size assumptions, would provide 90% power to detect a difference of at least 12.5% in conversion rates between the groups). This plan was endorsed by the EME programme, the DMEC and the TSC. This decision to continue recruitment was made before seeing any data or interim analyses. Consequently, a total of 471 patients had been randomised by the time the trial closed to recruitment. Under the original sample size assumptions, this provides around 86% power to detect a difference of at least 12.5% in conversion rates between the groups.

Randomisation

Randomisation took place as soon as possible after consent was obtained and after patients had completed their baseline patient-reported questionnaires (I-PSS, IIEF/FSFI, SF-36v2, MFI-20, EQ-5D). Randomisation took place as close to the date of surgery as possible. Surgeons were strongly encouraged to consent and randomise patients within 14 days of the planned surgery date whenever possible.

Following confirmation of written informed consent and eligibility, patients were randomised into the trial by authorised members of staff at the trial sites. Randomisation was performed centrally using the Clinical Trials Research Unit (CTRU) automated 24-hour telephone randomisation system. Authorisation codes and personal identification numbers (PINs), provided by the CTRU, were required to access the randomisation system.

Patients were randomised on a 1 : 1 basis to receive either robotic-assisted or standard laparoscopic rectal cancer surgery and were allocated a unique trial number. A computer-generated minimisation programme that incorporated a random element was used, with the following minimisation factors:

- treating surgeon
- patient sex (male or female)
- neo-adjuvant therapy (yes or no)
- nature of intended procedure (HAR, LAR or APR)

- body mass index (BMI) [calculated automatically from height (cm) and weight (kg) provided at randomisation and classified according to World Health Organization (WHO) criteria[29]]:
  - underweight/normal
  - overweight
  - obese class I
  - obese class II
  - obese class III.

Participating research sites were required to complete a log of all patients screened for eligibility who were not randomised either because they were ineligible or because they declined participation. Anonymised information was collected including:

- age
- sex
- date screened
- reason not eligible for trial participation
- eligible but declined and reason for this
- other reason for non-randomisation.

Blinding

As the two surgical procedures create incisions that can allow the patient to be blinded to the operative procedure performed, it arguably would have been scientifically preferable to blind patients to their surgical procedure, particularly in respect of patient-reported outcomes. However, it was anticipated that in practice maintaining the blind would have been extremely problematic (e.g. in countries such as the USA where private health-care insurance companies require disclosure of surgery details). Furthermore, it was anticipated that patients would also be seen by many health-care professionals throughout their time in the trial, increasing the risk that the blind may be broken. As a consequence, the trial design did not involve blinding patients to the operative procedure.

It should be noted that the trial end points are mainly objective measures and a central blinded assessment of these measures was included when possible (e.g. blinded central assessment of the quality of the plane of surgery).

Statistical methods

Unless otherwise stated, all analyses were prespecified and conducted on the intention-to-treat population (i.e. all randomised patients were categorised into treatment groups based on their randomisation, regardless of what treatment they subsequently received). All hypothesis tests were two-sided and conducted at the 5% level of significance. Estimates and their corresponding 95% confidence intervals (CIs) and *p*-values are presented for fixed effects. For the (random) surgeon effect, the intracluster correlation coefficient (ICC), estimated via the analysis-of-variance method, and bias-corrected bootstrapped 95% CIs are reported.

For most end points there was only a small number of missing data, such that a complete-case analysis was appropriate. For end points with non-negligible numbers of missing data, exploratory analyses were performed to consider the potential impact of the missing data. All models were fitted using SAS® version 9.4 (SAS Institute Inc., Cary, NC, USA).

All analyses, unless otherwise stated, adjusted for the minimisation factors only (see *Randomisation*). For each end point, sensitivity analyses to include adjustment for treating centre and country were considered; *Subgroup analyses* gives further details of this.

Primary end point: conversion to open surgery

The primary analysis was a complete-case analysis. Multilevel logistic regression was used to estimate the odds ratios (ORs) between treatment groups for conversion to open surgery, adjusting for all minimisation factors. All minimisation factors were included as fixed effects except intended operating surgeon, which was included as a random effect. A random intercept model was fitted first, then a model with both a random intercept and a random slope (i.e. random treatment effect) was fitted (hereafter referred to as the 'random slope' model). The need for the random slope term was assessed via consideration of a likelihood ratio test and the Akaike information criterion (AIC). The models were fitted using SAS version 9.4 glimmix procedure.

A number of prespecified sensitivity analyses were also performed.

Additional covariates

Fixity of tumour, whether or not the tumour was an obstructing tumour, T-stage or N-stage, whether or not the patient had abdominal surgery prior to their ROLARR operation and the level of scarring, whether or not adhesions were identified and whether or not there was a tumour perforation (non-iatrogenic) or abscess were all considered for inclusion in the model via examination of their effect on the model fit.

Actual operating surgeon

The primary analysis adjusted for the minimisation factors (i.e. the values of those factors that were used in the minimisation, regardless of whether or not those values were correct). In some cases, patients may have been allocated treatment under incorrect minimisation factors. In particular, their intended operating surgeon (used for minimisation) may not have been their actual operating surgeon. A sensitivity analysis was performed that incorporated actual operating surgeon rather than intended operating surgeon as a random effect in the model.

Learning effects

For each surgeon, the number of robotic-assisted and laparoscopic rectal operations relevant to the ROLARR trial performed by that surgeon was collected at regular intervals throughout the trial. From this, the number of ROLARR-relevant robotic-assisted and laparoscopic operations previously performed by the operating surgeon before each patient's operation was derived, assuming that the timings of all counted previous operations were uniformly distributed across the interval in which they occurred. These patient-level covariates ('number of previous robotic operations' and 'number of previous laparoscopic operations') were included in the multilevel model used in the primary analysis to explore potential associations between increased numbers of operations and patient outcomes. Interactions between the numbers of operations performed and the treatment effect were also considered.

Actual operation (post hoc)

The primary analysis adjusted for the minimisation factors (i.e. the values of those factors that were used in the minimisation, regardless of whether or not those values were correct). In some cases, patients may have been allocated treatment under incorrect minimisation factors. In particular, their intended procedure (used for minimisation) may not have been the actual procedure that they received. A sensitivity analysis was performed that incorporated actual procedure rather than intended procedure as a fixed effect in the model.

Key secondary end points

Circumferential resection margin positivity

The analysis of CRM+ was a complete-case analysis. Multilevel logistic regression was used to estimate the ORs between treatment groups for CRM+, adjusting for all minimisation factors. All minimisation factors were included as fixed effects except intended operating surgeon, which was included as a random effect. A random intercept model was fitted first, then a random slope model was fitted and the need for the random slope term was assessed via consideration of a likelihood ratio test and the AIC. The models were fitted using SAS version 9.4 glimmix procedure.

A number of prespecified sensitivity analyses were also performed.

Additional covariates

Fixity of tumour, T-stage and N-stage (post neo-adjuvant therapy) and whether or not there was a tumour perforation (non-iatrogenic)/abscess were all considered for inclusion in the model via examination of their effect on the model fit.

Three-year local recurrence

All patient follow-up, including follow-up beyond 3 years post randomisation, was incorporated into the analysis of local recurrence. Time to local recurrence was defined as the time from randomisation to the date of the relevant assessment (i.e. clinical, radiological and pathological) that first detected the local recurrence.

Differences in time to local recurrence between the treatment groups were estimated using a shared frailty model (Cox proportional hazards regression with mixed effects), including intended operating surgeon as a random effect. The models were fitted using SAS version 9.4 phreg procedure. The 3-year local recurrence rate was estimated using cumulative incidence functions for time to local recurrence, treating death as a competing risk.

Patients who were alive and without any local recurrence at the time of analysis were censored at the time they were last known to be alive and local recurrence free. If patients were lost to follow-up, they were also censored at the time they were last known to be alive and local recurrence free. Patients who died without any local recurrence were censored at date of death in analyses estimating treatment effects, but were classed as having a competing risk event in the analysis estimating incidence of local recurrence (as calculated using cumulative incidence functions) to avoid overestimation of cumulative incidence.[30] In certain non-standard circumstances (prespecified in the statistical analysis plan), patients were censored at time 0. Patients with non-standard circumstances who were censored at time 0 are summarised, and reasons are given in the results (see *Chapter 3, Disease-free survival*).

Further secondary end points

The analyses of further binary secondary end points – intraoperative complications, postoperative complications within 30 days, after 30 days and within 6 months, and quality of the plane of surgery (i.e. mesorectal plane Yes/No) – were complete-case analyses. Multilevel logistic regression was used to estimate the ORs between treatment groups for each end point, adjusting for all minimisation factors. All minimisation factors were included as fixed effects except intended operating surgeon, which was included as a random effect via a random intercept term. The models were fitted using SAS version 9.4 glimmix procedure.

For further continuous secondary end points [bladder function (I-PSS), sexual function in males (IIEF) and in females (FSFI), generic health-related QoL (SF-36v2) and fatigue (MFI-20)], multilevel generalised linear models were used to estimate the mean difference between treatment groups, adjusting for all minimisation factors and the baseline score. All minimisation factors were included as fixed effects except intended operating surgeon, which was included as a random effect via a random intercept term. The I-PSS, IIEF and FSFI were modelled using a two-level model: patients nested within surgeon. The SF-36v2 questionnaire and

MFI-20 were modelled using a three-level model: repeated assessments within patient within surgeon. The models were fitted using the SAS version 9.4 glimmix procedure.

All patient follow-up, including follow-up beyond 3 years post randomisation, was incorporated into the analysis of DFS and OS. For DFS and OS, shared frailty models were used to estimate the hazard ratios (HRs) between treatment groups, adjusting for all minimisation factors. All minimisation factors were included as fixed effects except intended operating surgeon, which was included as a random effect via a random intercept term. In certain non-standard circumstances (prespecified in the statistical analysis plan), patients were censored at time 0. Patients with non-standard circumstances who were censored at time 0 are summarised and reasons are given in the results (see *Chapter 3, Disease-free survival*). The models were fitted using SAS version 9.4 phreg procedure.

Subgroup analyses

Subgroup analyses relating to the primary end point across sex, BMI class and procedure received, as well as relating to CRM+ across sex, BMI class and T-stage, were performed. Subgroup analyses relating to each of local recurrence, DFS and OS across type of operation, T-stage and neo-adjuvant therapy were also performed. All subgroup analyses tested heterogeneity of the treatment effect across the subgroups and also estimated the treatment effect within each subgroup, via the addition of an appropriate interaction term to the primary analysis model.

Model diagnostics

Multilevel logistic regression models

Model fit was assessed by examining the raw residuals on the probability scale outputted from SAS version 9.4 glimmix procedure, for example the residual for patient *i*:

$$r_i = Y_i - \hat{p}_i, \tag{1}$$

in which:

$$Y_i = \begin{cases} 1, & \text{Patient had the event (e.g. was converted to open surgery)} \\ 0, & \text{Otherwise} \end{cases}, \tag{2}$$

and $\hat{p}_i$ is the predicted probability of the event (e.g. conversion to open surgery) for patient *i* [including empirical Bayes' estimate (EBE) of the random effect]. Index plots (plots of the raw residuals vs. patient identification) were used to identify potential outliers. Empirical probability plots were also used to assess model fit and identify potential outliers. These plots plotted the observed, ordered Pearson residuals (outputted from SAS version 9.4 glimmix procedure) against expected, ordered Pearson residuals under the model assumptions – analogous to a normal Q–Q plot for normal-errors regression. The expected sampling distributions of Pearson residuals were determined empirically via simulations. Specifically, for each simulation each patient's outcome was randomly drawn from a Bernoulli (p) distribution with:

$$p = \hat{p}_i, \tag{3}$$

for patient *i*.

The model was refitted to this simulated data set and the Pearson residuals recorded.

This was repeated 100 times to yield an empirical sampling distribution of Pearson residuals for each patient. In the empirical probability plot, the actual observed Pearson residuals for each patient were plotted against the median, 2.5th percentile and 97.5th percentile of the empirical sampling distribution. Observations were considered to be potential outliers if they lay outside the 2.5th percentile to 97.5th percentile range [analogous to considering Pearson residuals lying outside the interval (–2,2) to be potential outliers in a normal-errors regression].

Overly influential observations on the treatment effect regression coefficient were identified via the calculation of exponentiated delta-betas. The exponentiated delta-beta was calculated for each patient; for example, for patient *i*, the exponentiated delta-beta for the treatment effect regression coefficient was:

$$\exp(\beta_1^{(i)} - \beta_1) = \frac{\exp(\beta_1^{(i)})}{\exp(\beta_1)}, \quad (4)$$

in which β_1 is the regression coefficient for the treatment effect in the full model and $\beta_1^{(i)}$ is the treatment effect regression coefficient in the model where patient *i* has been omitted. Note that this is the ratio of the estimated ORs from the two models; for example, an exponentiated delta-beta for the treatment effect for patient *i* of 1.05 would imply that the inclusion of patient *i* increases the treatment effect OR estimate by 5% compared with the omission of patient *i*. The exponentiated delta-betas were plotted against patient identifier (ID) in order to visually identify highly influential observations.

Shared frailty models

Models were refitted as Cox proportional hazards models with robust standard errors, without the random effect for operating surgeon. This gave the same point estimates as the shared frailty model, and broadly similar standard errors. Deviance residuals were used to identify any potential outliers. The proportional hazards (PH) assumption for the treatment effect was assessed via a plot of the standardised Schoenfeld residuals over time, as well as a plot of the observed standardised score process versus simulated standardised score processes under PH. The PH assumption was also tested via the Supremum test. Exponentiated delta-betas (as described in *Multilevel logistic regression models*) were used to identify overly influential observations.

Health economic evaluation

An economic evaluation was performed using a UK NHS perspective to aid the development of an evidence base to support NHS service providers and budget holders in their decision-making processes. Costs associated with robotic surgery excluded acquisition and maintenance costs. The evaluation estimated the expected incremental cost-effectiveness of robotic resection compared with laparoscopic resection at 6 months. It was planned that this would be extrapolated using a decision-analytic model to estimate lifetime cost-effectiveness.

The ROLARR trial collected information on the nature of all initial resection operations using trial CRFs. This included information on the type of operation and resources used within this operation, including instrumentation and times for operation theatres and staff. CRF data also captured information on postoperative and distal complications on all trial patients.

However, many types of resource utilisation were not collected for all patients in the ROLARR trial. In particular, given the challenges of conducting research within global trials, data on resource utilisation after the initial operation were collected only on patients from the UK and the USA. As the adjuvant chemotherapy is likely to both vary widely and be expensive, there is a danger that any small differences at this stage will be both unrelated to the surgery received and, given the cost of chemotherapy, outweigh any cost differences that are related to the intended type of surgery. For this reason, the cost data do not attempt to consider the chemotherapies received or antinausea drugs attached to these chemotherapies.

For patients in the UK and USA, information was collected alongside the trial-related questionnaires at approximately 30 days and at 6 months. It was expected that data collection might be poor and as a result it would be necessary to impute data for a substantial number of these patients.

Data were collected on both primary care and secondary care, including contact with GPs and primary care physicians (including the location of any contacts), nurse contacts (including at primary care, district/at home nursing, and stoma nursing) and outpatient visits.

The analysis considered costs in GBP with 2015 as the base year, from an NHS-payer perspective. Given the focus of this perspective, when clinical practice appeared to differ in the USA, particularly around pain medication, the approach taken was the one that appeared to apply in the UK. This perspective also means that unit costs are those costs that apply within the NHS.

Costing individual resource utilisation

Surgery

Surgical costs were computed by first identifying an overall global average for resection surgeries, for which a weighted sum of non-elective complex large intestine operative costs was used. Individualised surgical costs allowed for both excess bed-days (at £326.11 per day) and differences in the time and staffing within the theatre. For the laparoscopic group, costs were calculated based on data provided on the operative team and time, and it was assumed that on average this group would cost the same as the baseline surgical cost. Therefore, those patients receiving a laparoscopic operation might have had a cheaper or more extensive operation than 'average' but would be similar overall. Given that robotic surgeries were expected to require longer use of the operating theatre (e.g. including greater setup time), incorporating staffing/time costs modifies the resection surgery costs to reflect this. The instrumentation costs were included separately for those items that, in the opinion of the chief investigator, would not necessarily be considered an automatic inclusion within the operating theatre. (So, for instance, although data were collected on suction, this is not costed.)

As staff costs appear below (*Table 1*), and instrument costs are assessed, it is not appropriate to include these. Excluding these costs, the use of theatres costs £339 per hour.

TABLE 1 Surgical unit costs

Surgical costs	Unit	Cost (£)	Source
Baseline resection surgery: operative cost	Per operation	8307.78	[a]*NHS Reference Costs 2014 to 2015*[31]
Baseline resection surgery: operative cost – excess bed-days	Per excess bed-day[b]	326.11	[a]*NHS Reference Costs 2014 to 2015*[31]
Other surgeries for complications	Per operation	8307.78	[c]*NHS Reference Costs 2014 to 2015*[31]
Surgeon	Per hour	138.00	Personal Social Services Research Unit[32]
Anaesthetist	Per hour	60.74	Personal Social Services Research Unit[32]
First surgical assistant (band 7)	Per hour	60.00	Personal Social Services Research Unit[32] (including qualifications)
Subsequent surgical assistants (band 5)	Per hour	43.00	Personal Social Services Research Unit[32] (including qualifications)
Operating theatre (no staff or specialist instruments)	Per hour	339.00	[d]Derived from Information Services Division, NHS National Services Scotland[33]

a Assumed to reflect weighted average of *NHS Reference Costs 2014 to 2015*,[31] as a weighted sum by subcategories of FZ74 complex large intestines. Non-elective only.
b Applies per diem to stays > 44 days (as weighted boundary).
c Assumed to reflect weighted average of *NHS Reference Costs 2014 to 2015*,[31] as a weighted sum by subcategories of FZ74 complex large intestines. Includes both elective and non-elective.
d All costs (medical, nursing, other staff, drugs, central sterile services department, other supplies) come out at £1172 per theatre hour.

Instrumentation costs were identified in discussion with the chief investigator who provided data on which of the instruments to specifically include (and which were more or less trivial given operating theatre costs) and unit costs for each instrument (*Table 2*).

An overall in-theatre cost was calculated for each surgery (where data were complete on the fields above) and calculated, including the theatre, staff and instrumental costs. The mean of these costs among the group who received laparoscopic surgery (as opposed to allocated) was identified, and this was subtracted from all individual cost figures to indicate how costs would be likely to differ from a typical operation. As such, the average laparoscopic difference is zero, although the robotic difference could be positive (if more expensive) or negative (if less expensive). The difference in the cost figure was then added back onto the reference cost to give an estimated cost for each individual surgery.

It should be noted that the analysis presented here does not include the cost of the surgical robot (when applicable), in order to provide an optimistic case figure for the cost-effectiveness of robotic surgery.

Other inpatient visits

The main inpatient costs assessed after the initial surgery were for stoma reversal operations and for other related colorectal surgeries identified from the CRFs. When other major related surgery was indicated, this was coded as an average of complex large intestine surgeries at £7621.24.[31]

Following approaches used elsewhere, stoma reversals are coded as elective intermediate procedures (FZ50Z, intermediate large intestine procedures, aged ≥ 19 years) at £1691.06 per case.[31]

The unit costs of all other inpatient visits are taken from *NHS Reference Costs 2014 to 2015*[31] and are shown in *Table 3*.

TABLE 2 Instrument unit costs

Robotic	Unit	Cost (£)
Aspirator	Each	150
Bipolar forceps	Each	150
Vessel sealer	Each	500
Graspers	Each	150
Haemolock	Initial	150
	Per clip	30
Hook	Each	150
Needle driver	Each	150
Scissors	Each	150
Stapler	Per firing	150
Laparoscopic and open		
Disposable Babcock	Each	150
Graspers	Each	100
Stapler (including open staplers)	Initial	300
	Per reload	80
Scissors	Each	100
Vascular clips	Each	80
Wound protector	Each	50
Wound retractor	Each	50
Vessel sealers (e.g. Ligasure)	Each	500

TABLE 3 Other inpatient procedure unit costs

Procedure	Code	Unit costs[a] (£)
Non-reversal stoma operations	FZ50Z. Elective inpatient	1691.06
Transient ischaemic attack	AA29F. Transient ischaemic attack with a CC score of 0–4 (Non-elective)	1252.81
Deep-vein thrombosis	YQ51E. Deep-vein thrombosis with a CC score of 0–2	1361.97
Pulmonary embolism	DZ09 K. Pulmonary embolus with interventions, with a CC score of 0–8	3509.12
Renal failure	Acute kidney injury without interventions, with a CC score of 0–3	1785.63
	LA07 K. Acute kidney injury with interventions, with a CC score of 0–5	3784.72
Abdominal infections, anastomotic leak	FZ36L Gastrointestinal infections with single intervention, with a CC score of 0–1	3610.31
Inpatient urinary tract infections	LA04S. Kidney or urinary tract infections, without interventions, with a CC score of 0–1	1502.55
Haemorrhage	FZ38P. Gastrointestinal bleed without interventions, with a CC score of 0–4	1370.09
Cardiac events	EB12C. Unspecified chest pain with a CC score of 0–4	1088.68
Protracted ileus	FZ13C. Minor therapeutic or diagnostic, general abdominal procedures, ≥ 19 years. Non-elective	3471.40
Urinary retention	LA09Q. General renal disorders without interventions, with a CC score of 0–2	1399.13
Gastrointestinal obstruction	FZ27G. Intermediate therapeutic general abdominal procedures, ≥ 19 years and over, with a CC score of 0	3335.27
High stoma output (not coded as serious though)	FZ50Z Intermediate large intestine procedures, ≥ 19 years	1836.19
Respiratory inpatient (non-infection)	DZ19 N. Other respiratory disorders without interventions, with a CC score of 0–4	1163.67
Not specified	Average of all elective inpatients, all sources	3573.02

CC, complication and comorbidity.
a All taken from *NHS Reference Costs 2014 to 2015*.[31]

Primary care

When possible, unit costs for GPs were taken from the Personal Social Services Research Unit (PSSRU) 2015[32] using qualification costs and direct care staff costs, with similar figures (without direct care staff costs) used for nurse-based visits. For primary care, GP costs assumed a mean duration of surgery visits of 11.7 minutes (£44) and of 11.4 minutes for home visits (plus 10 minutes of travel) (£81), and a cost of £27 for telephone consultations. Nurse costs were assessed based on location, assuming 15.5 minutes of contact for each visit, with an additional 10 minutes of travel for visits away from surgeries (surgery £14.47, other £21.25).

Outpatient and other health professional visits

Outpatient visits costed using unit costs (consultant-led outpatient attendances) from *NHS Reference Costs 2014 to 15*,[31] after grouping most visits into colorectal surgery, gastroenterology, medical oncology, medical ophthalmology, trauma and orthopaedics, nephrology, urology and others. Accident and emergency visits were costed using the overall average of all emergency medical attendances with the *NHS Reference Costs 2014 to 2015*.[31]

Contacts for many of the remaining health professionals were taken from the PSSRU costs, including occupational therapy (PSSRU 2015[32]) and counselling (PSSRU 2014[34]), and chiropody/podiatry (PSSRU 2010[35]), with costs inflated to 2015 figures using either the HCHS (Health and Community Health Service) price index for health or mid-point changes in *Agenda for Change* pay bands.[36]

Stoma costs

The costs of ongoing stoma were calculated from Jones,[37] who reported on figures from the Cwm Taf Health Board (NHS Wales) (*Table 4*). This assumes a monthly cost of £84 per 30 colostomy bags and £94 per 30 ileostomy bags. These 2011/12 figures were inflated to 2014/15 figures using the HCHS price index.[32] It should be noted that this is likely to underestimate the true costs of stoma, as this does not include items such as wipes, although there do not appear to be clear, published figures available that include such items.

Medication costs

Medication costs were found by coding responses from UK/US patients on a line-by-line basis initially and then across all responses for specified pharmaceuticals (*Table 5*). When possible, individual statements about frequency and duration of treatment were used to inform assumptions about utilisation. In cases where no statements were made to identify utilisation, the *British National Formulary*[38] was used to identify an indicative strength/dosage. Unit costs are taken from eMIT (the drugs and pharmaceutical electronic market information tool),[39] when possible, or the *NHS Indicative Drug Tariff*[40] figures, when not.

Medications for unrelated events (e.g. shingles, thyroid conditions, chronic obstructive pulmonary disease, glaucoma) were ignored. Information on dietary supplementation (e.g. vitamins) was provided infrequently but was not included.

For some pain medications, differences in clinical practice mean that medications have been recoded. For example, although hydrocodone is widely used in the USA, it does not appear within the (UK) *British National Formulary*;[38] codeine phosphate is used in preference to hydrocodone. Codeine phosphate is also used when several other medications (i.e. Norco®, codeine sulphate) are indicated.

Given the cost of chemotherapies and the relatively sparse information collected on these (and the danger that the costs involved mask all useful information), these have been ignored in the range of medications being considered. Furthermore, most of the stoma-related costs are removed, as the stoma unit costs specified in *Table 4* include a range of costs that may overlap. Chemotherapies as adjunct therapies and anti-emetic/anti-nausea drugs (including anti-psychotics) are also ignored, since the cost of these drugs risks swamping any useful information provided.

TABLE 4 Stoma unit costs

Stoma costs	Unit	Unit cost	Source
Colostomy costs @ two bags per day (2011/12)	Per 30 days	£84.00	Cwm Taf Health Board (Jones, 2015[37])
Ileostomy costs @ one bag per day (2011/12)	Per 30 days	£94.00	Cwm Taf Health Board (Jones, 2015[37])
Inflation between 2011/12 and 2014/15 using HCHS (293.1 vs. 282.5)		3.75%	
Stoma reversals	Per operation	£1691.06	[a]*NHS Reference Costs 2014 to 2015*[31]
Colostomy costs @ two bags per day (2011/12)	Per 30 days	£84.00	Cwm Taf Health Board (Jones, 2015[37])
Ileostomy costs @ one bag per day (2011/12)	Per 30 days	£94.00	Cwm Taf Health Board (Jones, 2015[37])
Inflation between 2011/12 and 2014/15 using HCHS (293.1 vs. 282.5)		3.75%	
Stoma reversals	Per operation	£1691.06	[a]*NHS Reference Costs 2014 to 2015*[31]

a Stoma reversals (intermediate procedures). FZ50Z Intermediate Large Intestine Procedures, 19 years and over. Elective inpatient.

TABLE 5 Medication unit costs

Symptom	Drug	Unit	Cost (£)
Cardiac/statins	Simvastatin	Per 28 units	0.16
Cardiac/statins	Atorvastatin	Per 28 units	0.49
Cardiac/statins	Rosuvastatin	Per 28 units	18.03
Cardiac/statins	Doxazosin	Per 28 units	0.19
Cardiac/statins	Candesartan	Per 28 units	0.55
Cardiac/statins	Losartan	Per 28 units	0.30
Cardiac/statins	Olmesartan (Benicar®)	Per 28 units	10.95
Cardiac/statins	Lisinopril	Per 28 units	0.29
Cardiac/statins	Perindopril	Per 30 units	0.61
Cardiac/statins	Ramipril	Per 28 units	0.27
Cardiac/statins	Bisprolol	Per 28 units	0.25
Cardiac/statins	Metaprolol (plus unspecified beta blocker)	Per 28 units	0.55
Cardiac/statins	Carvedilol	Per 28 units	0.59
Cardiac/statins	Propranolol	Per 56 units	1.67
Cardiac/statins	Atenolol	Per 28 units	0.18
Cardiac/statins	Amlodipine	Per 28 units	0.16
Cardiac/statins	Feoldipine	Per 28 units	0.55
Cardiac/statins	Lercanidipine	Per 28 units	1.42
Cardiac/statins	Nifedipine	Per 56 units	21.00
Cardiac/statins	Cartia	Per 56 units	41.87
Cardiac/statins	Furosemide	Per 28 units	0.13
Cardiac/statins	Indapamide	Per 28 units	1.02
Cardiac/statins	Esomeprazole	Daily	2.22
Cardiac/statins	Lansoprazole	Per 28 units	0.98
Cardiac/statins	Glytrin	Daily	1.13
Cardiac/statins	Amiodarone	Once	13.17
Pain	Tramadol[a]	Per 14 units	1.67
Pain	Paracetamol[b]	Per 16 units	0.13
Pain	Ibuprofen	Per 16 units	0.17
Pain	Codeine phosphate[c]	Per 28 units	0.37
Pain	Gabapentin	Per 100 units	1.30
Pain	Cocodamol	Per 30 units	0.65
Pain	Codrydramol	Per 30 units	0.47
Pain	Oramorph	Single use	1.89
Pain	Oxycodone	Per 56 units	6.06
Pain	Indomethacin	Single pack	0.55
Pain	Meloxicam	Single pack	0.43

continued

TABLE 5 Medication unit costs (*continued*)

Symptom	Drug	Unit	Cost (£)
Pain	Solaraze	Single pack	0.67
Pain	Naproxen	Per 28 units	0.70
Anticoagulants	Aspirin	Per 28 units	0.14
Anticoagulants	Clopidogrel	Per 30 units	4.58
Anticoagulants	Heparin	Per 10 units	16.62
Anticoagulants	Delteparin	Per 10 units	51.22
Anticoagulants	Warfarin	Per 28 units	0.25
Anticoagulants	Tinzaparin	Per 10 vials	105.66
Anticoagulants	Enoxaparin	Per 10 units	20.86
Anticoagulants	Rivaroxaban	Per 28 units	50.40
Antibiotics and immunological	Amoxicillin (plus unspecified antibiotics)	Per 21 units	0.45
Antibiotics and immunological	Trimethoprim	Per 14 units	0.89
Antibiotics and immunological	Nitrofurantoin	Per 28 units	3.57
Antibiotics and immunological	Cephalexin	Per 28 tablets	0.73
Antibiotics and immunological	Lexofloxacin	Per 5 tablets	0.92
Antibiotics and immunological	Bactrim	Per 28 units	3.03
Antibiotics and immunological	Oxytetracycline	Per 28 units	0.43
Antibiotics and immunological	Vancomycin	Per 28 units	32.90
Antibiotics and immunological	Ciprofloxacin	Per 20 units	0.42
Antibiotics and immunological	Metronidazole	Per 21 tablets	0.39
Antibiotics and immunological	Fluconazole	Per unit	0.22
Stool thickeners	Loperamide	Per 30 units	1.61
Stool thickeners	Atropine diphenoxylate	Per 100 units	10.74
Stool softeners/laxatives	Domperidone	Per 100 units	0.91
Stool softeners/laxatives	Lactulose (laxative if not clearly stated)	Per bottle	1.21
Stool softeners/laxatives	Docusate	Per 30 units	2.09
Stool softeners/laxatives	Metamucil	Per 10 units	4.22
Stool softeners/laxatives	Movicol/Laxido	Per 30 units	2.99
Stool softeners/laxatives	Clorphenamine	Per 28 units	0.84
Stool softeners/laxatives	Fluticasone	Per bottle	4.17
Other stomach/digestive	Mebeverine	Per unit	4.68
Other stomach/digestive	Ranitidine	Per 12 units	0.30
Other stomach/digestive	Omeprazole	Per 28 units	0.44
Urinary	Bendoflumethiazide	Per 28 units	0.11
Urinary	Tamsulosin	Per 30 units	0.73
Urinary	Solifenacin succinate	Per 30 units	27.62
Antidepressants/anti-anxiety	Alprazolam	Per 60 units	3.18
Antidepressants/anti-anxiety	Citalopram	Per 28 units	0.18

TABLE 5 Medication unit costs (*continued*)

Symptom	Drug	Unit	Cost (£)
Antidepressants/anti-anxiety	Diazepam	Per 28 units	0.23
Antidepressants/anti-anxiety	Lorazepam	Per 28 units	1.19
Antidepressants/anti-anxiety	Sertraline	Per 28 units	0.48
Antidepressants/anti-anxiety	Duloxetine	Per 28 units	22.40
Antidepressants/anti-anxiety	Amitripyline	Per 28 units	0.14
Sexual dysfunction	Tadalafil	Per 28 units	54.99
Sexual dysfunction	Sildenafil	Per 28 units	0.92
Sleeping pills	Zopiclone	Per 28 units	0.41

a Tramadol use varied substantially. Use costed as provided but when detail was lacking a 'default' case of 4 weeks/28 tablets was used.
b Short courses assumed to be six tablets per day. Other non-specified courses take an average length of 72 days.
c Where codeine sulphate, Norco and hydrocodone are indicated (USA), these replace codeine phosphate (as per UK). Codeine sulphate sees an average course of 101 pills taken in the subgroup when exact numbers are not specified, and 30 when stated as 'Norco' but not specified (as this reflects the utilisation in this subgroup).

Quality of life

The outcome measure for the economic evaluation was the quality-adjusted life-year (QALY). Health-related quality of life (HRQoL) was measured using the EQ-5D and valued using the standard UK tariff.[41] The EQ-5D data were obtained using English-language version questionnaires completed by patients recruited from UK and North American trial sites. The data were collected alongside resource data at baseline and at 30 days and at 6 months postoperatively. Multiple imputation methods were used to estimate HRQoL for those patients not completing this questionnaire. In this way, the analysis includes HRQoL for all patients in the trial, regardless of language.

Quality-of-life estimates were constructed using the EuroQol-5 Dimensions, three-level version (EQ-5D-3L), responses, valued using the standard UK tariff. Responses were in most cases valued as an area under the curve. The exception to this was when the data indicated that a stoma reversal operation occurred between 30 days and 6 months postoperatively; in this case, the 30-day figure was assumed to apply between 30 days and the reversal operation, and the 6-month figure was assumed to apply from the reversal date through to the end of follow-up.

As a sensitivity analysis (which has been run but is additional to those analyses displayed here), values were also inferred from the SF-36v2 data obtained from all patients in the ROLARR trial. Health-related quality-of-life figures were obtained as Short Form questionnaire-6 Dimensions (SF-6D) utilities by applying the algorithm developed by Brazier *et al*.[42]

Imputing costs and quality of life

Given that not all UK and US patients answered the medical resource utilisation questionnaire, imputation was necessary within the trial. Values were multiply imputed by category of variable and timing, using 100 imputations for each incomplete observation. By using a large number of multiple imputations, we aimed to more accurately reflect uncertainty.

The first set of variables imputed as chained equations related to the original inpatient admission, being the duration of surgery, the number of assistants and length of stay. These figures were imputed based on the procedure reviewed, whether or not this was a low anterior operation, whether or not there was evidence of locoregional spread, whether or not there had been any CT staging and using age/sex as

demographics. Once these figures for the incomplete variables were imputed (and so non-missing), then they were used to inform subsequent variables. These figures also allowed the calculation of a (complete) series of figures for operative costs.

Figures were then imputed to find both the number of other surgeries required and the number of days a stoma would be in place and the type of stoma.

Quality-of-life observations were imputed next, with the EQ-5D-3L and SF-6D data as chained equations using age/sex and information about baseline health conditions. For time periods after baseline, the previous observation for both the EQ-5D and the SF-6D were also used as predictors. With data now complete on these observations, and for the number of stoma days, QALYs could then be calculated.

Other costs were imputed based on the EQ-5D-3L utilities and observed complications, as represented by dummy variables representing common ($n > 10$) complications graded on the Clavien–Dindo classification to ≥ 3. As these equations were lengthy, each equation was examined and terms removed where $p > 0.200$. For all health professional and outpatient visits, the equations estimated the number of events. In order to turn these into costs, these utilisation figures were multiplied by the average cost of events observed within the data. There were no clear significant predictors of medication costs within the data. Medication costs were imputed by predicting first whether or not any medication costs existed for that patient, within the relevant period (i.e. post discharge within the first 30 days, between 30 days and 6 months), and then as random variables reflecting those patients with data in that period.

Analysis

Data for all cost items were combined together to form an estimate of total costs, alongside the estimated total number of QALYs within the first 6 months. Because this covers only a 6-month time period, the maximum number of QALYs that could be observed is 0.500 QALYs (or, more properly, 0.499, given that 182 days are used).

With these figures, total costs and costs within different cost categories are presented in terms of both tables and probabilistic sensitivity analyses. As missing data are multiply imputed, it is efficient to conduct the probabilistic sensitivity analysis by bootstrapping (sampling with replacement) among the relevant data set, selecting from the imputed data set until the number of patients in the initial sample has been reached. For example, if a laparoscopic group had 75 patients within a particular scenario, the multiply-imputed data set would have $75 \times 100 = 7500$ observations, and the procedure would choose one of these 7500 observations, 75 times, in order to obtain a resampled estimate. The results in terms of total costs and total QALYs for each group are then compared and assessed to identify the most cost-effective option. Repeating this resampling procedure 10,000 times provides an estimate of the distribution of incremental costs and incremental benefits between the two options under consideration. This also allows the calculation of cost-effectiveness acceptability curves, which display the probability of each of the options under consideration being cost-effective at different values of the cost-effectiveness threshold. By convention, the values of £20,000 and £30,000 per QALY are focused on,[43] although the evidence is that the true cost-effectiveness threshold may be lower than these figures.[44]

The primary analysis used imputed data for UK and US patients ($n = 190$). Secondary analyses were undertaken using the following:

- complete data for all patients ($n = 97$)
- imputed data for UK and US patients intended to receive low anterior surgery ($n = 135$)
- imputed data for all observations ($n = 471$).

Chapter 3 Results

Recruitment

Between 7 January 2011 and 30 September 2014, 1276 patients were assessed for eligibility by 40 surgeons from 26 sites across 10 countries (i.e. UK, Italy, Denmark, USA, Finland, South Korea, Germany, France, Australia and Singapore). In total, 471 (36.9%) of these patients were randomised: 234 to laparoscopic and 237 to robotic surgery (*Figure 1*). Four patients withdrew from data collection before their operation and

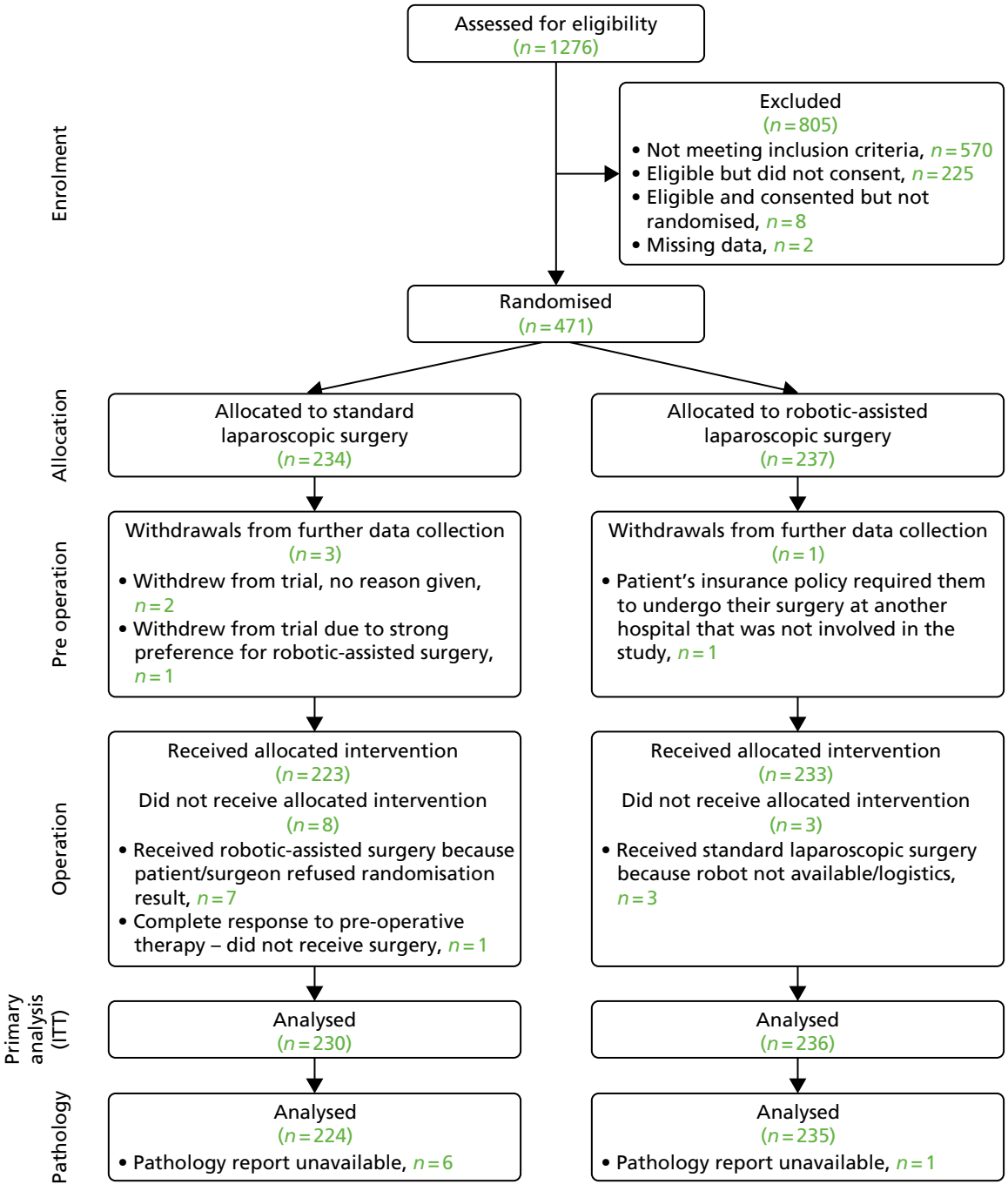


FIGURE 1 The CONSORT flow diagram. ITT, intention to treat.

one patient did not undergo surgery because of a complete clinical response to neo-adjuvant therapy. The remaining 466 patients underwent an operation, with 456 (97.9%) undergoing the allocated treatment.

Baseline data

All minimisation factors except treating surgeon are summarised by treatment group across all randomised patients in *Table 6*. The minimisation factor intended operating surgeon is summarised by treatment group for all randomised patients in *Table 7*. Summaries of selected additional baseline fields are given in *Table 8*.

Operative and pathology summaries

Crude summaries of operative and local pathology data fields are given in *Table 9*. The primary end point, conversion to open surgery, is summarised here but is considered in more detail in *Primary end point: conversion to open surgery*. The key secondary end point, CRM+, is also summarised here but is considered in more detail in *Key secondary end point: circumferential resection margin positivity (CRM+)*.

Table 10 presents the different pathways of intraoperative conversions between robotic, laparoscopic and open surgery. For example, the pathway 'Laparoscopic → Robotic' indicates that the patient's operation began as a standard laparoscopic operation, but was converted to a robotic operation intraoperatively. A pathway with only one type of operation indicates no conversions, for example 'Laparoscopic' indicates that the patient's operation began as a standard laparoscopic operation and was completed without conversion to robotic or open surgery.

TABLE 6 Minimisation factors by treatment group

Minimisation factor	Treatment group, *n* (%)		Total, *n* (%) (*N* = 471)
	Standard laparoscopic surgery (*N* = 234)	Robotic-assisted laparoscopic surgery (*N* = 237)	
Sex			
Male	159 (67.9)	161 (67.9)	320 (67.9)
Female	75 (32.1)	76 (32.1)	151 (32.1)
BMI classification			
Underweight/normal	87 (37.2)	93 (39.2)	180 (38.2)
Overweight	92 (39.3)	90 (38.0)	182 (38.6)
Obese class I	38 (16.2)	41 (17.3)	79 (16.8)
Obese class II	10 (4.3)	9 (3.8)	19 (4.0)
Obese class III	7 (3.0)	4 (1.7)	11 (2.3)
Neo-adjuvant therapy			
Yes	103 (44.0)	109 (46.0)	212 (45.0)
No	131 (56.0)	128 (54.0)	259 (55.0)
Intended procedure			
HAR	34 (14.5)	35 (14.8)	69 (14.6)
LAR	158 (67.5)	159 (67.1)	317 (67.3)
APR	42 (17.9)	43 (18.1)	85 (18.0)

TABLE 7 Recruitment by surgeon

Surgeon ID	Treatment group, *n* (%)		Total, *n* (%) (*N* = 471)
	Standard laparoscopic surgery (*N* = 234)	Robotic-assisted laparoscopic surgery (*N* = 237)	
1	20 (47.6)	22 (52.4)	42 (8.9)
2	20 (57.1)	15 (42.9)	35 (7.4)
3	17 (51.5)	16 (48.5)	33 (7.0)
4	17 (51.5)	16 (48.5)	33 (7.0)
5	11 (40.7)	16 (59.3)	27 (5.7)
6	12 (46.2)	14 (53.8)	26 (5.5)
7	13 (52.0)	12 (48.0)	25 (5.3)
8	11 (52.4)	10 (47.6)	21 (4.5)
9	9 (50.0)	9 (50.0)	18 (3.8)
10	10 (55.6)	8 (44.4)	18 (3.8)
11	9 (50.0)	9 (50.0)	18 (3.8)
12	9 (50.0)	9 (50.0)	18 (3.8)
13	7 (43.8)	9 (56.3)	16 (3.4)
14	6 (46.2)	7 (53.8)	13 (2.8)
15	6 (46.2)	7 (53.8)	13 (2.8)
16	5 (45.5)	6 (54.5)	11 (2.3)
17	3 (30.0)	7 (70.0)	10 (2.1)
18	5 (50.0)	5 (50.0)	10 (2.1)
19	3 (33.3)	6 (66.7)	9 (1.9)
20	4 (44.4)	5 (55.6)	9 (1.9)
21	3 (42.9)	4 (57.1)	7 (1.5)
22	5 (83.3)	1 (16.7)	6 (1.3)
23	1 (20.0)	4 (80.0)	5 (1.1)
24	2 (40.0)	3 (60.0)	5 (1.1)
25	3 (60.0)	2 (40.0)	5 (1.1)
26	3 (75.0)	1 (25.0)	4 (0.8)
27	3 (75.0)	1 (25.0)	4 (0.8)
28	0 (0.0)	3 (100.0)	3 (0.6)
29	2 (66.7)	1 (33.3)	3 (0.6)
30	2 (66.7)	1 (33.3)	3 (0.6)
31	1 (33.3)	2 (66.7)	3 (0.6)
32	3 (100.0)	0 (0.0)	3 (0.6)
33	1 (33.3)	2 (66.7)	3 (0.6)
34	1 (50.0)	1 (50.0)	2 (0.4)
35	1 (50.0)	1 (50.0)	2 (0.4)
36	2 (100)	0 (0.0)	2 (0.4)
37	2 (100)	0 (0.0)	2 (0.4)
38	1 (50.0)	1 (50.0)	2 (0.4)
39	0 (0.0)	1 (100.0)	1 (0.2)
40	1 (100)	0 (0.0)	1 (0.2)

TABLE 8 Baseline demographics

	Treatment group, *n* (%)		
	Laparoscopic surgery (*N* = 234)	Robotic surgery (*N* = 237)	Total, *n* (%) (*N* = 471)
Age (years), mean (SD)	65.5 (11.93)	64.4 (10.98)	64.9 (11.01)
ASA classification			
I: A normal healthy patient	52 (22.2)	39 (16.5)	91 (19.3)
II: A patient with mild systemic disease	124 (53.0)	150 (63.3)	274 (58.2)
III: A patient with severe systemic disease	52 (22.2)	46 (19.4)	98 (20.8)
IV: A patient with severe systemic disease that is a constant threat to life	1 (0.4)	0 (0.0)	1 (0.2)
Missing	5 (2.1)	2 (0.8)	7 (1.5)
Prior abdominal surgery			
Yes	67 (28.6)	62 (26.2)	129 (27.4)
No	162 (69.2)	174 (73.4)	336 (71.3)
Missing	5 (2.2)	1 (0.4)	6 (1.3)

TABLE 9 Summaries of operative and pathological variables

	Treatment group, *n* (%)		
Operative	Laparoscopic surgery (*N* = 230)	Robotic surgery (*N* = 236)	Total, *n* (%) (*N* = 466)
Operation performed			
HAR	19 (8.3)	28 (11.9)	47 (10.1)
LAR	165 (71.7)	152 (64.4)	317 (68.0)
APR	45 (19.6)	52 (22.0)	97 (20.8)
Other[a]	1 (0.4)	4 (1.7)	5 (1.1)
Operative time (minutes)			
Mean (SD)	261 (83.24)	298.5 (88.71)	280.0 (87.98)
Missing	4	1	5
Stoma formation			
Temporary	157 (68.3)	142 (60.2)	299 (64.2)
Permanent	49 (21.3)	53 (22.5)	102 (21.9)
No	24 (10.4)	41 (17.4)	65 (13.9)
Length of stay (days)			
Mean (SD)	8.2 (6.03)	8.0 (5.85)	8.1 (5.94)
Missing	13	14	27
Intraoperative conversion to open surgery			
Yes	28 (12.2)	19 (8.1)	47 (10.1)
No	202 (87.8)	217 (91.9)	419 (89.9)

TABLE 9 Summaries of operative and pathological variables (*continued*)

Operative	Treatment group, *n* (%)		
	Laparoscopic surgery (*N* = 230)	Robotic surgery (*N* = 236)	Total, *n* (%) (*N* = 466)
Pathology[b]			
T-stage			
0	24 (10.4)	22 (9.3)	46 (9.9)
1	20 (8.7)	24 (10.2)	44 (9.4)
2	61 (26.5)	64 (27.1)	125 (26.8)
3	114 (49.6)	117 (49.6)	231 (49.6)
4	8 (3.5)	5 (2.1)	13 (2.8)
Tx or missing	3 (1.3)	4 (1.7)	7 (1.5)
N-stage			
0	150 (65.2)	146 (61.9)	296 (63.5)
1	58 (25.2)	63 (26.7)	121 (26.0)
2	21 (9.1)	25 (10.6)	46 (9.9)
Missing	1 (0.4)	2 (0.8)	3 (0.6)
Lymph node yield			
Mean (SD)	24.1 (12.91)	23.2 (11.97)	23.6 (12.43)
Missing	9	1	10
Plane of surgery			
Mesorectal area (all specimens)			
Mesorectal plane	173 (75.2)	178 (75.4)	351 (75.3)
Intramesorectal plane	38 (16.5)	33 (14.0)	71 (15.2)
Muscularis propria plane	12 (5.2)	22 (9.3)	34 (7.3)
Missing	7 (3.1)	3 (1.3)	10 (2.1)
Sphincter area (APRs only)	(*n* = 45)	(*n* = 52)	(*n* = 97)
Levator plane	18 (40.0)	18 (34.6)	36 (37.1)
Sphincteric plane	19 (42.2)	22 (42.3)	41 (42.3)
Intrasphincteric/submucosal plane	5 (11.0)	9 (17.3)	14 (14.4)
Missing	3 (6.7)	3 (5.8)	6 (6.2)
CRM involvement	***(n = 224)***	***(n = 235)***	***(n = 459)***
Yes	14 (6.3)	12 (5.1)	26 (5.7)
No	210 (93.7)	223 (94.9)	433 (94.3)

a 'Other' operations: Laparoscopic group – 'Laparoscopic biopsy of peritoneum'. Robotic group – 'Dorsal pelvic exenteration, ureter resection distally right sided', 'Hartmann's procedure' (× 2), 'High anterior resection + subtotal colectomy'.
b Pathology data summarised here over the 466 patients who had an operation. CRM involvement summarised only over the 459 patients who had a pathology report available (i.e. the analysis population for that end point).

TABLE 10 Robotic and laparoscopic conversions

Intraoperative conversion pathway	Treatment group, *n* (%)		Total, *n* (%) (*N* = 471)
	Standard laparoscopic surgery, (*N* = 234)	Robotic-assisted laparoscopic surgery (*N* = 237)	
Laparoscopic	194 (82.9)	3 (1.3)	197 (41.8)
Laparoscopic → Open	28 (12.0)	0 (0.0)	28 (5.9)
Laparoscopic → Robotic	1 (0.4)	0 (0.0)	1 (0.2)
Robotic	7 (3.0)	209 (88.2)	216 (45.9)
Robotic → Open	0 (0.0)	14 (5.9)	14 (3.0)
Robotic → Laparoscopic	0 (0.0)	5 (2.1)	5 (1.1)
Robotic → Laparoscopic → Open	0 (0.0)	5 (2.1)	5 (1.1)
Did not receive surgery	1 (0.4)	0 (0.0)	1 (0.2)
Missing	3 (1.3)	1 (0.4)	4 (0.8)

Primary end point: conversion to open surgery

The rate of conversion to open surgery was 47 out of 466 (10.1%) patients overall: 28 out of 230 (12.2%) in the laparoscopic group and 19 out of 236 (8.1%) in the robotic group (unadjusted difference in proportions 4.12%, 95% CI –1.35% to 9.59%). There was no statistically significant difference between robotic surgery and conventional laparoscopic surgery with respect to odds of conversion (adjusted OR 0.614, 95% CI 0.311 to 1.211; $p = 0.16$).

The random intercept model was preferred to the random slope model, because the random slope model did not offer sufficient improvement in model fit, which is clear from both the non-significant likelihood ratio test result and the increase in AIC (see *Appendix 1*, *Table 51*).

Table 11 presents adjusted estimates of ORs and 95% CIs from the random intercept model, as well as crude summaries and unadjusted risk difference estimates and 95% CIs for conversion to open surgery by treatment group and also by each of the minimisation factors. The model shows significantly increased odds of conversion in obese patients versus underweight/normal patients (adjusted OR 4.691, 95% CI 2.080 to 10.581; $p < 0.01$) and in males versus females (adjusted OR 2.444, 95% CI 1.047 to 5.708; $p = 0.04$). Patients whose intended procedure was a LAR had a significantly higher rate of conversion than patients whose intended procedure was APR (adjusted OR 5.435, 95% CI 1.595 to 18.519; $p = 0.007$). Operating surgeon had a mild to moderate effect on odds of conversion, as reflected by the ICC estimate of 0.056 (95% CI 0.007 to 0.056).

Subgroup analyses

Odds ratios presented in *Tables 13–15* are derived from the linear combination of the estimated treatment (main effect) and treatment-by-subgroup interaction terms on the logit scale. The *p*-values are presented for the test of the treatment effect within each subgroup – this is the first column of *p*-values [e.g. in *Table 13* the test that the treatment effect is null (OR = 1) within the male subgroup is 0.0429]. The *p*-values are also presented for the test of heterogeneity of treatment effect across subgroups, the details of which are given in the footnotes of the tables (*Table 12*).

In the sex subgroup analysis, 39 out of 317 (12.3%) male patients underwent conversion to laparotomy: 25 out of 156 (16.0%) in the laparoscopic group and 14 out of 161 (8.7%) in the robotic group (unadjusted difference in proportions 7.3%, 95% CI 0.1% to 14.6%). There were 8 out of 149 (5.4%) female patients who underwent conversion to laparotomy: 3 out of 74 (4.1%) in the laparoscopic group and 5 out of 75 (6.7%) in the robotic group (unadjusted difference in proportions –2.6%, 95% CI –9.8% to 4.6%).

TABLE 11 Conversion to open surgery: adjusted estimates of ORs and 95% CIs from random intercept model

Effect: comparator group (vs. reference group)	Group [number of conversions/number of patients (%)]		Risk difference (unadjusted 95% CI)	OR (adjusted)	95% CI for OR (adjusted)	*p*-value
	Reference	Comparator				
Treatment: robotic surgery (vs. laparoscopic)	28/230 (12.2)	19/236 (8.1)	4.1 (–1.4 to 9.6)	0.614	0.311 to 1.211	0.16
Sex: male (vs. female)	8/149 (5.4)	39/317 (12.3)	–6.9 (–12.1 to –1.8)	2.444	1.047 to 5.708	0.04
BMI class: overweight (vs. underweight/normal)	13/179 (7.3)	9/180 (5.0)	2.3 (–2.7 to 7.2)	0.538	0.210 to 1.374	0.19
BMI class: obese (vs. underweight/normal)	13/179 (7.3)	25/107 (23.4)	–16.1 (–25.0 to –7.2)	4.691	2.080 to 10.581	0.0002
Previous radiotherapy or chemoradiotherapy: yes (vs. no)	27/262 (10.3)	20/204 (9.8)	0.5 (–5.0 to 6.0)	1.069	0.504 to 2.265	0.86
Intended procedure: HAR (vs. LAR)	37/312 (11.9)	6/68 (8.8)	3.0 (–4.6 to 10.7)	0.551	0.194 to 1.563	0.26
Intended procedure: APR (vs. LAR)	37/312 (11.9)	4/86 (4.7)	7.2 (1.5 to 12.9)	0.184	0.054 to 0.627	0.007

TABLE 12 Conversion to open surgery: estimate of the variance component relating to operating surgeon from random intercept model

Effect	Variance component		ICC	95% CI for ICC	
	Estimate	Standard error		Lower limit	Upper limit
Operating surgeon (random effect)	0.626	0.431	0.050	0.007	0.056

TABLE 13 Conversion to open surgery (subgroup analysis): ORs for treatment effect by sex

Effect	Surgery [number of conversions/number of patients (%)]		Risk difference (unadjusted 95% CI)	OR (adjusted 95% CI)[a]	*p*-value	
	Laparoscopic	Robotic				
Treatment in males: robotic surgery (vs. laparoscopic)	25/156 (16.0)	14/161 (8.7)	7.3 (0.1 to 14.6)	0.455 (0.209 to 0.987)	0.0429	0.0939[b]
Treatment in females: robotic surgery (vs. laparoscopic)	3/74 (4.1)	5/75 (6.7)	–2.6 (–9.8 to 4.6)	2.022 (0.425 to 9.621)	0.3757	

a Adjusted for BMI class, preoperative radiotherapy, intended procedure and operating surgeon.
b The *p*-value for the treatment effect is referring to a test of heterogeneity of treatment effect between the subgroups. ORs derived from the treatment term and treatment-by-sex interaction term.

TABLE 14 Conversion to open surgery (subgroup analysis): ORs for treatment effect by WHO obesity classification[29]

Effect	Surgery [number of conversions/number of patients (%)]		Risk difference (unadjusted 95% CI)	OR (adjusted 95% CI)[a]	*p*-value	
	Laparoscopic	Robotic				
Treatment in obese patients: robotic-assisted surgery (vs. laparoscopic)	15/54 (27.8)	10/53 (18.9)	8.9 (–7.0 to 24.8)	0.583 (0.212 to 1.602)	0.2944	0.6862[b]
Treatment in overweight patients: robotic-assisted surgery (vs. laparoscopic)	6/90 (6.7)	3/90 (3.3)	3.3 (–3.0 to 9.7)	0.508 (0.117 to 2.213)	0.3661	0.7509[b]
Treatment in underweight and normal patients: robotic-assisted surgery (vs. laparoscopic)	7/86 (8.1)	6/93 (6.5)	1.7 (–6.0 to 9.3)	0.751 (0.227 to 2.492)	0.6396	

a Adjusted for sex, preoperative radiotherapy, intended procedure and operating surgeon.
b The *p*-value for the treatment effect is referring to a (pairwise) test of heterogeneity of treatment effect between the subgroups. For example, the second *p*-value in the 'Treatment in obese patients' row refers to a test of heterogeneity of treatment effect between obese patients and underweight/normal patients.

TABLE 15 Conversion to open surgery (subgroup analysis): ORs for treatment effect by operation type

Effect	Surgery [number of conversions/number of patients (%)]		Risk difference (unadjusted 95% CI)	OR (adjusted 95% CI)[a]	*p*-value	
	Laparoscopic	Robotic				
Treatment (HAR): robotic-assisted surgery (vs. laparoscopic)	2/19 (10.5)	2/28 (7.1)	3.4 (–13.4 to 20.2)	0.771 (0.078 to 7.617)	0.8234	0.7106[b]
Treatment (APR): robotic-assisted surgery (vs. laparoscopic)	4/45 (8.9)	4/52 (7.7)	1.2 (–9.8 to 12.2)	0.705 (0.144 to 3.452)	0.6656	0.6833[b]
Treatment (LAR): robotic-assisted surgery (vs. laparoscopic)	22/165 (13.3)	11/152 (7.2)	6.1 (–0.5 to 12.7)	0.486 (0.210 to 1.123)	0.0909	

a Adjusted for sex, BMI class, preoperative radiotherapy and operating surgeon.
b The *p*-value for the treatment effect is referring to a (pairwise) test of heterogeneity of treatment effect between the subgroups. For example, the second *p*-value in the 'Treatment (HAR)' row refers to a test of heterogeneity of treatment effect between patients who underwent HAR and patients who underwent LAR.
Five patients underwent a procedure other than HAR, APR or LAR, 1 in the laparoscopic treatment group (no conversion to open surgery) and four in the robotic treatment group (two conversions). These patients were excluded from this model.

A Wald test of interaction between treatment effect and sex in the adjusted model yielded $p = 0.094$. This acts as moderate evidence that the difference between treatment groups is different for males and females. Furthermore, the estimated adjusted OR for conversion to laparotomy (robotic vs. conventional laparoscopic) in males is 0.455 (95% CI 0.209 to 0.987; $p = 0.043$), suggesting that there may in fact be a significant benefit of robotic surgery compared with laparoscopic surgery in terms of odds of conversion in male patients.

No substantial interactions between treatment effect and BMI or type of operation were found. The treatment effect OR in patients who underwent LAR was 0.486 (95% CI 0.210 to 1.123; $p = 0.091$), which may warrant further investigation into a potential benefit of robotic surgery in this group of patients.

Sensitivity analysis: learning effects

For 464 out of 471 (98.5%) patients we had sufficient data to derive the level of experience of the operating surgeon, expressed in terms of 'number of previous laparoscopic operations performed' and 'number of previous robotic operations performed' by the operating surgeon, at the time of the patient's operation. *Table 16* presents the distribution across patients of the previous experience of the operating surgeon who performed their operation. The amount of previous laparoscopic and robotic experience varied widely between participating surgeons, and there was a clear disparity between laparoscopic and robotic experience.

The patient-level variables 'number of previous laparoscopic operations performed by the operating surgeon, at the time of the patient's operation' and 'number of previous robotic operations performed by the operating surgeon, at the time of the patient's operation' were added to the primary analysis model as centred, linear fixed effects terms. Interactions between each of these variables and the treatment allocation were also included. The resulting estimated treatment effect ORs at various levels of operating surgeon laparoscopic and robotic experience are presented in *Table 17*. The model suggests that increasing operating surgeon robotic experience notably affects the treatment effect OR in favour of robotic surgery, regardless of the level of laparoscopic experience. The full fitted model is given in *Tables 18* and *19*, with untransformed estimates (i.e. on the log-odds scale).

TABLE 16 Previous experience of operating surgeons

	Previous experience (number of patients)	
Statistic	Laparoscopic ($n = 464$)	Robotic ($n = 464$)
Mean (SD)	152.5 (178.38)	67.9 (48.75)
Median (range)	91.4 (10.0–853.0)	49.5 (10.3–183.0)
(Q1, Q3)	(44.9, 180.1)	(30.4, 101.3)

Q1, first interquartile; Q3, third interquartile.

TABLE 17 Estimated treatment effect OR by surgeon experience

	Surgeon's experience level (number of previous operations)			
Effect	Laparoscopic	Robotic	OR (robotic vs. laparoscopic)	95% CI for OR
Primary analysis model	–	–	0.614	0.311 to 1.211
Treatment: robotic surgery (vs. laparoscopic)				
Learning effects model	45	30	0.961	0.336 to 2.747
Treatment: robotic surgery (vs. laparoscopic)		50	0.691	0.277 to 1.721
		100	0.303	0.090 to 1.018
	91	30	0.963	0.383 to 2.424
		50	0.692	0.317 to 1.513
		100	0.303	0.096 to 0.959
	180	30	0.966	0.416 to 2.245
		50	0.694	0.336 to 1.437
		100	0.304	0.094 to 0.989

TABLE 18 Conversion to open surgery (learning effects): adjusted estimates of ORs and 95% CIs from random intercept model including covariates related to operating surgeon's experience

Effect	Estimate	Standard error	Pr > |t|	95% CI
Intercept	–3.0056	0.5794	< 0.0001	–4.1787 to –1.8326
Treatment: robotic surgery (vs. laparoscopic)	–0.6618	0.3893	0.0899	–1.4271 to 0.1035
Sex: male (vs. female)	0.8688	0.4355	0.0467	0.01283 to 1.7248
BMI class: overweight (vs. underweight/normal)	–0.6763	0.4832	0.1623	–1.6262 to 0.2735
BMI class: obese (vs. underweight/normal)	1.4781	0.4257	0.0006	0.6413 to 2.3149
Previous radiotherapy or chemoradiotherapy: yes (vs. no)	0.1537	0.3924	0.6956	–0.6177 to 0.9250
Intended procedure: HAR (vs. LAR)	–0.5234	0.5342	0.3278	–1.5734 to 0.5267
Intended procedure: APR (vs. LAR)	–1.7021	0.6267	0.0069	–2.9340 to –0.4702
Surgeon's laparoscopic experience level (number of previous operations)	–0.00038	0.001759	0.8307	–0.00383 to 0.003082
Surgeon's robotic experience level (number of previous operations)	–0.00232	0.006330	0.7141	–0.01476 to 0.01012
Interaction term: treatment × surgeon's laparoscopic experience level	0.000037	0.002728	0.9891	–0.00533 to 0.005399
Interaction term: treatment × surgeon's robotic experience level	–0.01651	0.009887	0.0958	–0.03594 to 0.002929

TABLE 19 Conversion to open surgery (learning effects): estimate of the variance component from random intercept model

	Variance component	
Effect	Estimate	Standard error
Operating surgeon (random effect)	0.640	0.429

Sensitivity analysis: actual operating surgeon

The overall proportion of patients who were operated on by a surgeon other than their intended operating surgeon was low [42/466 (9.0%)] and occurred mainly at high recruiting sites, so the discrepancies were not influential on the results.

Adjusting for actual operating surgeon instead of intended operating surgeon made little difference to the model estimates. In particular, the treatment effect estimate (OR) adjusting for actual operating surgeon was 0.612 (95% CI 0.310 to 1.207), a negligible change from the primary analysis model.

Further details are given in *Appendix 1*, *Sensitivity analysis: actual operating surgeon – further details*.

Sensitivity analysis: actual procedure

The numbers of patients whose actual procedure was different from their intended procedure (the stratification factor) are summarised in *Table 20*. There were 65 out of 466 patients (13.9%) who had a procedure other than their 'intended procedure'. The most common discrepancy was for patients whose intended procedure was a HAR to actually undergo a LAR.

TABLE 20 Actual procedure performed vs. intended procedure

Actual procedure performed	Intended procedure, *n* (%)			
	HAR (*N* = 68)	LAR (*N* = 312)	APR (*N* = 86)	Total (*N* = 466)
HAR	37 (54.4)	10 (3.2)	0 (0.0)	47 (10.1)
LAR	29 (42.6)	283 (90.7)	5 (5.8)	317 (68.0)
APR	1 (1.5)	15 (4.8)	81 (94.2)	97 (20.8)
Other	1 (1.5)	4 (1.3)	0 (0.0)	5 (1.1)

Adjusting for actual procedure instead of intended procedure had a minor effect on model estimates. The treatment effect estimate (OR) adjusting for actual procedure was 0.572 (95% CI 0.289 to 1.132), which was a minor change from the primary analysis model, although it still points to the same conclusion.

Further details are given in *Appendix 2*.

Key secondary end point: circumferential resection margin positivity (CRM+)

A total of 459 (98.5%) patients of the 466 who had an operation had complete pathology data available (*Table 21*). In that group, 26 out of 459 (5.7%) patients were CRM+, 14 out of 224 (6.25%) patients in the laparoscopic group and 12 out of 235 (5.11%) patients in the robotic group (unadjusted difference in proportions 1.14%, 95% CI –3.10% to 5.38%). There was no statistically significant difference in the odds of CRM+ between the groups (adjusted OR 0.785, 95% CI 0.350 to 1.762; $p = 0.56$). It should be noted that the variance component estimate for operating surgeon is 0, and consequently there is not a valid standard error estimate for this. This indicates that the variation of odds of CRM+ between surgeons was negligible (*Table 22*).

TABLE 21 Circumferential resection margin positivity: adjusted estimates of ORs and 95% CIs from random intercept model

Effect: comparator group (vs. reference group)	Group [number of conversions/number of patients (%)]		Risk difference (unadjusted 95% CI)	OR (adjusted)	95% CI for OR (adjusted)	*p*-value
	Reference	Comparator				
Treatment: robotic surgery (vs. laparoscopic)	14/224 (6.3)	12/235 (5.1)	1.1 (–3.1 to 5.4)	0.785	0.350 to 1.762	0.5566
Sex: male (vs. female)	4/148 (2.7)	22/311 (7.1)	–4.4 (–8.2 to –0.5)	2.770	0.928 to 8.270	0.0679
BMI class: overweight (vs. underweight/normal)	12/178 (6.7)	12/176 (6.8)	–0.1 (–5.3 to 5.2)	1.099	0.471 to 2.563	0.8272
BMI class: obese (vs. underweight/normal)	12/178 (6.7)	2/105 (1.9)	4.8 (0.3 to 9.4)	0.263	0.057 to 1.216	0.0872
Previous radiotherapy or chemoradiotherapy: yes (vs. no)	13/258 (5.0)	13/201 (6.5)	–1.4 (–5.8 to 2.9)	1.136	0.491 to 2.628	0.7647
Intended procedure: HAR (vs. LAR)	16/308 (5.2)	2/68 (2.9)	2.3 (–7.0 to 2.5)	0.593	0.129 to 2.736	0.5022
Intended procedure: APR (vs. LAR)	16/308 (5.2)	8/83 (9.6)	–4.4 (–11.3 to 2.4)	2.010	0.799 to 5.056	0.1377

TABLE 22 Circumferential resection margin positivity: estimate of the variance component from random intercept model

Effect	Variance component	
	Estimate	Standard error
Operating surgeon (random effect)	0.000	N/A

N/A, not applicable.

Subgroup analyses

As a result of the low frequency of CRM+, many subgroups had insufficient numbers to yield meaningful estimates of the treatment effect.

Further details are given in *Appendix 2*.

Key secondary end point: 3-year local recurrence

Follow-up times are summarised in *Table 23*. Median follow-up time from randomisation was 3.1 years. The seven patients with missing data in *Table 23* had non-standard circumstances; three patients had benign disease (two laparoscopic, one robotic), one patient did not undergo surgery (laparoscopic) and three patients had palliative surgery only (one laparoscopic, two robotic). These seven patients were censored at time 0.

A local recurrence was observed in 30 out of 471 (6.4%) patients, 14 out of 234 (6.0%) in the laparoscopic group and 16 out of 237 (6.8%) in the robotic group. The date of local recurrence was defined as the date of the relevant assessment (i.e. clinical, radiological and pathological) that first detected the local recurrence.

The estimated cumulative incidence of local recurrence in each treatment group is presented in *Figure 2* (note that the *y*-axis is truncated to 0–0.1). At 3 years, the estimated difference (robotic minus laparoscopic) in cumulative incidence of local recurrence is 0.002 (95% CI –0.041 to 0.046).

Table 24 presents the estimated adjusted HRs and corresponding 95% CIs and Wald test *p*-values from the shared frailty model. There is not a statistically significant difference between the treatment groups. The estimated adjusted HR suggests that a patient undergoing robotic surgery is 1.137 (95% CI 0.554, 2.335; $p = 0.756$) times more likely to experience local recurrence than a patient undergoing laparoscopic surgery, all else being equal.

There appears to be a substantial difference in probability of local recurrence between males and females, with males much more likely to experience the event (adjusted HR 3.184, 95% CI 1.109 to 9.174; $p = 0.031$). This is reflected in the plot of cumulative incidence by sex in *Figure 3* (note that the *y*-axis is truncated to 0–0.1).

TABLE 23 Length of follow-up from randomisation, by treatment group

Length of follow-up from randomisation (years)	Treatment group		
	Standard laparoscopic surgery ($n = 234$)	Robotic-assisted laparoscopic surgery ($n = 237$)	Total ($n = 471$)
Mean (SD)	3.1 (1.07)	3.2 (1.12)	3.2 (1.10)
Median (range)	3.1 (0.0–6.1)	3.1 (0.0–6.0)	3.1 (0.0–6.1)
(Q1, Q3)	(3.0, 4.0)	(3.0, 3.9)	(3.0, 4.0)
Missing	4	3	7

Q1, first interquartile; Q3, third interquartile.

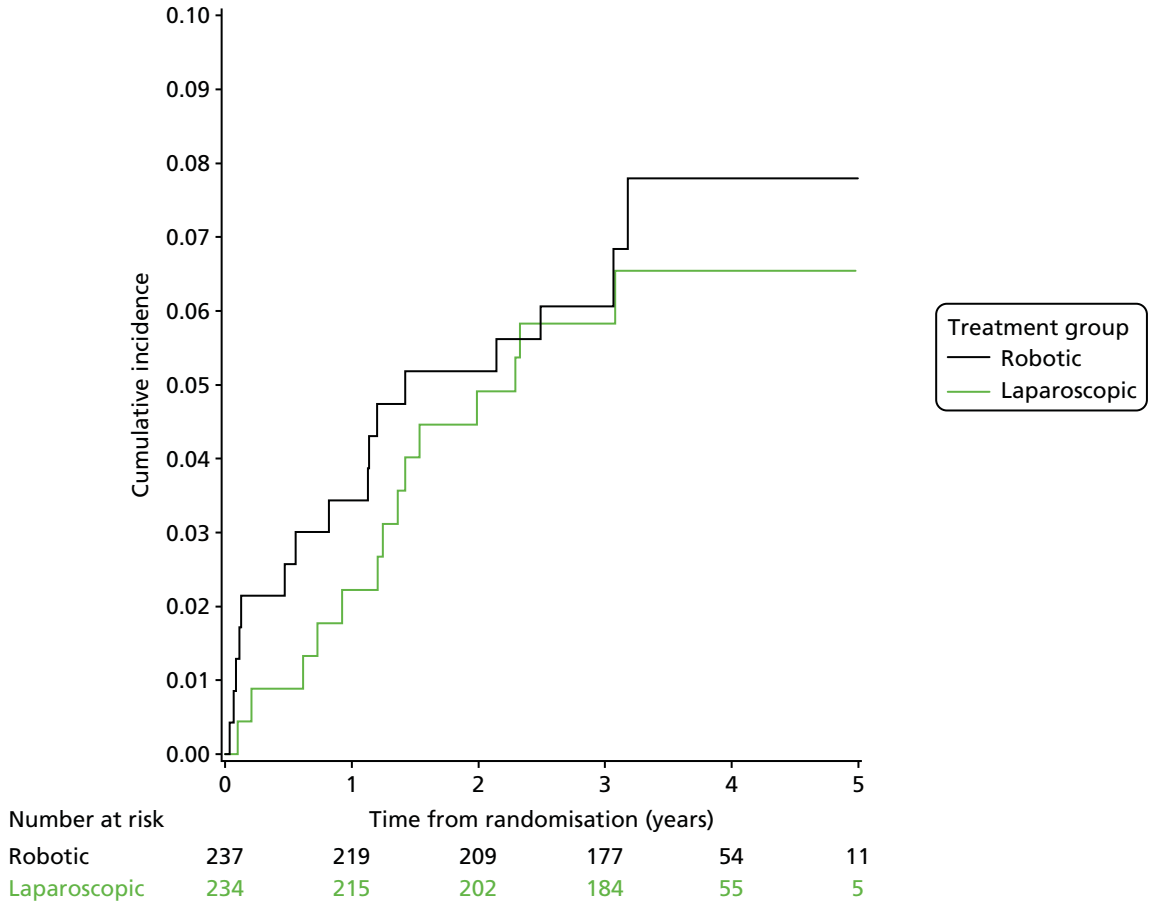


FIGURE 2 Estimated cumulative incidence of local recurrence, by treatment group.

TABLE 24 Three-year local recurrence: adjusted estimates of HRs and 95% CIs from random shared frailty model

Parameter	HR	95% CI	*p*-value
Treatment allocation: robotic (vs. laparoscopic)	1.137	0.554 to 2.335	0.7257
Sex: male (vs. female)	3.184	1.109 to 9.174	0.0314
Neo-adjuvant therapy: yes (vs. no)	1.083	0.510 to 2.299	0.8361
BMI classification obese (vs. underweight/normal)	0.954	0.345 to 2.634	0.927
BMI classification overweight (vs. underweight/normal)	1.366	0.603 to 3.095	0.4545
Intended procedure HAR (vs. LAR)	0.645	0.187 to 2.224	0.4873
Intended procedure APR (vs. LAR)	1.07	0.423 to 2.707	0.886

Subgroup analyses

None of the prespecified subgroup analyses yielded meaningful evidence of an interaction between treatment effect and subgroup, or evidence of a treatment effect within any individual subgroup. Given the clear (main) effect of sex on local recurrence, and the clinical plausibility of a potential difference of treatment effect by sex, an ad-hoc sex subgroup analysis was performed. Similarly, this subgroup analysis showed no evidence of a subgroup by treatment interaction and no significant treatment effect within either subgroup.

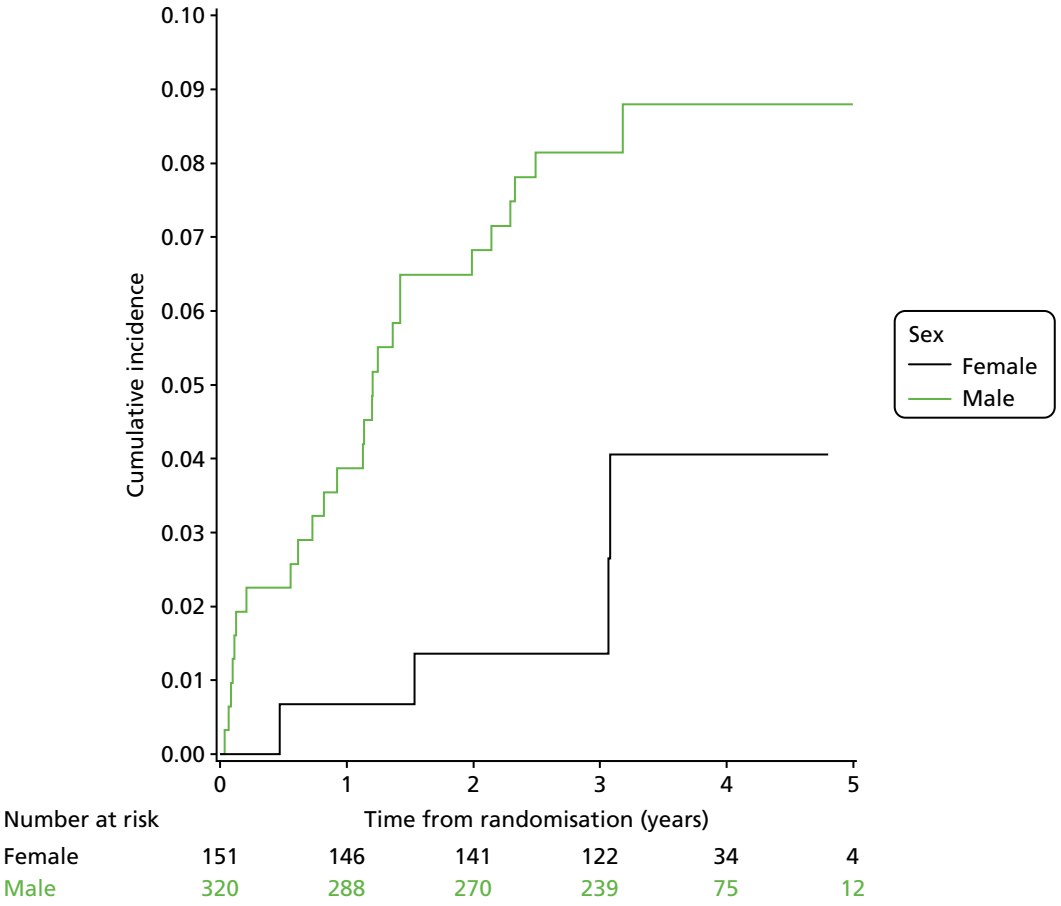


FIGURE 3 Estimated cumulative incidence of local recurrence by sex.

Intraoperative complications

A total of 70 out of 466 (15.0%) patients had an intraoperative complication, 34 out of 230 (14.8%) in the laparoscopic group and 36 out of 236 (15.3%) in the robotic group (unadjusted risk difference 0.5%, 95% CI –6.0% to 7.0%). *Table 25* presents the numbers of patients experiencing different types of intraoperative complications. The most common intraoperative complications were damage to an organ/structure, significant haemorrhage and surgical equipment failure. *Table 26* presents the multilevel logistic regression model. There was no significant difference between the groups (adjusted OR 1.020, 95% CI 0.599 to 1.736; $p = 0.94$). There is significant evidence of a difference in the odds of having an intraoperative complication between males and females (adjusted OR 3.083, 95% CI 1.543 to 6.158; $p = 0.0015$). Note that the variance component estimate for operating surgeon is 0 and consequently there is not a valid standard error estimate for this. This indicates that the variation of odds of CRM+ between surgeons was negligible (*Table 27*).

Thirty-day postoperative complications

A total of 151 out of 466 (32.4%) patients had a postoperative complication within 30 days of their operation, 73 out of 230 (31.7%) in the laparoscopic group and 78 out of 236 (33.1%) in the robotic group (unadjusted risk difference –1.3%, 95% CI –9.8% to 7.2%). *Table 28* presents the numbers of patients who experienced different types of postoperative complications within 30 days of their operation. The most common were gastrointestinal complications (including anastomotic leak), surgical site infections and urinary complications. *Tables 29* and *30* present the multilevel logistic regression model. There was no significant difference between the groups (adjusted OR 1.043, 95% CI 0.689 to 1.581; $p = 0.84$). There is significant evidence of a difference in the odds of having a postoperative complication within 30 days of an operation between males and females (adjusted OR 3.083, 95% CI 1.573 to 4.183; $p = 0.0002$).

TABLE 25 Numbers of patients experiencing intraoperative complications

Intraoperative complications	Treatment group, *n* (%)	
	Laparoscopic surgery (*N* = 230)	Robotic surgery (*N* = 236)
Damage to organ/structure	5 (2.2)	11 (4.7)
Significant haemorrhage	11 (4.8)	4 (1.7)
Equipment failure	6 (2.6)	8 (3.4)
Faecal contamination	6 (2.6)	7 (3.0)
Anastomotic complication	6 (2.6)	7 (3.0)
Iatrogenic tumour perforation	3 (1.3)	2 (0.8)
Inadequate tumour localisation/clearance	2 (0.9)	2 (0.8)
Respiratory event	2 (0.9)	1 (0.4)
Cardiac event	1 (0.4)	1 (0.4)
Overall	34 (14.8)	36 (15.3)

Counts are the number of patients who experienced the complication; the categories are not mutually exclusive.

TABLE 26 Intraoperative complications: adjusted estimates of ORs and 95% CIs from random intercept model

Effect: comparator group (vs. reference group)	Group [number of conversions/number of patients (%)]		Risk difference (unadjusted 95% CI)	OR (adjusted)	95% CI for OR (adjusted)	*p*-value
	Reference	Comparator				
Treatment: robotic surgery (vs. laparoscopic)	34/230 (14.8)	36/236 (15.3)	–0.5 (–6.0 to 7.0)	1.020	0.599 to 1.736	0.9426
Sex: male (vs. female)	11/149 (7.4)	59/317 (18.6)	–11.2 (–17.2 to –5.2)	3.083	1.543 to 6.158	0.0015
BMI class: overweight (vs. underweight/normal)	25/179 (14.0)	30/180 (16.7)	–2.7 (–10.2 to 4.7)	1.280	0.699 to 2.344	0.4222
BMI class: obese (vs. underweight/normal)	25/179 (14.0)	15/107 (14.0)	–0.1 (–8.4 to 8.3)	0.939	0.456 to 1.931	0.8634
Previous radiotherapy or chemoradiotherapy: yes (vs. no)	24/262 (9.2)	46/204 (22.6)	–13.4 (–20.1 to –6.7)	3.480	1.955 to 6.192	< 0.0001
Intended procedure: HAR (vs. LAR)	53/312 (17.0)	9/68 (13.2)	3.8 (–5.3 to 12.8)	1.143	0.502 to 2.601	0.7504
Intended procedure: APR (vs. LAR)	53/312 (17.0)	8/86 (9.3)	7.7 (0.3 to 15.1)	0.403	0.179 to 0.908	0.0284

TABLE 27 Intraoperative complications: estimate of the variance component from random intercept model

Effect	Variance component	
	Estimate	Standard error
Operating surgeon (random effect)	0.0	N/A

N/A, not applicable.

TABLE 28 Numbers of patients experiencing postoperative complications within 30 days of their operation

30-day complications	Treatment group, *n* (%)	
	Laparoscopic surgery (*N* = 230)	Robotic surgery (*N* = 236)
Gastrointestinal complication	40 (17.4)	35 (14.8)
Surgical site infection	19 (8.3)	21 (8.9)
Urinary complication	14 (6.1)	17 (7.2)
Respiratory complication	6 (2.6)	4 (1.7)
Cardiac complication	6 (2.6)	3 (1.3)
Other	12 (5.2)	17 (7.2)
Overall	73 (31.7)	78 (33.1)

Counts are the number of patients who experienced the complication; the categories are not mutually exclusive.

TABLE 29 Thirty-day postoperative complications: adjusted estimates of ORs and 95% CIs from random intercept model

Effect: comparator group (vs. reference group)	Group [number of conversions/number of patients (%)]		Risk difference (unadjusted 95% CI)	OR (adjusted)	95% CI for OR (adjusted)	*p*-value
	Reference	Comparator				
Treatment: robotic surgery (vs. laparoscopic)	73/230 (31.7)	78/236 (33.1)	–1.3 (–9.8 to 7.2)	1.043	0.689 to 1.581	0.8407
Sex: male (vs. female)	30/149 (20.1)	121/317 (38.2)	–18.0 (–26.4 to –9.7)	2.565	1.573 to 4.183	0.0002
BMI class: overweight (vs. underweight/normal)	53/179 (29.6)	52/180 (28.9)	0.1 (–8.7 to 10.1)	0.946	0.578 to 1.548	0.8236
BMI class: obese (vs. underweight/normal)	53/179 (29.6)	46/107 (43.0)	–13.4 (–24.9 to –1.9)	1.758	1.022 to 3.024	0.0417
Previous radiotherapy or chemoradiotherapy: yes (vs. no)	75/262 (28.6)	76/204 (37.3)	–8.6 (–17.2 to –0.3)	1.432	0.906 to 2.264	0.1241
Intended procedure: HAR (vs. LAR)	101/312 (32.4)	15/68 (22.1)	10.3 (–21.5 to 0.8)	0.599	0.304 to 1.180	0.1383
Intended procedure: APR (vs. LAR)	101/312 (32.4)	35/86 (40.7)	–8.3 (–19.9 to 3.3)	1.278	0.740 to 2.209	0.3778

TABLE 30 Thirty-day postoperative complications: estimate of the variance component from random intercept model

Effect	Variance component	
	Estimate	Standard error
Operating surgeon (random effect)	0.286	0.213

Six-month postoperative complications (after 30 days)

A total of 72 out of 466 (15.5%) patients had a postoperative complication after 30 days and within 6 months of their operation, 38 out of 230 (16.5%) in the laparoscopic group and 34 out of 236 (14.4%) in the robotic group (unadjusted risk difference 2.1%, 95% CI –4.5% to 8.7%). *Table 31* presents the numbers of patients to experience different types of postoperative complications after 30 days and within 6 months of their operation. The most common was gastrointestinal complication (including anastomotic leak). *Tables 32* and *33* present the multilevel logistic regression model. There was no significant difference between the groups (adjusted OR 0.719, 95% CI 0.411 to 1.258; $p = 0.25$).

TABLE 31 Numbers of patients experiencing postoperative complications after 30 days and within 6 months of their operation

6-month complications (after 30 days)	Treatment group, *n* (%)	
	Laparoscopic surgery ($N = 230$)	Robotic surgery ($N = 236$)
Gastrointestinal complication	18 (7.8)	20 (8.5)
Urinary complication	6 (2.6)	7 (3.0)
Surgical site infection	8 (3.5)	4 (1.7)
Respiratory complication	3 (1.3)	2 (0.8)
Cardiac complication	1 (0.4)	0 (0.0)
Cerebrovascular complication	1 (0.4)	0 (0.0)
Other	12 (5.2)	7 (3.0)
Overall	38 (16.5)	34 (14.4)

Counts are the number of patients who experienced the complication; the categories are not mutually exclusive.

TABLE 32 Six-month postoperative complications: adjusted estimates of ORs and 95% CIs from random intercept model

Effect: comparator group (vs. reference group)	Group [number of conversions/number of patients (%)]		Risk difference (unadjusted 95% CI)	OR (adjusted)	95% CI for OR (adjusted)	*p*-value
	Reference	Comparator				
Treatment: robotic surgery (vs. laparoscopic)	38/230 (16.5)	34/236 (14.4)	2.1 (–4.5 to 8.7)	0.719	0.411 to 1.258	0.2468
Sex: male (vs. female)	19/149 (12.8)	53/317 (16.7)	–4.0 (–10.7 to 2.8)	1.230	0.654 to 2.313	0.5197
BMI class: overweight (vs. underweight/normal)	31/179 (17.3)	24/180 (13.3)	4.0 (–3.5 to 11.4)	0.715	0.371 to 1.378	0.3156
BMI class: obese (vs. underweight/normal)	31/179 (17.3)	17/107 (15.9)	1.4 (–7.4 to 10.3)	0.663	0.316 to 1.390	0.2754
Previous radiotherapy or chemoradiotherapy: yes (vs. no)	31/262 (11.8)	41/204 (20.1)	–8.3 (–15.0 to –1.5)	1.704	0.906 to 3.206	0.0979
Intended procedure: HAR (vs. LAR)	50/312 (16.0)	6/68 (8.8)	7.2 (–0.7 to 15.1)	0.620	0.228 to 1.686	0.3479
Intended procedure: APR (vs. LAR)	50/312 (16.0)	16/86 (18.6)	–2.6 (–11.8 to 6.6)	1.166	0.561 to 2.423	0.6794

TABLE 33 Six-month postoperative complications: estimate of the variance component from random intercept model

Effect	Variance component	
	Estimate	Standard error
Operating surgeon (random effect)	1.343	0.708

Thirty-day operative mortality

Death within 30 days of operation was a rare event, with 2 out of 230 (0.87%) and 2 out of 236 (0.85%) events in the standard laparoscopic and robotic groups, respectively. All deaths involved a septic complication and were related to the surgical intervention. Owing to the small number of events, sophisticated statistical models were not fitted.

Patient self-reported bladder function

Higher I-PSS indicates worse bladder function and is measured on a scale of 0–35.

Baseline characteristics of the population of patients with complete I-PSS data, and a comparison with those patients with missing I-PSS data, are given in *Appendix 4*.

Figure 4 visualises the distribution of I-PSS scores at 6 months post operation in the two treatment groups. The distribution of scores is very similar between the groups.

Tables 34 and *35* present the multilevel generalised linear model. Normal errors were assumed, so the estimates represent differences in the mean I-PSS. The estimated difference in mean I-PSS (robotic minus standard) is –0.7426 (95% CI –2.0722 to 0.5870; $p = 0.2726$). On the 35-point scale, this is a very small effect size that is also not statistically significant. The baseline score is highly prognostic of the 6-month score. The estimated difference in mean I-PSS between patients with a difference in baseline score of 10 points, all else being equal, is 4.20 (95% CI 3.23 to 5.17; $p < 0.0001$).

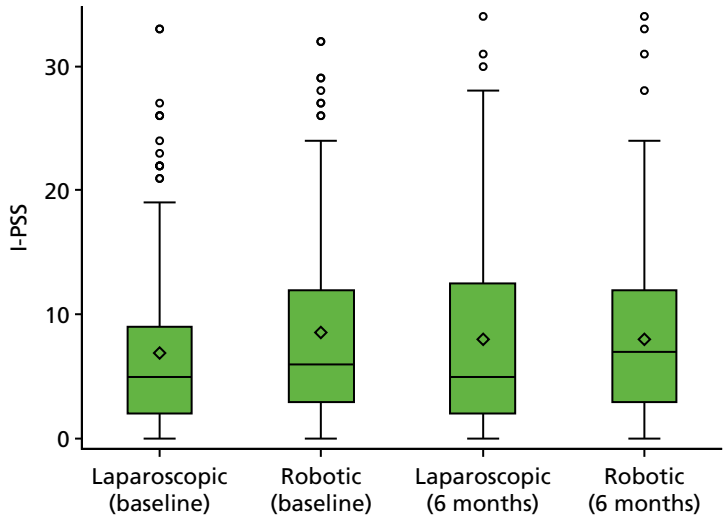


FIGURE 4 Box plot of observed I-PSS values at baseline and at 6 months post randomisation, by treatment group.

TABLE 34 The I-PSS: adjusted estimates of mean effects and 95% CIs from random intercept model

Effect	Estimate	Standard error	*p*-value	95% CI
Intercept	3.8249	0.9557	0.0003	1.8867 to 5.7631
Treatment: robotic-assisted surgery (vs. standard)	–0.7426	0.6757	0.2726	–2.0722 to 0.5870
Sex: male (vs. female)	1.7798	0.7425	0.0171	0.3188 to 3.2407
BMI class: overweight (vs. underweight/normal)	0.3740	0.7741	0.6293	–1.1493 to 1.8973
BMI class: obese (vs. underweight/normal)	0.2473	0.9268	0.7898	–1.5764 to 2.0710
Previous neo-adjuvant therapy: yes (vs. no)	–1.1450	0.7345	0.1201	–2.5903 to 0.3003
Intended procedure: HAR (vs. LAR)	–1.0208	1.0117	0.3138	–3.0116 to 0.9699
Intended procedure: APR (vs. LAR)	2.7760	0.9326	0.0031	0.9410 to 4.6111
Baseline I-PSS (1-unit increase)	0.4198	0.04933	< 0001	0.3228 to 0.5169

TABLE 35 The I-PSS: estimate of the variance component from random intercept model

Parameter	Subject	Estimate	Standard error
Intercept	Surgeon	1.2834	1.4209
Residual		38.9462	3.1275

Patient self-reported sexual function: males

A higher IIEF score indicates better sexual function; the score is measured on a scale of 5–75.

Baseline characteristics of the population of patients with complete IIEF data, and a comparison with those patients with missing IIEF data, are given in *Appendix 5*.

Figure 5 visualises the distribution of IIEF scores at 6 months post operation in the two treatment groups. The distribution of scores is very similar between the treatment groups. Median IIEF scores at 6 months were notably lower than at baseline in both groups.

Tables 36 and *37* present the multilevel generalised linear model. Normal errors were assumed, so the estimates represent differences in the mean IIEF score. The estimated difference in mean IIEF (robotic minus standard) is –0.8020 (95% CI –5.7039 to 4.1000; $p = 0.7468$). On the 70-point scale, this is a very small effect size that is also not statistically significant.

Patient self-reported sexual function: females

A higher FSFI score indicates better sexual function; the score is measured on a scale of 2–36.

Baseline characteristics of the population of patients with complete FSFI data, and a comparison with those patients with missing FSFI data, are given in *Appendix 6*.

Figure 6 visualises the distribution of FSFI scores at 6 months post operation in the two treatment groups. The distribution of scores is very similar between the treatment groups.

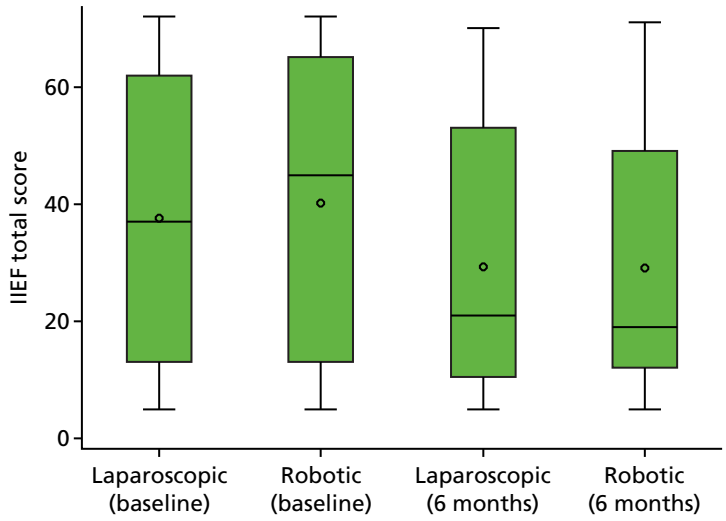


FIGURE 5 Box plot of observed IIEF values at baseline and at 6 months post randomisation, by treatment group.

TABLE 36 The IIEF: adjusted estimates of mean effects and 95% CIs from random intercept model

Effect	Estimate	Standard error	Pr > |t|	95% CI
Intercept	9.7690	3.6018	0.0104	2.4493 to 17.0887
Treatment: robotic-assisted surgery (vs. standard)	–0.8020	2.4793	0.7468	–5.7039 to 4.1000
BMI class: overweight (vs. underweight/normal)	–0.7386	2.9556	0.8030	–6.5823 to 5.1051
BMI class: obese (vs. underweight/normal)	3.0106	3.3590	0.3717	–3.6307 to 9.6519
Previous neo-adjuvant therapy: yes (vs. no)	–5.1767	3.0462	0.0915	–11.1996 to 0.8462
Intended procedure: HAR (vs. LAR)	7.2280	3.7064	0.0532	–0.1001 to 14.5562
Intended procedure: APR (vs. LAR)	–0.7213	3.6837	0.8450	–8.0046 to 6.5620
Baseline total IIEF score (1-unit increase)	0.5171	0.05045	< 0001	0.4174 to 0.6169

TABLE 37 The IIEF: estimate of the variance component from random intercept model

Parameter	Subject	Estimate	Standard error
Intercept	Surgeon	51.9161	29.1957
Residual		250.47	29.2417

Tables 38 and *39* present the multilevel generalised linear model. Normal errors were assumed, so the estimates represent differences in the mean FSFI score. The estimated difference in the mean FSFI score (robotic minus standard) is –1.2309 (95% CI –6.0030 to 3.5413; $p = 0.6010$). On the 34-point scale, this is a small effect size that is also not statistically significant.

Patient self-reported generic health

The SF-36 is a multipurpose, short-form health survey with 36 questions. It has eight scales of functional health: physical functioning, social functioning, role limitation physical, role limitation emotional, mental health, vitality, pain and general health that are scored on a 0–100 scale. It also provides a physical component score (PCS) and a mental component score (MCS), which are combinations of the eight scales.

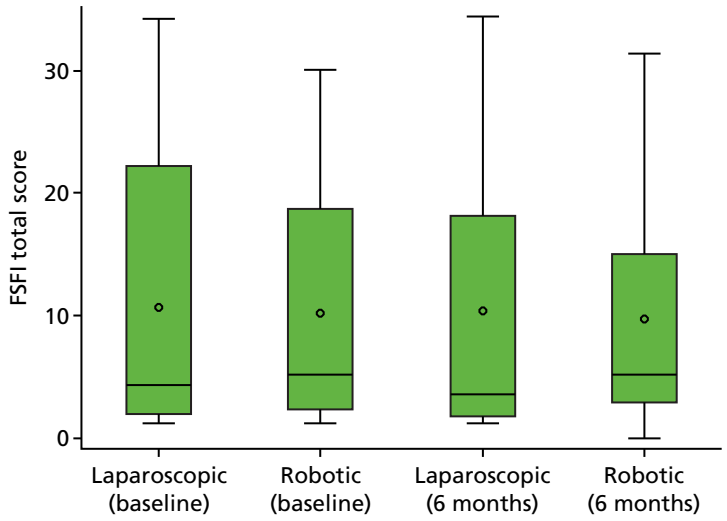


FIGURE 6 Box plot of observed FSFI values at baseline and at 6 months post randomisation, by treatment group.

TABLE 38 The FSFI: adjusted estimates of mean effects and 95% CIs from random intercept model

Effect	Estimate	Standard error	Pr > |t|	95% CI
Intercept	9.0710	3.2464	0.0116	2.2762 to 15.8657
Treatment: robotic-assisted surgery (vs. standard)	–1.2309	2.3258	0.6010	–6.0030 to 3.5413
BMI class: overweight (vs. underweight/normal)	4.1518	2.7584	0.1439	–1.5079 to 9.8116
BMI class: obese (vs. underweight/normal)	–0.9541	3.2873	0.7738	–7.6992 to 5.7909
Previous neo-adjuvant therapy: yes (vs. no)	–0.8097	2.7129	0.7676	–6.3761 to 4.7567
Intended procedure: HAR (vs. LAR)	–0.7669	3.2401	0.8147	–7.4151 to 5.8813
Intended procedure: APR (vs. LAR)	–4.9505	3.1579	0.1286	–11.4300 to 1.5289
Baseline FSFI score (1-unit increase)	0.4629	0.1147	0.0004	0.2275 to 0.6982

TABLE 39 The FSFI: estimate of the variance component from random intercept model

Parameter	Subject	Estimate	Standard error
Intercept	Surgeon	0.1703	10.2019
Residual		70.5888	16.9383

The SF-36 was collected at baseline, and at 30 days and at 6 months post operation.

A higher score indicates a better QoL.

Baseline characteristics of the population of patients with complete generic QoL data, and a comparison with those patients with missing generic QoL data, are given in *Appendix 7*.

Tables 40 and *41* show the multilevel linear model for the PCS and MCS, respectively. *Figures 7* and *8* illustrate the model estimates and 95% CIs at baseline and at 1 month and 6 months post randomisation of the average PCS and MCS respectively, split by treatment group. The estimated average difference in PCS between

TABLE 40 The SF-36v2 PCS: adjusted estimates of mean effects and 95% CIs from random intercept model

Effect	Estimate	Standard error	Pr > |t|	95% CI
Intercept	54.5476	1.6592		
30 days	–7.3421	1.7918	< 0.0001	–10.8595 to –3.8247
6 months	–2.1738	1.8099	0.2301	–5.7266 to 1.3791
Treatment: robotic-assisted laparoscopic surgery (vs. standard)	–0.1220	0.7672	0.8737	–1.6281 to 1.3840
Sex: female (vs. male)	–1.4117	0.6549	0.0314	–2.6974 to –0.1260
Neo-adjuvant therapy: no (vs. yes)	3.0683	0.8419	0.0003	1.4156 to 4.7210
Intended procedure: APR (vs. HAR)	–2.8637	1.4321	0.0459	–5.6750 to –0.05246
Intended procedure: LAR (vs. HAR)	–0.2736	1.1468	0.8115	–2.5249 to 1.9776
BMI class: obese (vs. underweight/normal)	–0.2153	1.0415	0.8363	–2.2597 to 1.8292
BMI class: overweight (vs. underweight/normal)	0.9225	0.8818	0.2958	–0.8085 to 2.6535
ASA grade: (II vs. I)	–3.8012	1.0400	0.0003	–5.8427 to –1.7597
ASA grade: (III vs. I)	–6.6250	1.2870	< 0001	–9.1513 to –4.0986
Robotic-assisted laparoscopic surgery and 30-day interaction	0.4651	0.8664	0.5916	–1.2357 to 2.1659
Robotic-assisted laparoscopic surgery and 6-month interaction	0.6086	0.8777	0.4882	–1.1143 to 2.3315
ASA grade II and 30-day interaction	2.4549	1.1467	0.0326	0.2039 to 4.7058
ASA grade III and 30-day interaction	3.4690	1.3853	0.0125	0.7495 to 6.1884
ASA grade II and 6-month interaction	0.5546	1.1382	0.6262	–1.6797 to 2.7889
ASA grade III and 6-month interaction	2.7739	1.3844	0.0455	0.05625 to 5.4916
No neo-adjuvant therapy and 30-day interaction	–3.4295	0.9152	0.0002	–5.2262 to –1.6329
No neo-adjuvant therapy and 6-month interaction	–2.9066	0.9271	0.0018	–4.7266 to –1.0867
APR and 30-day interaction	–2.6589	1.5995	0.0969	–5.7989 to 0.4810
APR and 6-month interaction	0.5959	1.6210	0.7133	–2.5862 to 3.7780
LAR and 30-day interaction	–1.9226	1.2971	0.1387	–4.4688 to 0.6237
LAR and 6-month interaction	–0.2858	1.3196	0.8286	–2.8764 to 2.3047
Obese and 30-day interaction	–2.6187	1.1657	0.0250	–4.9071 to –0.3303
Obese and 6-month interaction	–1.5758	1.1784	0.1816	–3.8891 to 0.7375
Overweight and 30-day interaction	–0.5450	0.9780	0.5775	–2.4647 to 1.3748
Overweight and 6-month interaction	–0.3320	0.9915	0.7378	–2.2784 to 1.6143

the groups at baseline is negligible: –0.1220 (95% CI –1.6281 to 1.3840; $p = 0.8737$). This is also the case at 1 month and 6 months post randomisation, as shown by the small magnitude and large *p*-values for the estimates of interaction between treatment effect and time (see *Table 40*). The estimated average difference in MCS between the groups at baseline is also negligible, –0.4875 (95% CI –2.6008 to 1.6258; $p = 0.6508$). Again this does not change notably over time, as seen in *Figure 8* and by the small magnitude and large *p*-values of the estimates of interaction between time and treatment effect in *Table 32*.

TABLE 41 The SF-36v2 MCS: adjusted estimates of mean effects and 95% CIs from random intercept model

Effect	Estimate	Standard error	Pr > |t|	95% CI
Intercept	40.4101	3.3900		
30 days	–2.3060	2.3656	0.3300	–6.9498 to 2.3378
6 months	4.4232	2.3888	0.0645	–0.2661 to 9.1125
Treatment: robotic-assisted laparoscopic surgery (vs. standard)	–0.4875	1.0766	0.6508	–2.6008 to 1.6258
Sex: female (vs. male)	–3.0157	0.9511	0.0016	–4.8828 to –1.1486
Neo-adjuvant therapy: no (vs. yes)	1.6362	1.1905	0.1697	–0.7008 to 3.9733
Intended procedure: APR (vs. HAR)	–2.8133	2.0057	0.1611	–6.7505 to 1.1240
Intended procedure: LAR (vs. HAR)	–0.7372	1.6047	0.6461	–3.8873 to 2.4130
BMI class: obese (vs. underweight/normal)	0.8777	1.4551	0.5466	–1.9786 to 3.7340
BMI class: overweight (vs. underweight/normal)	–0.2983	1.2335	0.8090	–2.7198 to 2.1232
Age	0.1339	0.04350	0.0022	0.04845 to 0.2193
ASA grade: (II vs. I)	0.2121	1.4829	0.8863	–2.6988 to 3.1230
ASA grade: (III vs. I)	–0.6905	1.8218	0.7048	–4.2668 to 2.8859
ASA grade: (IV vs. I)	–13.0801	11.6604	0.2623	–35.9699 to 9.8097
Robotic-assisted laparoscopic surgery and 30-day interaction	2.0753	1.1435	0.0699	–0.1694 to 4.3200
Robotic-assisted laparoscopic surgery and 6-month interaction	0.2681	1.1576	0.8169	–2.0042 to 2.5404
ASA grade II and 30-day interaction	0.5556	1.5142	0.7137	–2.4168 to 3.5281
ASA grade III and 30-day interaction	0.1340	1.8290	0.9416	–3.4564 to 3.7243
ASA grade II and 6-month interaction	–2.4933	1.5015	0.0972	–5.4408 to 0.4541
ASA grade III and 6-month interaction	–1.5419	1.8273	0.3990	–5.1289 to 2.0451
No neo-adjuvant therapy and 30-day interaction	–1.1184	1.2080	0.3548	–3.4897 to 1.2530
No neo-adjuvant therapy and 6-month interaction	–1.0979	1.2230	0.3696	–3.4987 to 1.3029
APR and 30-day interaction	–0.4158	2.1120	0.8440	–4.5618 to 3.7302
APR and 6-month interaction	–1.0937	2.1408	0.6096	–5.2960 to 3.1087
LAR and 30-day interaction	–1.4976	1.7126	0.3822	–4.8595 to 1.8644
LAR and 6-month interaction	–1.9039	1.7421	0.2748	–5.3237 to 1.5159
Obese and 30-day interaction	–1.7664	1.5395	0.2516	–4.7885 to 1.2557
Obese and 6-month interaction	0.4705	1.5564	0.7625	–2.5849 to 3.5258
Overweight and 30-day interaction	0.4266	1.2903	0.7410	–2.1063 to 2.9596
Overweight and 6-month interaction	1.2397	1.3073	0.3433	–1.3266 to 3.8059

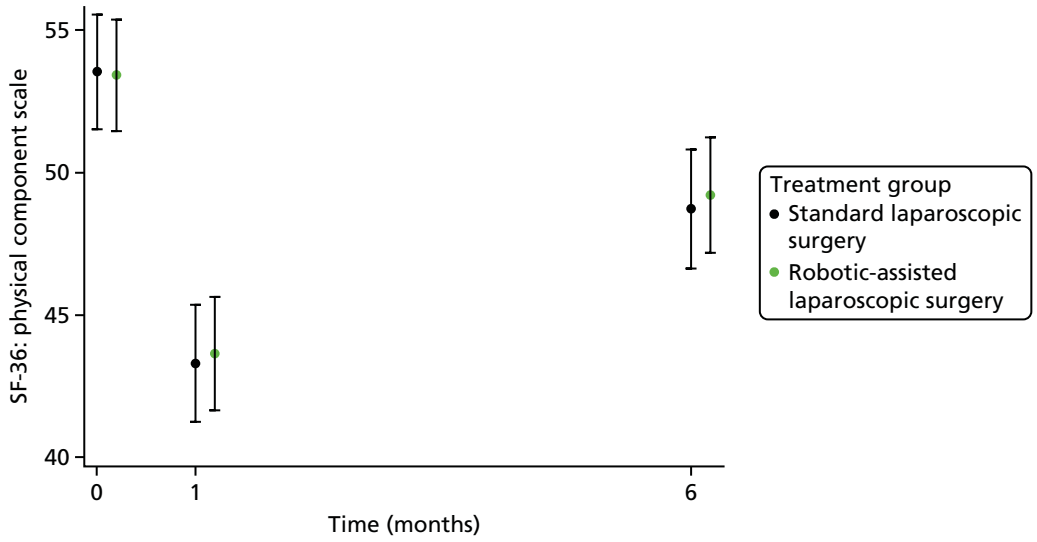


FIGURE 7 Adjusted estimates and 95% CIs of mean SF-36v2 PCS values at baseline, at 1 month and at 6 months post randomisation, by treatment group.

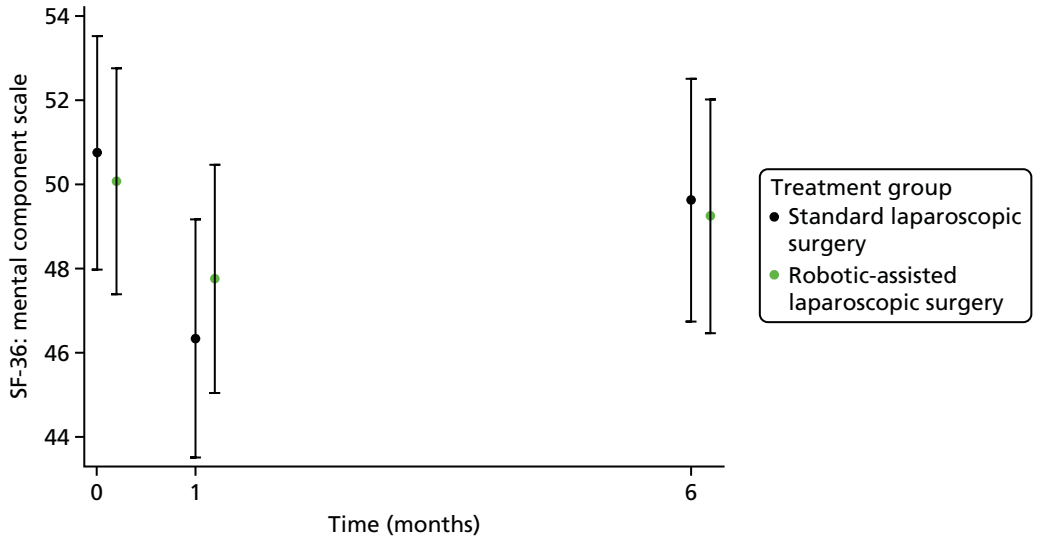


FIGURE 8 Adjusted estimates and 95% CIs of mean SF-36v2 MCS values at baseline, at 1 month and at 6 months post randomisation, by treatment group.

Patient self-reported fatigue

The MFI-20 is a self-report instrument. It contains 20 statements that cover different aspects of fatigue.

These 20 items are organised in five scales: general fatigue, physical fatigue, reduced activity, reduced motivation and mental fatigue.

The scores per item run from 1 to 5. For each scale, consisting of four items, a total score is calculated by summation of the scores of the individual items. Scores can range from the minimum of 4 to the maximum of 20. The use of a total score over all 20 items is not recommended.

The MFI-20 was collected at baseline, at 30 days post operation and at 6 months post operation. A higher score indicates more fatigue.

Baseline characteristics of the population of patients with complete MFI-20 data, and a comparison with those patients with missing MFI-20 data, are given in *Appendix 8*.

Figure 9 illustrates the model estimates and 95% CIs at baseline, at 1 month and at 6 months post randomisation for each of the five scales, split by treatment group. The estimated differences between the treatment groups in scores at baseline were as follows: general fatigue –0.2517 (95% CI –0.5965 to 1.0999; $p = 0.5603$), physical fatigue 0.3964 (95% CI –0.4404 to 1.2332; $p = 0.3527$), reduced activity –0.1634 (95% CI –0.9777 to 0.6510; $p = 0.6938$), reduced motivation –0.03917 (95% CI –0.7324 to 0.6540; $p = 0.9117$) and mental fatigue 0.1374 (95% CI –0.6626 to 0.9374; $p = 0.7360$). All of these differences are small and none are statistically significant. Furthermore, this lack of a notable difference between the groups persists over time, as seen in *Figure 9* and in the non–significant interaction terms in the model; further details of the models are given in *Appendix 8*.

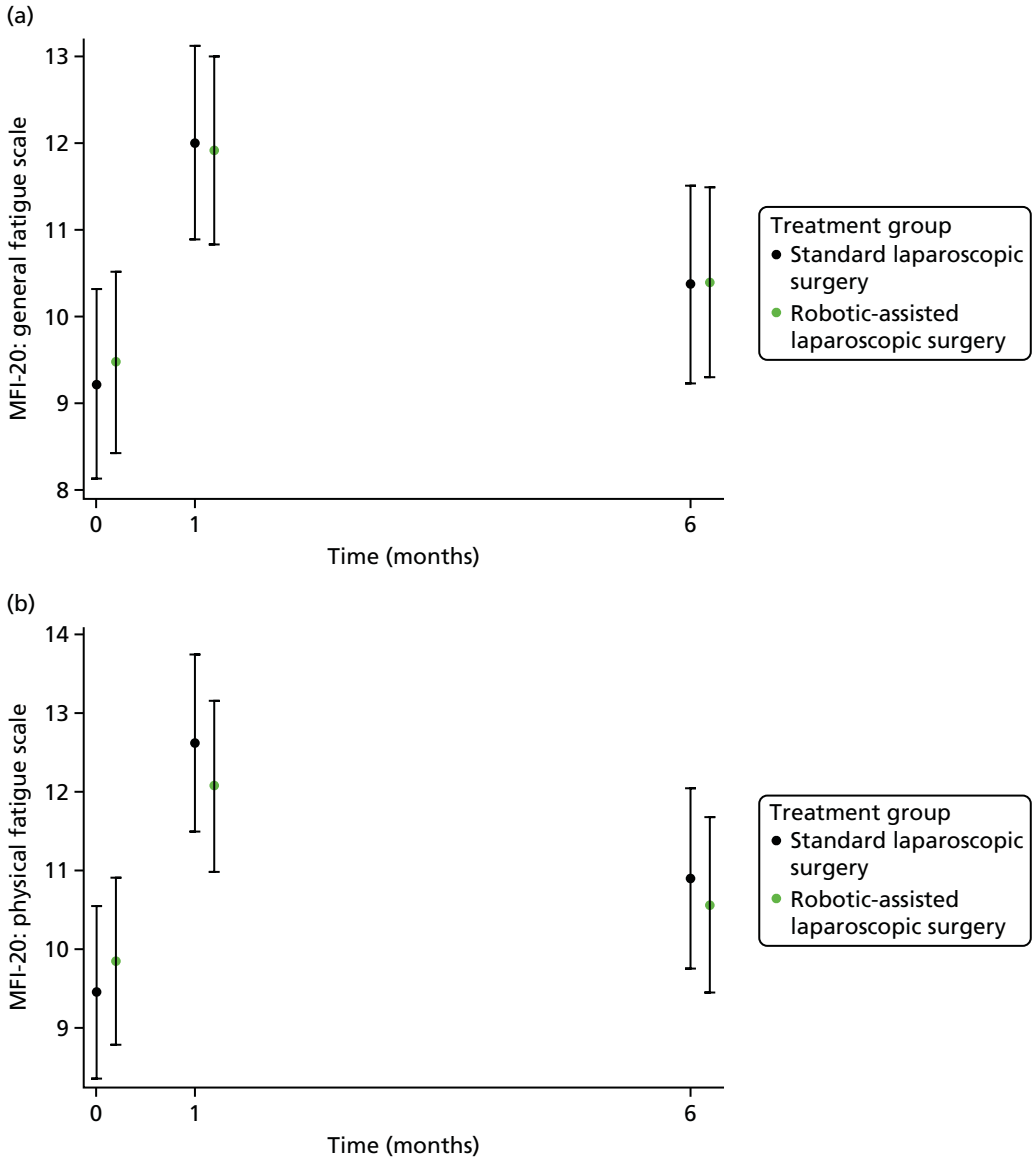


FIGURE 9 Adjusted estimates and 95% CIs of mean values of each of the five scales of the MFI-20 at baseline, at 1 month and at 6 months post randomisation, by treatment group. (a) General fatigue; (b) physical fatigue; (c) reduced activity; (d) reduced motivation; and (e) mental fatigue. (*continued*)

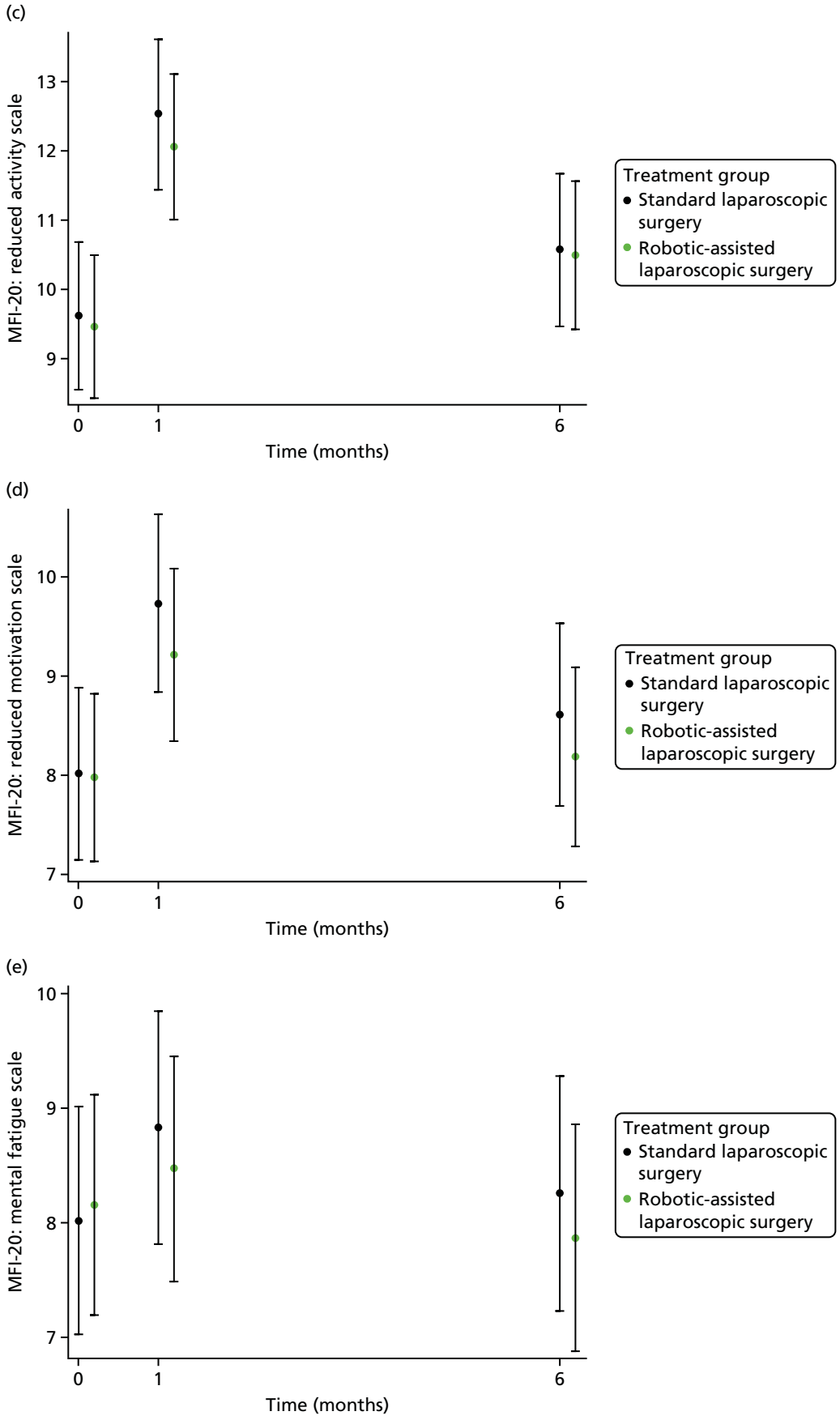


FIGURE 9 Adjusted estimates and 95% CIs of mean values of each of the five scales of the MFI-20 at baseline, at 1 month and at 6 months post randomisation, by treatment group. (a) General fatigue; (b) physical fatigue; (c) reduced activity; (d) reduced motivation; and (e) mental fatigue.

Plane of surgery

A total of 456 out of 471 (96.8%) patients had a returned pathology report with data for the mesorectal plane assessment. There were 351 out of 456 (77.0%) patients' specimens graded as mesorectal plane in the pathology report, 178 out of 233 (76.4%) in the laparoscopic group and 173 out of 223 (77.6%) in the robotic group (unadjusted risk difference 1.2%, 95% CI –6.5% to 8.9%). *Table 42* presents the crude summary of plane of resection (mesorectum) between the treatment groups. *Tables 43* and *44* present the

TABLE 42 Observed planes of resection (mesorectum), by treatment group

Mesorectum plane	Treatment group, *n* (%)		Total (*N* = 471), *n* (%)
	Standard laparoscopic surgery (*N* = 234)	Robotic-assisted laparoscopic surgery (*N* = 237)	
Mesorectal fascial plane	173 (73.9)	178 (75.1)	351 (74.5)
Intramesorectal plane	38 (16.2)	33 (13.9)	71 (15.1)
Muscularis propria plane	12 (5.1)	22 (9.3)	34 (7.2)
Missing	11 (4.7)	4 (1.7)	15 (3.2)

TABLE 43 Mesorectal plane of surgery: adjusted estimates of ORs and 95% CIs from random intercept model

Effect: comparator group (vs. reference group)	Group [number of conversions/number of patients (%)]		Risk difference (unadjusted 95% CI)	OR (adjusted)	95% CI for OR (adjusted)	*p*-value
	Reference	Comparator				
Treatment: robotic surgery (vs. laparoscopic)	173/223 (77.6)	178/233 (76.4)	1.2 (–6.5 to 8.9)	0.943	0.565 to 1.572	0.8211
Sex: male (vs. female)	122/149 (81.9)	229/307 (74.6)	7.3 (–0.6 to 15.2)	0.729	0.411 to 1.295	0.2808
BMI class: overweight (vs. underweight/normal)	142/177 (80.2)	134/175 (76.6)	3.7 (–4.9 to 12.3)	0.851	0.458 to 1.580	0.6086
BMI class: obese (vs. underweight/normal)	142/177 (80.2)	75/104 (72.1)	8.1 (–2.3 to 18.5)	0.905	0.457 to 1.795	0.7757
Previous radiotherapy or chemoradiotherapy: yes (vs. no)	197/256 (77.0)	154/200 (77.0)	–0.1 (–7.8 to 7.7)	0.796	0.435 to 1.454	0.4569
Intended procedure: HAR (vs. LAR)	251/308 (81.5)	55/68 (80.9)	0.6 (–9.7 to 10.9)	0.901	0.411 to 1.977	0.7943
Intended procedure: APR (vs. LAR)	251/308 (81.5)	45/80 (56.3)	25.2 (13.5 to 37.0)	0.358	0.185 to 0.694	0.0024

TABLE 44 Mesorectal plane of surgery: estimate of the variance component from random intercept model

Effect	Variance component	
	Estimate	Standard error
Operating surgeon (random effect)	2.236	1.021

multilevel logistic regression model. There was no significant difference of the odds of 'mesorectal plane' between the groups (adjusted OR 0.943, 95% CI 0.565 to 1.572; $p = 0.821$). Patients undergoing APR have notably lower odds of a mesorectal plane grading [adjusted OR (vs. LAR) 0.358, 95% CI 0.185 to 0.694; $p = 0.0024$].

Disease-free survival

A recurrence was observed in 73 out of 471 (15.5%) patients, 38 out of 234 (16.2%) in the laparoscopic group and 35 out of 237 (14.8%) in the robotic group.

The date of recurrence was defined as the date of the relevant assessment (i.e. clinical, radiological and pathological) that first detected the recurrence.

Kaplan–Meier estimates of DFS in each treatment group are presented in *Figure 10*.

Tables 45 and *46* present the estimated HRs and corresponding 95% CIs and Wald test *p*-values from the shared frailty model. There is no statistically significant difference between the treatment groups. The estimated adjusted HR suggests that a patient undergoing robotic surgery is 1.030 (95% CI 0.713 to 1.489; $p = 0.874$) times more likely to experience a recurrence or a new primary cancer or to die than a patient undergoing laparoscopic surgery, all else being equal.

There appears to be a substantial difference in DFS between patients undergoing different types of operation. All else being equal, those patients undergoing an APR are most likely to have a recurrence, a new primary cancer or to die (adjusted HR vs. LAR: 1.602, 95% CI 1.035 to 2.479; $p = 0.034$), and those patients undergoing HAR are least likely to have a recurrence, a new primary cancer or to die (adjusted HR vs. LAR 0.421, 95% CI 0.191 to 0.925; $p = 0.031$).

Subgroup analyses

None of the prespecified subgroup analyses yielded meaningful evidence of an interaction between treatment effect and subgroup, or evidence of a treatment effect within any individual subgroup. Given the clear (main) effect of sex on DFS, and the clinical plausibility of a potential difference of treatment effect by sex, an ad hoc sex subgroup analysis was performed. Similarly, this subgroup analysis showed no evidence of a subgroup by treatment interaction and no significant treatment effect within either subgroup.

Further details are given in *Appendix 9*.

Overall survival

Death was observed for 46 out of 471 (9.8%) patients, 23 out of 234 (9.8%) in the laparoscopic group and 23 out of 237 (9.7%) in the robotic group.

Kaplan–Meier estimates of OS in each treatment group are presented in *Figure 11*.

Tables 47 and *48* present the estimated HRs and corresponding 95% CIs and Wald test *p*-values from the shared frailty model. There is not a statistically significant difference between the treatment groups. The estimated adjusted HR suggests that a patient undergoing robotic surgery is 0.945 (95% CI 0.530 to 1.686; $p = 0.848$) times more likely to die than a patient undergoing laparoscopic surgery, all else being equal.

There appears to be a notable difference in the risk of death between males and females. All else being equal, the probability of death in males is 2.187 (95% CI 1.017 to 4.700; $p = 0.045$) times greater than in females.

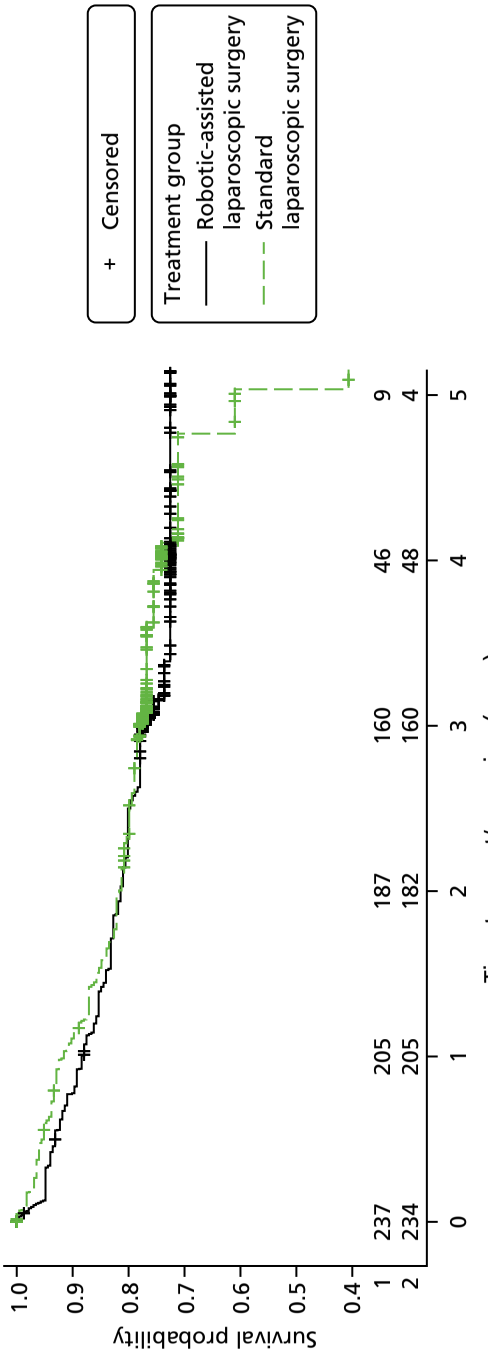


FIGURE 10 Kaplan–Meier plot of disease-free survival, by treatment group. Product-limit survival estimates with number of patients at risk. Note that the *y*-axis is truncated to 0.4–1.0.

TABLE 45 Disease-free survival: adjusted estimates of HRs and 95% CIs from random shared frailty model

Parameter	HR	95% CI	*p*-value
Treatment allocation: robotic (vs. laparoscopic)	1.030	0.713 to 1.489	0.8736
Sex: male (vs. female)	1.487	0.973 to 2.272	0.0665
Neo-adjuvant therapy: yes (vs. no)	1.259	0.857 to 1.848	0.2401
BMI classification: obese (vs. underweight/normal)	0.875	0.523 to 1.462	0.6095
BMI classification: overweight (vs. underweight/normal)	1.274	0.840 to 1.934	0.2542
Intended procedure: APR (vs. LAR)	1.602	1.035 to 2.479	0.0344
Intended procedure: HAR (vs. LAR)	0.421	0.191 to 0.925	0.0313

TABLE 46 Disease-free survival: estimate of the variance component from random intercept model

	Variance component	
Effect	Estimate	Standard error
Operating surgeon (random effect)	0.0271	0.0745

Subgroup analyses

None of the prespecified subgroup analyses yielded meaningful evidence of an interaction between treatment effect and subgroup, or evidence of a treatment effect within any individual subgroup. Given the clear (main) effect of sex on OS, and the clinical plausibility of a potential difference of treatment effect by sex, an ad-hoc sex subgroup analysis was performed. Similarly, this subgroup analysis showed no evidence of a subgroup by treatment interaction and no significant treatment effect within either subgroup.

Further details are given in *Appendix 9*.

Health economics

The results of the primary analysis are reported in *Table 49*. There are very similar QoL figures across the two groups, with a difference of 0.014 QALYs across the first 6 months of treatment. Across both groups, the pattern is of baseline EQ-5D utilities being highest pre-surgery (0.810 vs. 0.828) and noticeably lower at 30 days (0.680 vs. 0.700). The utilities are much closer to their pre-surgery values by 6 months (0.774 vs. 0.798).

There is an overall cost difference of £980 across the two groups of the trial over the first 6 months of treatment. Across the different categories of costs in the model, robotic surgery is around £1085 more expensive than laparoscopic surgery. The main drivers of the higher operative costs for robotic surgery are (1) the duration of surgery (357 minutes and 408 minutes, respectively), which has a knock-on effect on the cost of theatre time and staff, and (2) the use of surgical instruments. There is very little difference in the number of staff in attendance (mean number of assistants 1.7 and 1.63). As more stomas are formed with laparoscopic surgery, the mean costs of both reversal surgeries (£585 vs. £481; £104) and stoma supplies (£547 vs. £486; £61) are slightly higher for this form of surgery. Mean costs are marginally higher for the robotic group in terms of medications and other health professional contacts (e.g. outpatients, GPs, nurses, etc.), although these differences are small (£590 vs. £656; –£66).

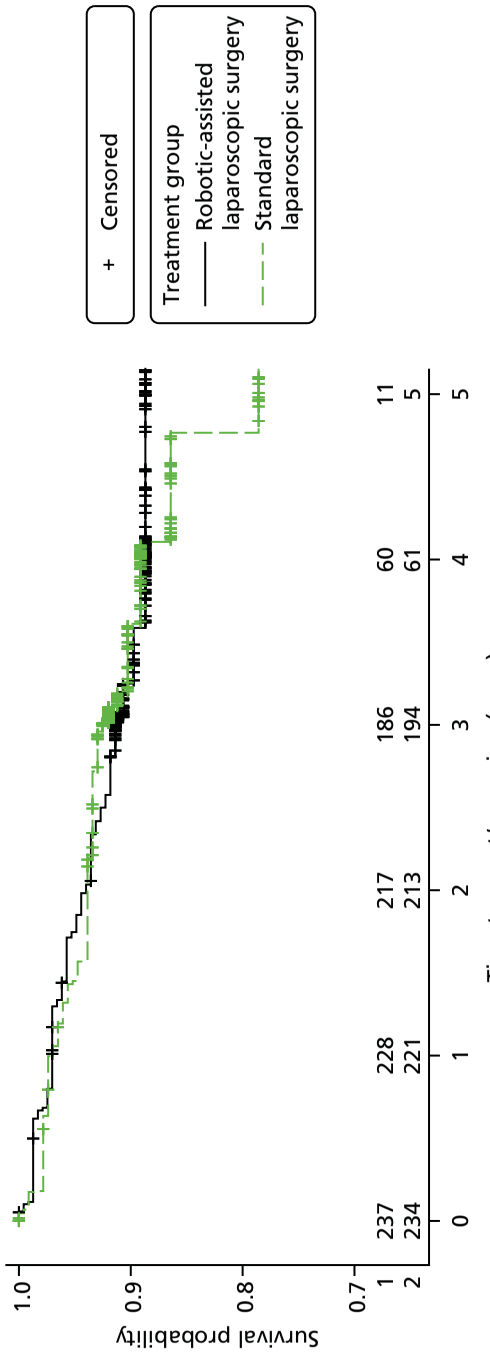


FIGURE 11 Kaplan–Meier plot of OS, by treatment group. Product-limit survival estimates with number of patients at risk. Note that the *y*-axis is truncated to 0.7–1.0.

TABLE 47 Overall survival: adjusted estimates of HRs and 95% CIs from random shared frailty model

Parameter	HR	95% CI	*p*-value
Treatment allocation: robotic (vs. laparoscopic)	0.945	0.530 to 1.686	0.8483
Sex: male (vs. female)	2.187	1.017 to 4.700	0.0450
Neo-adjuvant therapy: yes (vs. no)	1.380	0.756 to 2.522	0.2945
BMI classification: obese (vs. underweight/normal)	0.577	0.258 to 1.290	0.1804
BMI classification: overweight (vs. underweight/normal)	0.652	0.339 to 1.255	0.2007
Intended procedure: APR (vs. LAR)	1.442	0.741 to 2.804	0.2815
Intended procedure: HAR (vs. LAR)	0.520	0.155 to 1.739	0.2881

TABLE 48 Overall survival: estimate of the variance component from random intercept model

	Variance component	
Effect	Estimate	Standard error
Operating surgeon (random effect)	< 0.001	0.1649

TABLE 49 Results of primary scenario: imputed data for all UK and US patients ($n = 190$)

	Treatment group							
	Laparoscopic surgery ($n = 95$)				Robotic surgery ($n = 95$)			
Parameter	Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall figures								
QALY	0.364	0.097	–0.029	0.499	0.378	0.089	0.023	0.499
Total costs (£)	10,874	2676	6245	26,969	11,853	2940	7330	22,872
Health and QoL								
Baseline	0.810	0.192	0.085	1.000	0.828	0.181	0.131	1.000
30 days	0.680	0.229	–0.458	1.000	0.700	0.244	–0.116	1.000
6 months	0.774	0.221	–0.287	1.000	0.798	0.201	–0.248	1.000
Stoma formed (%)	91	28	0	100	81	39	0	100
Days with stoma	146.35	58.71	–	182.00	132.92	69.30	–	182.00
Cost breakdown (£)								
Total costs	10,874	2676	6245	26,969	11,853	2940	7330	22,872
Initial surgery	8423	1443	5249	12,466	9508	1219	6591	14,328
Stoma reversals	585	805	–	1691	481	763	–	1691
Stoma supplies	547	292	–	1057	486	351	–	1057
Other surgery	729	2245	–	15,242	723	2233	–	7621
All other costs	590	494	–	5178	656	468	–	3606

Across this scenario, the expected costs for robotic surgery are higher (£980) and robotic surgery provides marginally more health (QALY gain 0.014). These figures combine for an estimated incremental cost-effectiveness ratio (ICER) of £69,837 per QALY. This figure is well in excess of a standard threshold of £20,000–30,000 per QALY and suggests that robotic surgery, even when not including the sizeable cost of the robot, is not cost-effective. The implication for this is that, even if this means that the robot is lying idle, it is less cost-effective to use the robot for rectal resections.

The cost-effectiveness acceptability curve (*Figure 12*) for this scenario provides an estimate of the likelihood that this is cost-effective at a cost-effectiveness threshold of £20,000 or £30,000 per QALY. Across the cases calculated, robotic surgery is cost-effective only 10% of the time when the threshold is taken to be £20,000 per QALY and only 20% of the time when the threshold is £30,000 per QALY.

Across the other scenarios (*Table 50*), there was no strong suggestion of cost-effectiveness. For those patients who intended to receive low anterior surgery at randomisation, the QoL benefit for robotic surgery is very close to zero; there is a benefit of 0.002 QALYs and the ICER for this scenario is nearly £700,000 per QALY. Although there is still some small uncertainty about cost-effectiveness (there is a 10% chance that robotic surgery is cost-effective at £30,000 per QALY), there is no clear case for cost-effectiveness in this subgroup. Although the corresponding cost-effectiveness would be higher among the other intended surgical groups at baseline, there is no clinical reason to justify the consideration of these subgroups.

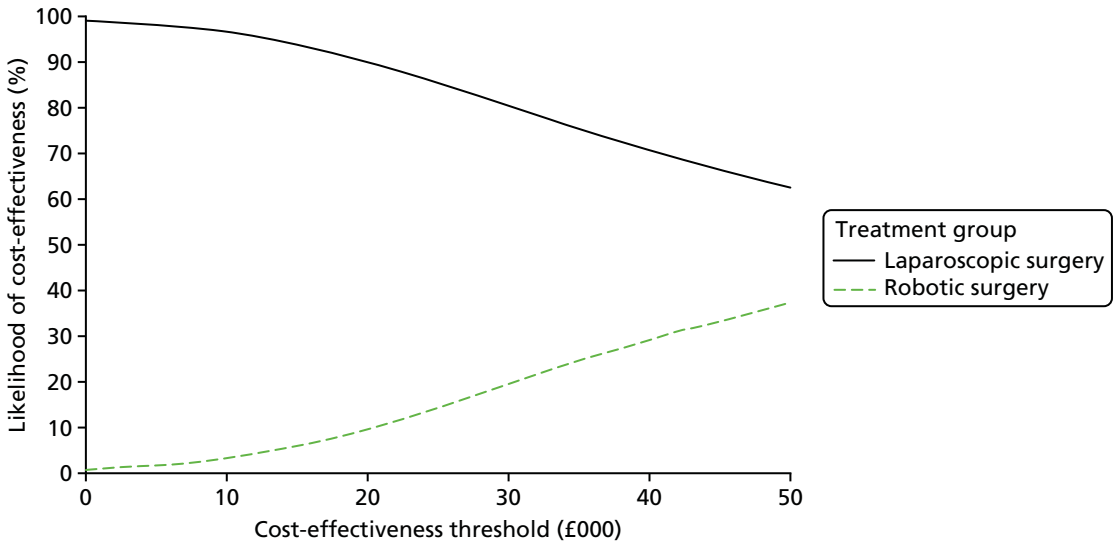


FIGURE 12 Cost-effectiveness acceptability curve for the primary analysis.

TABLE 50 Results of the secondary scenarios: complete data for all patients ($n = 97$), imputed data for UK and US patients intended to receive low anterior surgery ($n = 135$), imputed data for all observations ($n = 471$)

Scenario	Incremental cost (£)	Incremental QALY	Incremental cost-effectiveness ratio (£ per QALY)	Likelihood of cost-effectiveness (%)	
				£20,000/QALY	£30,000/QALY
Baseline (UK, USA) ($n = 190$)	980	0.014	69,837	9.8	19.5
Low anterior surgery (UK, USA) ($n = 135$)	1096	0.002	698,000	5.9	9.8
Complete cases ($n = 97$)	1241	0.028	43,844	16.1	29.7
All patients ($n = 471$)	743	0.004	172,943	2.7	5.6

Across the two remaining sensitivity analyses, the ICER varied greatly according to the group of patients considered. Only 51% of UK and US patients in the subgroup had complete data on all costing categories and all QoL values. Data on QoL were more complete, with 133 patients providing enough information to allow QALYs to be computed. In contrast, only 99 patients provided full cost data. (Two patients provided cost data but did not provide full data for QALYs.)

In the subgroup with complete data, the ICER was lower than the baseline value, although still above the standard values for thresholds, at around £44,000 per QALY. It is worth noting that, here, the chance of cost-effectiveness at £20,000 and £30,000 per QALY is still only 16% and 30%, respectively.

However, the complete-case data appear slightly misleading. Across the 133 patients observed to have sufficient data to calculate QALYs, there is a 0.017 benefit (vs. 0.014 in the base case and 0.028 in the complete case). Across the 99 patients with full cost data, the cost difference observed (£1169) is much more similar to those in both the complete case (£1241) and the base case (£980). This suggests that the base-case results are broadly consistent with the overall data set and more representative than the complete-case example.

In contrast, the imputation across the pan-world trial (as opposed to the USA and the UK) produces quite different numbers. In this case, the benefits estimated are much lower than in the baseline case, leading to an ICER of > £170,000 per QALY.

Chapter 4 Discussion

In this study, which, to our knowledge, is the largest randomised trial of robotic-assisted surgery for patients with rectal adenocarcinoma suitable for curative resection, there were no statistically significant differences in the rates of conversion to laparotomy for robotic-assisted laparoscopic surgery compared with conventional laparoscopic surgery (8.1% vs. 12.2%, respectively), and there were no statistically significant differences in resection margin positivity, complication rates or QoL at 6 months. There is insufficient evidence to conclude that robotic-assisted laparoscopic surgery, compared with conventional laparoscopic surgery, reduces the risk of conversion to laparotomy when performed by surgeons of varying experience with robotic surgery.

The primary outcome measure was conversion to open surgery, based on the hypothesis that the technological advantages of the robot should facilitate rectal cancer resection and avoid the need to convert to an open operation. The sample size calculations were based on best available evidence in 2009 and included the largest randomised clinical trial of laparoscopic rectal cancer surgery, the MRC CLASICC trial, which reported a 34% conversion rate to open surgery.[5] A 25% conversion rate from laparoscopic to open surgery was assumed, giving a sample size of 400 patients to demonstrate a 50% relative reduction in conversion rate with robotic surgery. The actual overall conversion rate turned out to be much lower, at 10.1%. A similar reduction in conversion rates with time has been reported in other laparoscopic rectal cancer trials: COLOUR II 16%,[7] ACOSOG Z6501 11%,[11] ALaCaRT 9%.[10] In our trial, a difference in conversion rate between laparoscopic (12.2%) and robotic (8.1%) surgery was not statistically significant. The higher overall conversion in patients undergoing LAR than those undergoing APR probably reflects the fact that the majority of the oncological component of the operation is performed from the perineum in the abdominoperineal approach and is less affected by the laparoscopic approach. Similarly, the higher overall conversion rates for males than females, and obese patients than underweight/normal patients probably reflects the increasing technical difficulty of carrying out these procedures on these patients.

The sensitivity analysis exploring learning effects suggested a potential benefit of robotic surgery when performed by surgeons with substantial prior robotic experience, regardless of their level of laparoscopic experience. This suggests that the majority of participating surgeons were experts in laparoscopic surgery, but still in their learning phases for robotic surgery, and that at the higher end of the spectrum of experience in robotic surgery there is evidence of a benefit (in terms of conversion rate) over standard laparoscopic surgery.

In almost all of the subgroup analyses, there were insufficient numbers of patients to produce statistically meaningful comparisons between the groups regarding the need to convert to an open operation. However, differences were apparent in the conversion rates for the laparoscopic and robotic groups in males, with robotic surgery appearing to offer a benefit. Although results yielded by a subgroup analysis must be interpreted with caution, the moderate evidence of interaction between sex and treatment effect, the evidence of a difference between treatments in males, and the clinical plausibility of the robot facilitating dissections in the narrower male pelvis with more operator-controlled retraction, better optics and instrument precision, all warrant further investigation into the potential benefit of robotic surgery in this subgroup of technically challenging patients.

The experience of the participating surgeons was also evident in the low CRM+ rate (overall 5.7%), which was lower than in previous laparoscopic rectal cancer trials: COLOR II 10%, ACOSOG Z6501 12.1%, ALaCaRT 7%. Pathological grading of the plane of surgery showed a good standard, with mesorectal plane surgery observed in 76% overall. This is lower than that reported in COLOR II (88%) and ALaCaRT (87%), but similar to ACOSOG Z5601 (72.9%), and is probably because of the recognised variation in reporting between pathologists. In our trial, reporting of pathological plane of surgery was standardised to the method described by Nagtegaal and Quirke.[23] Despite this, there was considerable variation between

local reporting of the plane of surgery and that reported on central review, illustrating the subjectivity in assessment and the need to interpret the results of other trials with caution.

The complication rates following laparoscopic and robotic surgery were similar and there were no safety issues attributable to the robotic system. Overall 30-day mortality was low at 0.9%, in keeping with the results of meta-analyses.[16,17] The leading causes of intraoperative morbidity were iatrogenic damage to an organ/structure and significant haemorrhage. In contrast to other studies, haemorrhage was not more frequently associated with robotic surgery.[13] Rectal cancer surgery is a high-risk intervention with 32.4% of patients experiencing a complication within 30 days and, after that, 15.5% of the patients experiencing complications between 30 days and 6 months. Complications related to the gastrointestinal tract, including anastomotic leak, were not surprisingly the most common cause of morbidity at both time points. Despite advances in operative technique and perioperative care, the high morbidity associated with rectal cancer surgery has not declined over the past few decades. There is a need for more sophisticated preoperative stratification systems to enable surgeons to predict patients at risk of complications, to stratify surgery accordingly and to put pathways in place to prevent complications from occurring and minimise the consequences should they occur.

Previous studies have shown that both laparoscopic and robotic rectal cancer surgery can result in bladder and sexual dysfunction, but suggest that recovery is earlier following robotic surgery.[19] The analysis of bladder and sexual function present in the ROLARR study was undertaken at the same time points and using the same research questionnaires as the previously reported studies. The findings do not support the published data; there was no significant benefit to postoperative bladder or sexual function from the use of the robot. Notably, there was little change in any of the I-PSS, IIEF and FSFI scores between baseline and 6 months, suggesting that the ROLARR surgeons were accomplished in autonomic nerve preservation and that clinically relevant bladder and sexual dysfunction were an infrequent event.

The case for laparoscopic surgery, rather than open surgery, for colorectal cancer is well established with proven benefits in terms of a shorter stay in hospital and a faster return to normal function. However, it has been difficult to demonstrate an advantage for the laparoscopic technique in terms of improvements in QoL. Similarly, in the ROLARR study, we have not been able to show an obvious advantage for robotic surgery over laparoscopic surgery in terms of QoL. A small benefit for robotic surgery was seen in QoL using the EQ-5D score, but no advantage over laparoscopic surgery was seen in either the physical components or the mental components in the SF-36v2 QoL analysis. Similarly, there was no difference in recovery following robotic surgery or laparoscopic surgery, as measured by the MFI-20 questionnaire. This is probably not surprising given that the extent of the surgery performed is not influenced by the surgical approach, with the robot serving as an alternative tool to enable a laparoscopic operation to be performed. Similar trends were noted using both the SF-36v2 questionnaire and the MFI-20 questionnaire, with the predictable deterioration in QoL at the 1-month follow-up time point and a period of 6 months being required before QoL returned towards baseline. This has implications for patients being scheduled for rectal cancer surgery who should be counselled about a prolonged recovery period, which extends far beyond the immediate postoperative period. Women, in particular, appear to be more likely to experience a protracted recovery than males.

Results from the health economics analysis suggest that robotic rectal cancer surgery is unlikely to be cost saving. The mean difference per operation, excluding the acquisition and maintenance costs, was £980 and was driven by longer operating theatre time and increased costs for robotic instruments. This concurs with previously reported data, which consistently report longer operating times associated with the robot.[21] When considering robotic surgery as a whole, rather than just rectal cancer surgery, one has to consider the cost of the purchase and maintenance of the system, the operational life and the total utilisation of the robot per year for all robotic procedures. Estimates of acquisition costs in 2017 varied between £0.43M and £1.8M with maintenance costs between £0.06M and £0.12M per year.[45] Assuming a mid-point acquisition cost of £1.12M and a mid-point maintenance cost of £0.896M per year, with an operational life/amortisation period of 7 years,[46,47] the total cost of a robot would be around £1.738M. Estimates for

total utilisation of the robot per year in 2017 varied between 819,000 and 843,000 procedures across 3919 installed systems, or 1505 procedures per robot over 7 years.[45] This gives the total fixed costs of around £1155 per procedure, in addition to the variable costs. Alternatively stated, the net benefits (excluding fixed costs) of any robotic procedure included in a set of cost-effective procedures needs to be positive, and the whole set of cost-effective procedures needs to have an average net benefit of at least £1155. On average, all robot procedures combined must exceed this figure, with all procedures making at least some positive contribution. On the basis of the evidence presented here, robotic rectal cancer surgery does not appear to provide a positive contribution and does not appear to be justified given the extra costs and the equivalency of clinical outcomes.

Analysis of the long-term outcomes, local recurrence, and DFS and OS, failed to show a difference between the treatment groups and conformed to the recognised patterns seen following rectal cancer surgery. Interestingly, local recurrence was more common in males than in females, which might reflect the increased technical difficulty in operating in the narrower male pelvis, although there was no benefit from robotic surgery compared with laparoscopic surgery in male patients. Although there was no difference between the treatment groups in DFS, those patients undergoing APR were most at risk of disease recurrence, and those undergoing HAR were least at risk. Males fared worse in terms of OS, which might be related to the higher risk of local recurrence, but will also be influenced by the generally shorter life expectancy for males, with no difference whether or not the operation was performed with the robot or by standard laparoscopy.

Limitations

The ROLARR study had several limitations. The much lower than anticipated rate of conversion to laparotomy limits the ability to provide conclusive evidence about our primary question of how robotic surgery compares with conventional laparoscopic surgery in terms of the odds of conversion. However, the fact that no statistically significant differences between the treatment groups were seen in any of the end points does suggest that robotic surgery, when performed by surgeons with varying robotic experience, does not confer a clinically important benefit over laparoscopic surgery in the short term.

No blinding to treatment allocation was incorporated into this trial. Our primary end point and the measure of mortality will certainly be unaffected by this, as an objective end point. However, there is the potential for end points that are not completely objective to have been affected. In our pathology end points, including CRM+, we have guarded against this by carrying out a blinded central review of pathology assessments.

Despite enforcing a mandatory minimum experience level for surgeon participation, operations in this trial were performed, on average, by a surgeon considered to be an expert in conventional laparoscopic surgery and who may still have been in their learning phase for robotic surgery. The prespecified sensitivity analysis of learning effects addresses this, by extending the primary analysis model to analyse the interaction between operating surgeon experience and the treatment effect.

The primary analysis adjusted only for stratification factors (including operating surgeon) and thus it did not include an adjustment for treating site in particular. A (prespecified) adjustment for treating site was considered in a sensitivity analysis, but model estimation issues were caused by the small sizes of the resulting strata, resulting in no meaningful output.

Chapter 5 Conclusions

The ROLARR study provides the first rigorous evaluation of robotic-assisted surgery compared with laparoscopic surgery for rectal cancer. It has failed to show an advantage for the robotic technique, although interesting trends have been noted. In particular, the trend to reduced conversions in males, and perhaps those patients undergoing LAR, deserves further investigation. A registry is currently being designed under the auspices of the European Society of Coloproctology that should enable data on several hundred robotic rectal cancer operations to be collected in a relatively short time frame and might provide further insight into the trends observed in the ROLARR study.

The health economic evaluation performed in the ROLARR study concludes that robotic rectal cancer surgery is not cost-effective compared with laparoscopic surgery. Although it is tempting to generalise this to wider surgical practice and the health-care provision of future robotic services, it should be borne in mind that the ROLARR study investigated only a single robotic system, the da Vinci surgical robot, which was the only system that was commercially available at the time. There have subsequently been rapid developments in other surgical robotic platforms, with several expected to come to market within the next couple of years. Future systems promise to be more competitive in terms of costs, with per-procedure costs challenging those of laparoscopic surgery. The health economic data from the ROLARR study will be beneficial to commercial companies developing robotic systems, in particular highlighting the need to bring the cost of robotic instruments down in order to be competitive with laparoscopic surgery.

Any judgement about the future of robotic surgery based on the ROLARR study should be tempered with future developments borne in mind. The situation is further complicated by the recent debate about the benefits of laparoscopic surgery that has been stirred by the recent publication of two large randomised trials comparing laparoscopic with open surgery for rectal cancer: the ALaCarte and ACOSOG trials. Both these studies failed to show the non-inferiority of laparoscopic surgery compared with open surgery for rectal cancer in terms of a short-term composite pathological outcome. It is therefore not clear whether or not any future analysis of robotic rectal cancer surgery should include an assessment of open surgery as well as laparoscopic techniques. In the UK's NHS, the adoption of laparoscopic rectal cancer surgery is probably too advanced to countenance reverting back to open surgery, unless there is hard long-term evidence to suggest otherwise.

Acknowledgements

We are indebted to the patients who participated in this trial. We would also like to thank the TSC (Michael Parker, Lynn Faulds Wood, Gareth Griffiths and Hamish Macdonald), the DMEC (Azad Najmaldin, Nigel Scott and Craig Ramsay), additional members of the Trial Management Group (Jennifer Barrie, Sarah Perry, Christopher Garbett, Helen Howard and Fiona Collinson), and the CTRU Project Team for their important contributions.

Patient and public involvement

Patients were involved in all stages of the design and delivery of the ROLARR trial. Christopher Garbett provided valuable input into all aspects of the study and served on the Trial Management Committee, and Lynn Faulds Wood served on the Trail Steering Committee.

The following institutions and surgeons participated in the trial.

Aarhus University Hospital (Denmark): Soren Laurberg, Niels Thomassen and Lene H Iversen; Aria Health (USA): Luca Giordano; Augusta Kranken Anstalt (Germany): Benno Mann; Azienda Ospedaliera SS Antonio e Biagio (Italy): Giuseppe Spinoglio; Barnet & Chase Farm Hospitals NHS Trust (UK): Daren Francis; Centre Oscar Lambret (France): Mehrdad Jafari; Christie Hospital (UK): Chelliah Selvasekar; Duke University (USA): Linda Farkas and Michael Hopkins; European Institute of Oncology: Unit of Integrated Abdominal Surgery (Italy) – Fabrizio Luca and Roberto Biffi; European Institute of Oncology: Unit of Minimally Invasive Surgery (Italy) – Paolo Pietro Bianchi and Roberto Biffi; Frimley Park (UK): Henry Tilney and Mark Gudgeon; Gangnam Severance Hospital (South Korea): Kang Young Lee; Hospital Herlev, Copenhagen University (Denmark): Jacob Rosenburg, Henrik Loft Jakobsen, Mads Bundgaard, Jens Ravn Eriksen, Jesper Olsen and Thomas Bent Harvald; Jackson South Community Hospital (USA): Gustavo Plasencia and Henry J Lujan; John Muir Medical Center (USA): Samuel Oommen; Leeds Teaching Hospital Trust (UK): David Jayne, Richard Baker and Julian Hance; National University Hospital (Singapore): Charles Tsang; Oulu University Hospital (Finland): Tero Rautio; Peter MacCallum Cancer Centre (Australia): Andrew Craig Lynch; Roskilde Hospital (Denmark): Per Jess, Michael Seiersen, Ole Roikjaer and Steffen Brisling; Royal Surrey County Hospital (UK): Tim Rockall; San Pio X Hospital (Italy): Jacques Megevand; Torbay Hospital (UK): Stephen Mitchell; University of California, Irvine Medical Center (USA): Alessio Pigazzi, Joseph C. Carmichael and Steven Mills; University of Pisa (Italy): Luca Morelli; Washington University in St. Louis School of Medicine (USA): Elisa Birnbaum and Paul Wise.

The following people were local pathologists for the trial:

Aarhus University Hospital (Denmark): Rikke Hagemann-Madsen and Katrine Stribolt; Aria Health (USA): Peter Farano, Thomas Rizzo Jr and Behnaz Toorkey; Augusta Kranken Anstalt (Germany): Stathis Philippou and Konrad Morgenroth; Azienda Ospedaliera SS Antonio e Biagio (Italy): Narciso Mariani, Paola Barbieri and Paola Re; Barnet & Chase Farm Hospitals NHS Trust (UK): Khurram Chaudhary and Anupam Joshi; Centre Oscar Lambret (France): Charles André; Christie Hospital (UK): Bipasha Chakrabarty, Guy Betts, Igor Racu-Amoasii and Khalifa Sawalem; Duke University (USA): Carol Filomena; European Institute of Oncology: Unit of Integrated Abdominal Pelvic Surgery (Italy) – Angelica Sonzogni and Luca Bottiglieri; European Institute of Oncology: Unit of Minimally Invasive Surgery (Italy) – Angelica Sonzogni and Luca Bottiglieri; Frimley Park (UK): Salome Beeslaar and George Kousparos; Gangnam Severance Hospital (South Korea): Soon Won Hong; Hospital Herlev, Copenhagen University (Denmark): Jill Levin Langhoff and Peter Ingeholm; Jackson South Community Hospital (USA): Rebeca Porto; John Muir Medical Center (USA): Barry Latner; Leeds Teaching Hospital Trust (UK): Padmini Prasad and Heike Grabsch; National University Hospital (Singapore): Teh Ming, Fredrik Petersson and Wang Shi; Oulu University Hospital (Finland): Markus Mäkinen and Tuomo Karttunen; Peter MacCallum Cancer Centre (Australia): Phillip Moss and Catherine Mitchell; Roskilde Hospital (Denmark): Peter Engel, Susanne Eiholm, Matteo Biagini and

Anne Mellon Mogensen; Royal Surrey County Hospital (UK): Izhar Bagwan; San Pio X Hospital (Italy): Claudio Clemente, Anna Maria Ferrari, Stefania Rao and Cristina Campidelli; Torbay Hospital (UK): Nick Ryley, Ian Buley and Maria Consuelo Garrido; University of California, Irvine Medical Center (USA): Mark Li-Cheng Wu, Roberta Edwards and Philip Carpenter; University of Pisa (Italy): Daniela Campani and Luca Emanuele Pollina; Washington University in St. Louis School of Medicine (USA): Ilke Nalbantoglu.

Contributions of authors

David Jayne (Professor of Surgery) contributed to the design of the study, patient recruitment, data interpretation and manuscript writing.

Alessio Pigazzi (Consultant Surgeon) contributed to the design of the study, patient recruitment, data interpretation and manuscript writing.

Helen Marshall (Senior Statistician) contributed to the study design and oversight, data analysis and interpretation, and manuscript writing.

Julie Croft (Senior Clinical Triallist) contributed to the study design, study coordination and manuscript writing.

Neil Corrigan (Senior Statistician) contributed to the study design and oversight, data analysis and interpretation, and manuscript writing.

Joanne Copeland (Clinical Triallist) contributed to the study coordination and manuscript writing.

Philip Quirke (Professor of Pathology) designed the pathology protocol, provided training to other centres, reviewed the pathology and interpreted the pathology data.

Nicholas West (Academic Consultant Histopathologist) contributed to the design of the pathology protocol, reviewed the pathology and interpreted the pathology data.

Richard Edlin (Health Economist) contributed to the design of the cost-effectiveness study, data analysis and manuscript writing.

Claire Hulme (Professor of Health Economics) contributed to the design of the cost-effectiveness study, data analysis and manuscript writing.

Julia Brown (Professor of Statistics and Director of the Leeds Clinical Trial and Research Unit) contributed to the study design and oversight, data analysis and interpretation, and manuscript writing. All authors reviewed the final manuscript.

Publications

Collinson FJ, Jayne DG, Pigazzi A, Tsang C, Barrie JM, Edlin R, *et al*. An international, multicentre, prospective, randomised, controlled, unblinded, parallel-group trial of robotic-assisted versus standard laparoscopic surgery for the curative treatment of rectal cancer. *Int J Colorectal Dis* 2012;**27**:233–41.

Jayne D, Pigazzi A, Marshall H, Croft J, Corrigan N, Copeland J, *et al*. Effect of robotic-assisted vs. conventional laparoscopic surgery on risk of conversion to open laparotomy among patients undergoing resection for rectal cancer: the ROLARR randomized clinical trial. *JAMA* 2017;**318**:1569–80.

Corrigan N, Marshall H, Croft J, Copeland J, Jayne D, Brown J. Exploring and adjusting for potential learning curve effects in ROLARR: a randomised controlled trial comparing robotic-assisted vs. standard laparoscopic surgery for rectal cancer resection. *Trials* 2018;**19**:339–49.

Data-sharing statement

All available data can be obtained by contacting the corresponding author.

Patient data

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease, develop new treatments, monitor safety, and plan NHS services. Patient data should be kept safe and secure, to protect everyone's privacy, and it's important that there are safeguards to make sure that it is stored and used responsibly. Everyone should be able to find out about how patient data are used. #datasaveslives You can find out more about the background to this citation here: https://understandingpatientdata.org.uk/data-citation.

References

1. MacFarlane JK, Ryall RD, Heald RJ. Mesorectal excision for rectal cancer. *Lancet* 1993;**341**:457–60. https://doi.org/10.1016/0140-6736(93)90207-W
2. Jayne DG, Guillou PJ, Thorpe H, Quirke P, Copeland J, Smith AM, *et al.* Randomized trial of laparoscopic-assisted resection of colorectal carcinoma: 3-year results of the UK MRC CLASICC Trial Group. *J Clin Oncol* 2007;**25**:3061–8. https://doi.org/10.1200/JCO.2006.09.7758
3. Nelson H, Sargent DJ, Wieand HS, Fleshman J, Anvari M, Stryker SJ, *et al.* A comparison of laparoscopically assisted and open colectomy for colon cancer. *N Engl J Med* 2004;**350**:2050–9. https://doi.org/10.1056/NEJMoa032651
4. Tang CL, Eu KW, Tai BC, Soh JG, MacHin D, Seow-Choen F. Randomized clinical trial of the effect of open versus laparoscopically assisted colectomy on systemic immunity in patients with colorectal cancer. *Br J Surg* 2001;**88**:801–7. https://doi.org/10.1046/j.1365-2168.2001.01781.x
5. Guillou PJ, Quirke P, Thorpe H, Walker J, Jayne DG, Smith AM, *et al.* Short-term endpoints of conventional versus laparoscopic-assisted surgery in patients with colorectal cancer (MRC CLASICC trial): multicentre, randomised controlled trial. *Lancet* 2005;**365**:1718–26. https://doi.org/10.1016/S0140-6736(05)66545-2
6. Jayne DG, Thorpe HC, Copeland J, Quirke P, Brown JM, Guillou PJ. Five-year follow-up of the Medical Research Council CLASICC trial of laparoscopically assisted versus open surgery for colorectal cancer. *Br J Surg* 2010;**97**:1638–45. https://doi.org/10.1002/bjs.7160
7. van der Pas MH, Haglind E, Cuesta MA, Fürst A, Lacy AM, Hop WC, Bonjer HJ, COlorectal cancer Laparoscopic or Open Resection II (COLOR II) Study Group. Laparoscopic versus open surgery for rectal cancer (COLOR II): short-term outcomes of a randomised, phase 3 trial. *Lancet Oncol* 2013;**14**:210–18. https://doi.org/10.1016/S1470-2045(13)70016-0
8. Bonjer HJ, Deijen CL, Abis GA, Cuesta MA, van der Pas MH, de Lange-de Klerk ES, *et al.* A randomized trial of laparoscopic versus open surgery for rectal cancer. *N Engl J Med* 2015;**372**:1324–32. https://doi.org/10.1056/NEJMoa1414882
9. Kang SB, Park JW, Jeong SY, Nam BH, Choi HS, Kim DW, *et al.* Open versus laparoscopic surgery for mid or low rectal cancer after neoadjuvant chemoradiotherapy (COREAN trial): short-term outcomes of an open-label randomised controlled trial. *Lancet Oncol* 2010;**11**:637–45. https://doi.org/10.1016/S1470-2045(10)70131-5
10. Stevenson AR, Solomon MJ, Lumley JW, Hewett P, Clouston AD, Gebski VJ, *et al.* Effect of laparoscopic-assisted resection vs open resection on pathological outcomes in rectal cancer: the ALaCaRT randomized clinical trial. *JAMA* 2015;**314**:1356–63. https://doi.org/10.1001/jama.2015.12009
11. Fleshman J, Branda M, Sargent DJ, Boller AM, George V, Abbas M, *et al.* Effect of laparoscopic-assisted resection vs open resection of stage II or III rectal cancer on pathologic outcomes: the ACOSOG Z6051 randomized clinical trial. *JAMA* 2015;**314**:1346–55. https://doi.org/10.1001/jama.2015.10529
12. Pigazzi A, Ellenhorn JD, Ballantyne GH, Paz IB. Robotic-assisted laparoscopic low anterior resection with total mesorectal excision for rectal cancer. *Surg Endosc* 2006;**20**:1521–5. https://doi.org/10.1007/s00464-005-0855-5
13. Hellan M, Anderson C, Ellenhorn JD, Paz B, Pigazzi A. Short-term outcomes after robotic-assisted total mesorectal excision for rectal cancer. *Ann Surg Oncol* 2007;**14**:3168–73. https://doi.org/10.1245/s10434-007-9544-z

14. Baik SH, Ko YT, Kang CM, Lee WJ, Kim NK, Sohn SK, *et al.* Robotic tumor-specific mesorectal excision of rectal cancer: short-term outcome of a pilot randomized trial. *Surg Endosc* 2008;**22**:1601–8. https://doi.org/10.1007/s00464-008-9752-z

15. Keller DS, Hashemi L, Lu M, Delaney CP. Short-term outcomes for robotic colorectal surgery by provider volume. *J Am Coll Surg* 2013;**217**:1063–9.e1. https://doi.org/10.1016/j.jamcollsurg.2013.07.390

16. Ortiz-Oshiro E, Sánchez-Egido I, Moreno-Sierra J, Pérez CF, Díaz JS, Fernández-Represa JÁ. Robotic assistance may reduce conversion to open in rectal carcinoma laparoscopic surgery: systematic review and meta-analysis. *Int J Med Robot* 2012;**8**:360–70. https://doi.org/10.1002/rcs.1426

17. Mirnezami AH, Mirnezami R, Venkatasubramaniam AK, Chandrakumaran K, Cecil TD, Moran BJ. Robotic colorectal surgery: hype or new hope? A systematic review of robotics in colorectal surgery. *Colorectal Dis* 2010;**12**:1084–93. https://doi.org/10.1111/j.1463-1318.2009.01999.x

18. Xiong B, Ma L, Huang W, Zhao Q, Cheng Y, Liu J. Robotic versus laparoscopic total mesorectal excision for rectal cancer: a meta-analysis of eight studies. *J Gastrointest Surg* 2015;**19**:516–26. https://doi.org/10.1007/s11605-014-2697-8

19. Kim JY, Kim NK, Lee KY, Hur H, Min BS, Kim JH. A comparative study of voiding and sexual function after total mesorectal excision with autonomic nerve preservation for rectal cancer: laparoscopic versus robotic surgery. *Ann Surg Oncol* 2012;**19**:2485–93. https://doi.org/10.1245/s10434-012-2262-1

20. Kim CW, Baik SH, Roh YH, Kang J, Hur H, Min BS, *et al.* Cost-effectiveness of robotic surgery for rectal cancer focusing on short-term outcomes: a propensity score-matching analysis. *Medicine* 2015;**94**:e823. https://doi.org/10.1097/MD.0000000000000823

21. Ramji KM, Cleghorn MC, Josse JM, MacNeill A, O'Brien C, Urbach D, Quereshy FA. Comparison of clinical and economic outcomes between robotic, laparoscopic, and open rectal cancer surgery: early experience at a tertiary care center. *Surg Endosc* 2016;**30**:1337–43. https://doi.org/10.1007/s00464-015-4390-8

22. Collinson FJ, Jayne DG, Pigazzi A, Tsang C, Barrie JM, Edlin R, *et al.* An international, multicentre, prospective, randomised, controlled, unblinded, parallel-group trial of robotic-assisted versus standard laparoscopic surgery for the curative treatment of rectal cancer. *Int J Colorectal Dis* 2012;**27**:233–41. https://doi.org/10.1007/s00384-011-1313-6

23. Nagtegaal ID, Quirke P. What is the role for the circumferential margin in the modern treatment of rectal cancer? *J Clin Oncol* 2008;**26**:303–12. https://doi.org/10.1200/JCO.2007.12.7027

24. Keller DS, McGee MF, Goyal S, Nobel T, O'Brien Ermlich B, Cheruvu VK, Delaney CP. Construct validation and comparison of a novel postoperative quality-of-life metric and the Short Form-36 in colorectal surgery patients. *Surgery* 2013;**154**:690–5. https://doi.org/10.1016/j.surg.2013.06.037

25. Smets EM, Garssen B, Bonke B, De Haes JC. The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *J Psychosom Res* 1995;**39**:315–25. https://doi.org/10.1016/0022-3999(94)00125-O

26. Barry MJ, Fowler FJ, O'Leary MP, Bruskewitz RC, Holtgrewe HL, Mebust WK, Cockett AT, The Measurement Committee of the American Urological Association. The American Urological Association symptom index for benign prostatic hyperplasia. *J Urol* 1992;**148**:1549–57. https://doi.org/10.1016/S0022-5347(17)36966-5

27. Rosen RC, Riley A, Wagner G, Osterloh IH, Kirkpatrick J, Mishra A. The international index of erectile function (IIEF): a multidimensional scale for assessment of erectile dysfunction. *Urology* 1997;**49**:822–30. https://doi.org/10.1016/S0090-4295(97)00238-0

28. Rosen R, Brown C, Heiman J, Leiblum S, Meston C, Shabsigh R, *et al.* The Female Sexual Function Index (FSFI): a multidimensional self-report instrument for the assessment of female sexual function. *J Sex Marital Ther* 2000;**26**:191–208. https://doi.org/10.1080/009262300278597

29. World Health Organization. *Body Mass Index (BMI) classification*. URL: www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi (accessed 11 August 2019).

30. Satagopan JM, Ben-Porat L, Berwick M, Robson M, Kutler D, Auerbach AD. A note on competing risks in survival data analysis. *Br J Cancer* 2004;**91**:1229–35. https://doi.org/10.1038/sj.bjc.6602102

31. Department of Health and Social Care. *NHS Reference Costs 2014 to 2015*. URL: https://gov.uk/government/publications/nhs-reference-costs-2014-to-2015 (accessed 17 October 2018).

32. Curtis L, Burns A. *Unit Costs of Health and Social Care 2015*. Canterbury: Personal Social Services Research Unit, University of Kent; 2015.

33. Information Services Division, NHS National Services Scotland. *Theatres*. URL: www.isdscotland.org/Health-Topics/Finance/Costs/Detailed-Tables/Theatres.asp

34. Curtis L. *Unit Costs of Health and Social Care 2014*. Canterbury: Personal Social Services Research Unit, University of Kent; 2014.

35. Curtis, L. *Unit Costs of Health and Social Care 2010*. Canterbury: Personal Social Services Research Unit, University of Kent; 2010.

36. NHS Employers. *Agenda for Change Pay Bands and Points from 1 April 2015 (England)*. URL: http://nhsemployers.org/-/media/Employers/Documents/Pay-and-reward/AfC-pay-bands-from-1-April-2015.pdf (accessed 17 October 2018).

37. Jones S. *Value of the Nurse Led Stoma Care Clinic. Royal Glamorgan & Ysbyty Cwm Rhondda Hospitals*. 2015. URL: www.rcn.org.uk/professional-development/research-and-innovation/innovation-in-nursing/case-studies-demonstrating-the-value-of-nursing/sheila-jones (accessed 11 August 2019).

38. Joint Formulary Committee. *British National Formulary*. London: BMJ Group and Pharmaceutical Press. URL: https://bnf.org/products/bnf-online/ (accessed 17 October 2018).

39. Department of Health and Social Care. *Drugs and Pharmaceutical Electronic Market Information Tool* (*eMIT*). URL: https://gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit (accessed 17 October 2018).

40. NHS Business Services Authority. *Drug Tariff*. URL: https://nhsbsa.nhs.uk/pharmacies-gp-practices-and-appliance-contractors/drug-tariff (accessed 17 October 2018).

41. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;**35**:1095–108. https://doi.org/10.1097/00005650-199711000-00002

42. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002;**21**:271–92. https://doi.org/10.1016/S0167-6296(01)00130-8

43. National Institute for Health and Care Excellence (NICE). *Guide to the Methods of Technology Appraisal*. 2013. URL: www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781 (accessed 11 August 2019).

44. Claxton K, Martin S, Soares M, Rice N, Spackman E, Hinde S *et al.* Methods for the estimation of the NICE cost-effectiveness threshold. *Health Technol Assess* 2015;**19**(14). https://doi.org/10.3310/hta19140

45. Intuitive Surgical, Inc. *Annual Report 2018*. URL: https://isrg.gcs-web.com/static-files/31b5c428-1d95-4c01-9c85-a7293bac5e05 (accessed 11 August 2019).

46. Iavazzo C, Papadopoulou EK, Gkegkes ID. Cost assessment of robotics in gynecologic surgery: a systematic review. *J Obstet Gynaecol Res* 2014;**40**:2125–34. https://doi.org/10.1111/jog.12507

47. Coronado PJ, Fasero M, Magrina JF, Herraiz MA, Vidart JA. Comparison of perioperative outcomes and cost between robotic-assisted and conventional laparoscopy for transperitoneal infrarenal para-aortic lymphadenectomy (TIPAL). *J Minim Invasive Gynecol* 2014;**21**:674–81. https://doi.org/10.1016/j.jmig.2014.01.023

Appendix 1 Primary end point (conversion to open surgery) analysis: further details

TABLE 51 Deviance and AIC values for likelihood ratio test for conversion

Model	AIC	Deviance	Difference in deviance	*p*-value (likelihood ratio test)
Random intercept	276.94	258.94	1.22	0.269
Random slope	277.72	257.72		

Model diagnostics

Residual index plot

Figure 13 presents the index plot of raw residuals (including EBEs of random effects) on the probability scales (*y*-axis) versus patient ID (*x*-axis). Residual r_i for patient i is calculated as:

$$r_i = Y_i - \hat{p}_i, \tag{5}$$

where

$$Y_i = \begin{cases} 1, & \text{Patient converted to open surgery} \\ 0, & \text{Otherwise} \end{cases}, \tag{6}$$

and $\hat{p}_i$ is the predicted probability of conversion to open surgery for patient i (including EBE of the random effect). Residuals with larger absolute values indicate poorer model fit. Patient 374 (labelled in *Figure 13*) stands out as an instance of poor model fit, with the model yielding a predicted probability of conversion of 0.676, but they were not converted. Many of the residuals for patients who did convert to open surgery (shown in green in *Figure 13*) are large, indicating that the model fitted low probabilities of conversion for those patients, but this is reasonable and perhaps expected because conversion to open surgery was an infrequent event. The empirical probability plot helps us to objectively determine what magnitude of residual is 'expected'.

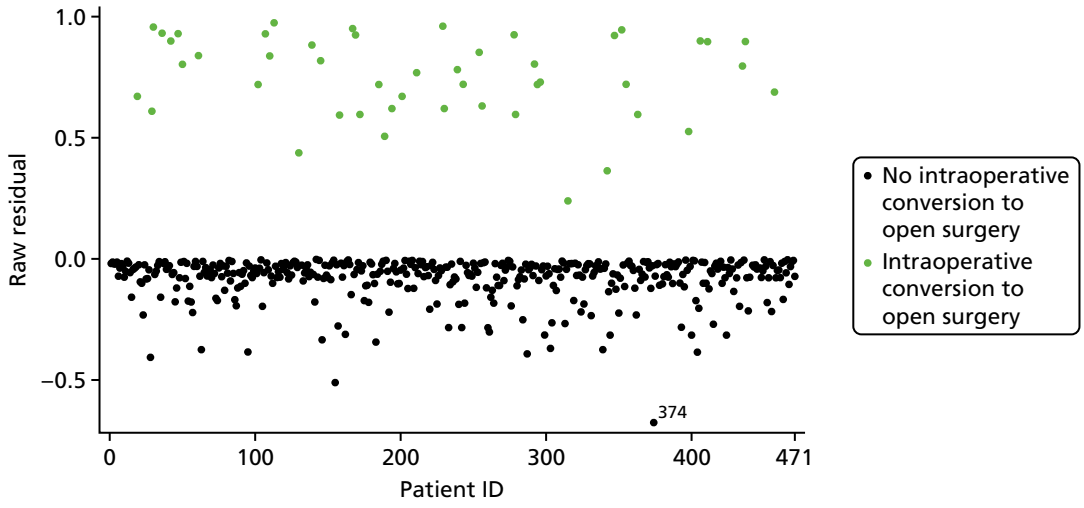


FIGURE 13 Conversion to open surgery: index plot of raw residuals of the random intercept model. Scatterplot: raw residuals on the probability scale (including EBEs of random effects) by observed outcome.

Figure 14 presents the empirical probability plot for the primary analysis model, which can be used to compare actual Pearson residuals with expected Pearson residuals. The *y*-axis is the actual Pearson residual value, the *x*-axis is the empirical median Pearson residual expected under our fitted model assumptions. The dots represent patients. If the model was a perfect fit, then we would expect all of the dots to lie on the reference line. The band in *Figure 14* represents the interval between the empirical 2.5th percentile and 97.5th percentile empirical Pearson residual. No values lie outside this region, indicating that we do not have any substantial outliers.

Delta-betas

Figure 15 presents the plot of exponentiated delta-betas (*y*-axis) versus patient ID. Exponentiated delta-betas further from 1 indicate greater influence of the observation on the estimated treatment effect. Conversions to open surgery demonstrably have greater influence on the estimated treatment effect, which is perhaps expected given that conversion to open surgery was an infrequent event. Patient 374 stands out as having high influence for a non-conversion to open surgery. Patient 374 was in the robotic treatment group and their exponentiated delta-beta is 1.051, indicating that the estimated OR for conversion to open surgery (robotic vs. laparoscopic) increases by a factor of 1.051 when they are removed from the model, from the original 0.614 (95% CI 0.311 to 1.211; $p = 0.16$) to 0.645 (95% CI 0.325 to 1.280; $p = 0.21$).

Further investigation of outliers

The observation for patient 374 is genuine. It is just a relatively unexpected outcome, rather than erroneous data. Patient 374 was male, obese and underwent a LAR, all indicating a higher risk of conversion to open surgery according to the model. Furthermore, their operating surgeon converted 10 out of 33 (30.3%) of their ROLARR patients to open surgery: a much higher rate than average. The model therefore estimated a 67.6% probability of conversion to open surgery for patient 374 but they were not converted, hence the large magnitude of the residual. The EBE of this surgeon's effect on the odds of conversion was that they increased the odds by a factor of around 5 (i.e. a patient being operated on by this surgeon was estimated to be five times more likely to convert to open surgery than a patient operated on by the average surgeon, according

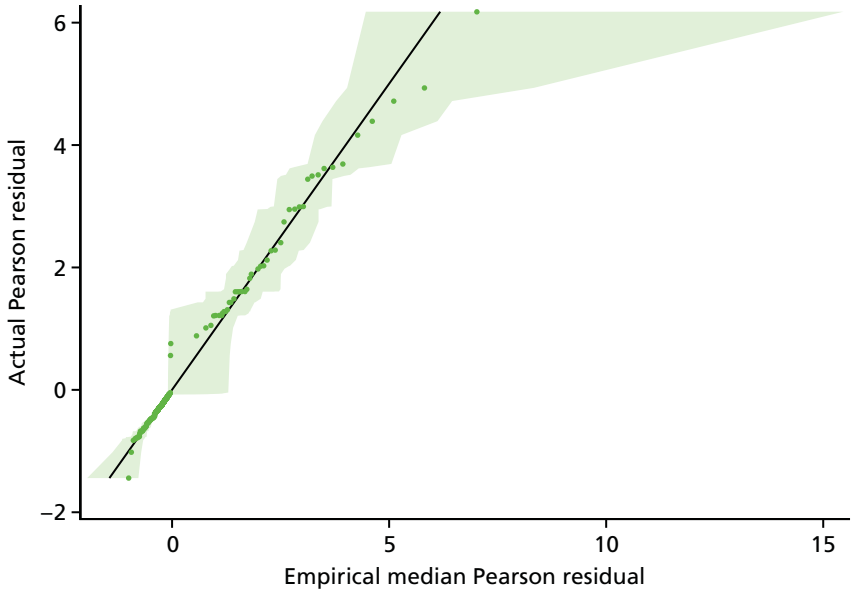


FIGURE 14 Conversion to open surgery: empirical probability plot (including simulated envelope of 2.5th–97.5th percentile empirical Pearson residual) of raw residuals of the random intercept model.

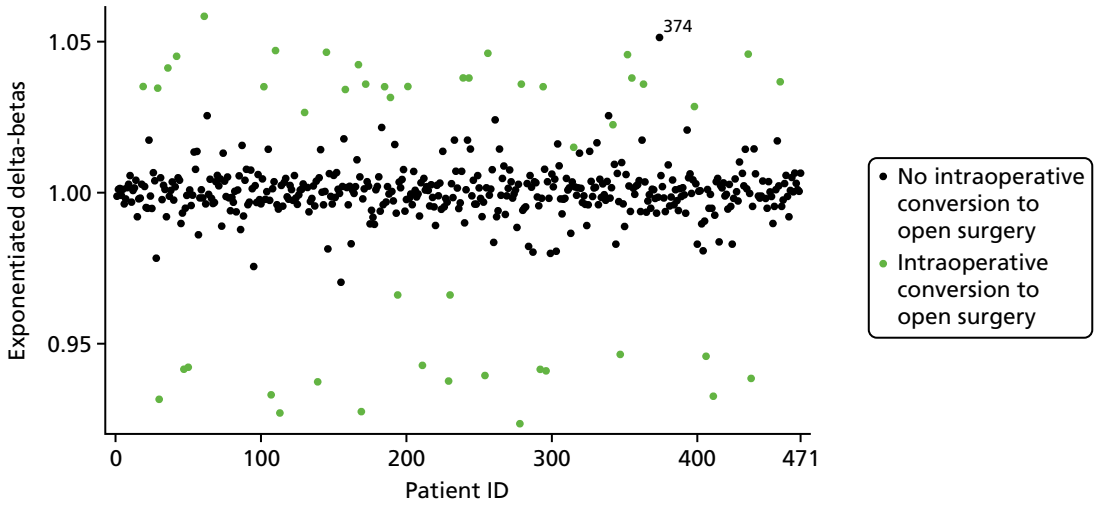


FIGURE 15 Conversion to open surgery: plot of exponentiated delta-betas from the random intercept model. Scatterplot: random intercept model delta-betas by observed outcome.

to this model). This is a larger effect than any of the other patient factors. However, patient 374 was this surgeon's last patient in the ROLARR study, and their level of robotic experience changed substantially throughout their participation in the study, from 35 previous robotic operations before their first patient to 121 before patient 374. Most of the surgeon's conversions came during their earlier ROLARR cases. The learning effects analysis (see *Chapter 3*, *Sensitivity analysis: learning effects*) accounts for this improvement over time and gives a predicted probability of conversion of 42.3% for patient 374, yielding a smaller residual that lies more comfortably within the expected range of residual values. It seems that the large residual in this model is therefore mainly because of poor fit as a consequence of assuming no learning effects.

Sensitivity analysis: actual operating surgeon – further details

TABLE 52 Sensitivity analysis by actual operating surgeon, by treatment group

Surgeon other than randomised	Treatment group, *n* (%)		Total (*N* = 466), *n* (%)
	Standard laparoscopic surgery (*N* = 230)	Robotic-assisted laparoscopic surgery (*N* = 236)	
Yes	16 (7.0)	26 (11.0)	42 (9.0)
No	214 (93.0)	210 (89.0)	424 (91.0)

TABLE 53 Mixed-effects logistic regression model adjusting for actual operating surgeon instead of intended operating surgeon

Effect: comparator group (vs. reference group)	Group [number of conversions/number of patients (%)]		Risk difference (unadjusted 95% CI)	OR (adjusted)	95% CI for OR (adjusted)	*p*-value
	Reference	Comparator				
Treatment: robotic surgery (vs. laparoscopic)	28/230 (12.2)	19/236 (8.1)	4.1 (–1.4 to 9.6)	0.612	0.310 to 1.207	0.16
Sex: male (vs. female)	8/149 (5.4)	39/317 (12.3)	–6.9 (–12.1 to 1.8)	2.449	1.046 to 5.734	0.04
BMI class: overweight (vs. underweight/normal)	13/179 (7.3)	9/180 (5.0)	2.3 (–2.7 to 7.2)	0.547	0.215 to 1.396	0.21
BMI class: obese (vs. underweight/normal)	13/179 (7.3)	25/107 (23.4)	–16.1 (–25.0 to –7.2)	4.649	2.069 to 10.448	0.0002
Previous radiotherapy or chemoradiotherapy: yes (vs. no)	27/262 (10.3)	20/204 (9.8)	0.5 (–5.0 to 6.0)	1.094	0.512 to 2.338	0.82
Intended procedure: HAR (vs. LAR)	37/312 (11.9)	6/68 (8.8)	3.0 (–4.6 to 10.7)	0.564	0.199 to 1.598	0.28
Intended procedure: APR (vs. LAR)	37/312 (11.9)	4/86 (4.7)	7.2 (1.5 to 12.9)	0.185	0.054 to 0.631	0.007

TABLE 54 Actual operating surgeon instead of intended operating surgeon: estimate of the variance component from random intercept model

Effect	Variance component	
	Estimate	Standard error
Operating surgeon (random effect)	0.612	0.412

Sensitivity analysis: actual procedure – further details

The most notable effect that adjusting for actual procedure had on the model estimates was on the effect of APR (vs. LAR), which is attenuated compared with the primary analysis model. Specifically, in this model (*Table 57*) the OR is 0.433 (95% CI 0.165 to 1.134; $p = 0.088$) compared with 0.184 (95% CI 0.054 to 0.627; $p = 0.007$) in the primary analysis model.

TABLE 55 Actual procedure performed, by treatment group

Procedure performed	Treatment group, *n* (%)		
	Standard laparoscopic surgery (*N* = 230)	Robotic- assisted laparoscopic surgery (*N* = 236)	Total (*N* = 466), *n* (%)
HAR	19 (8.3)	28 (11.9)	47 (10.1)
LAR	165 (71.7)	152 (64.4)	317 (68.0)
APR	45 (19.6)	52 (22.0)	97 (20.8)
Other	1 (0.4)	4 (1.7)	5 (1.1)

TABLE 56 Details of 'other' procedures

Patient ID	Treatment allocation	Intended procedure	Details of procedure
113	Robotic	LAR	Dorsal pelvic exenteration, ureter resection distally right sided
139	Robotic	LAR	Hartmann's procedure
143	Standard	LAR	Laparoscopic biopsy of peritoneum
183	Robotic	LAR	HAR + subtotal colectomy
429	Robotic	HAR	Hartmann's procedure

TABLE 57 Mixed-effects logistic regression model adjusting for actual procedure instead of intended procedure

Effect: comparator group (vs. reference group)	Group [number of conversions/number of patients (%)]		Risk difference (unadjusted 95% CI)	OR (adjusted)	95% CI for OR (adjusted)	*p*-value
	Reference	Comparator				
Treatment: robotic surgery (vs. laparoscopic)	28/230 (12.2)	19/236 (8.1)	4.1 (–1.4 to 9.6)	0.572	0.289 to 1.132	0.1083
Sex: male (vs. female)	8/149 (5.4)	39/317 (12.3)	–6.9 (–12.1 to 1.8)	2.401	1.034 to 5.573	0.0416
BMI class: overweight (vs. underweight/normal)	13/179 (7.3)	9/180 (5.0)	2.3 (–2.7 to 7.2)	0.562	0.221 to 1.432	0.2264
BMI class: obese (vs. underweight/normal)	13/179 (7.3)	25/107 (23.4)	–16.1 (–25.0 to –7.2)	4.374	1.972 to 9.700	0.0003
Previous radiotherapy or chemoradiotherapy: yes (vs. no)	27/262 (10.3)	20/204 (9.8)	0.5 (–5.0 to 6.0)	1.050	0.509 to 2.166	0.8956
Procedure: HAR (vs. LAR)	33/317 (10.4)	4/47 (8.5)	1.9 (–6.8 to 10.6)	0.718	0.209 to 2.464	0.5975
Procedure: APR (vs. LAR)	33/317 (10.4)	8/97 (8.3)	2.2 (–4.3 to 8.6)	0.433	0.165 to 1.134	0.0883
Procedure: other (vs. LAR)	33/317 (10.4)	2/5 (40.0)	–29.6 (–72.7 to 13.5)	5.934	0.689 to 51.079	0.1047

TABLE 58 Actual procedure instead of intended procedure: estimate of the variance component from random intercept model

	Variance component	
Effect	Estimate	Standard error
Operating surgeon (random effect)	0.535	0.389

Appendix 2 Key secondary end point: circumferential resection margin positivity (CRM+)

Model diagnostics

Residual index plot

Figure 16 presents the index plot of raw residuals (including EBEs of random effects) on the probability scales (*y*-axis) versus patient ID (*x*-axis). Residual r_i for patient i is calculated as:

$$r_i = Y_i - \hat{p}_i, \tag{7}$$

in which:

$$Y_i = \begin{cases} 1, & \text{CRM+} \\ 0, & \text{Otherwise} \end{cases}, \tag{8}$$

and $\hat{p}_i$ is the predicted probability of CRM+ for patient i (including EBE of the random effect). Residuals with larger absolute values indicate poorer model fit. Many of the residuals for patients who did have CRM+ (shown in green in *Figure 16*) are large, indicating that the model fitted low probabilities of CRM+ for those patients, but this is reasonable and perhaps expected because CRM+ was a rare event. There are no clear outliers in *Figure 16*. The empirical probability plot helps us to objectively determine what magnitude of residual is 'expected'.

Figure 17 presents the empirical probability plot for the primary analysis model, which can be used to compare actual Pearson residuals with expected Pearson residuals. The *y*-axis is the actual Pearson residual value and the *x*-axis is the empirical median Pearson residual expected under our fitted model assumptions. The dots represent patients. If the model was a perfect fit, then we would expect all of the dots to lie on the reference line. The band in *Figure 17b* represents the interval between the empirical 2.5th percentile and 97.5th percentile empirical Pearson residual. No values lie outside this region, indicating that we do not have any substantial outliers.

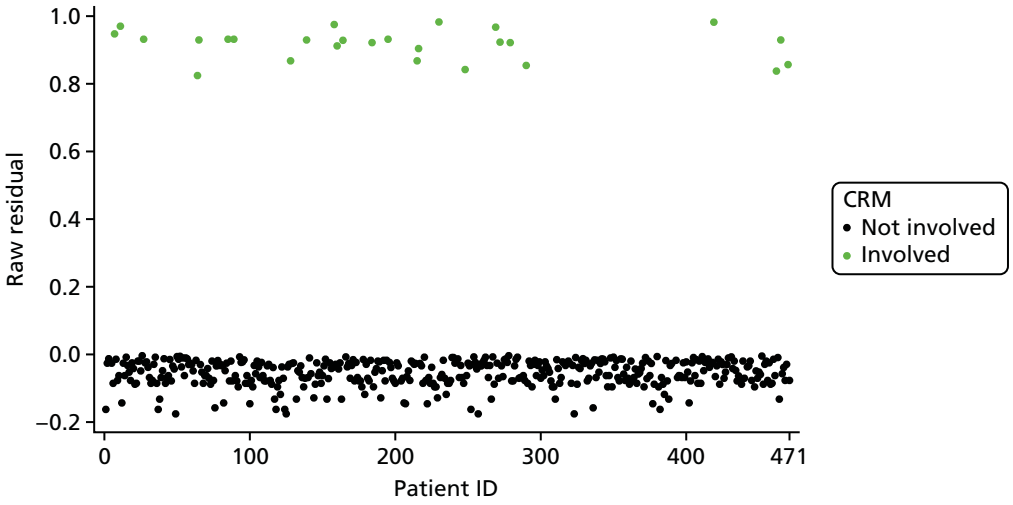


FIGURE 16 Circumferential resection margin positivity: index plot of raw residuals of the random intercept model. Scatterplot: raw residuals on the probability scale (including EBEs of random effects) by observed outcome.

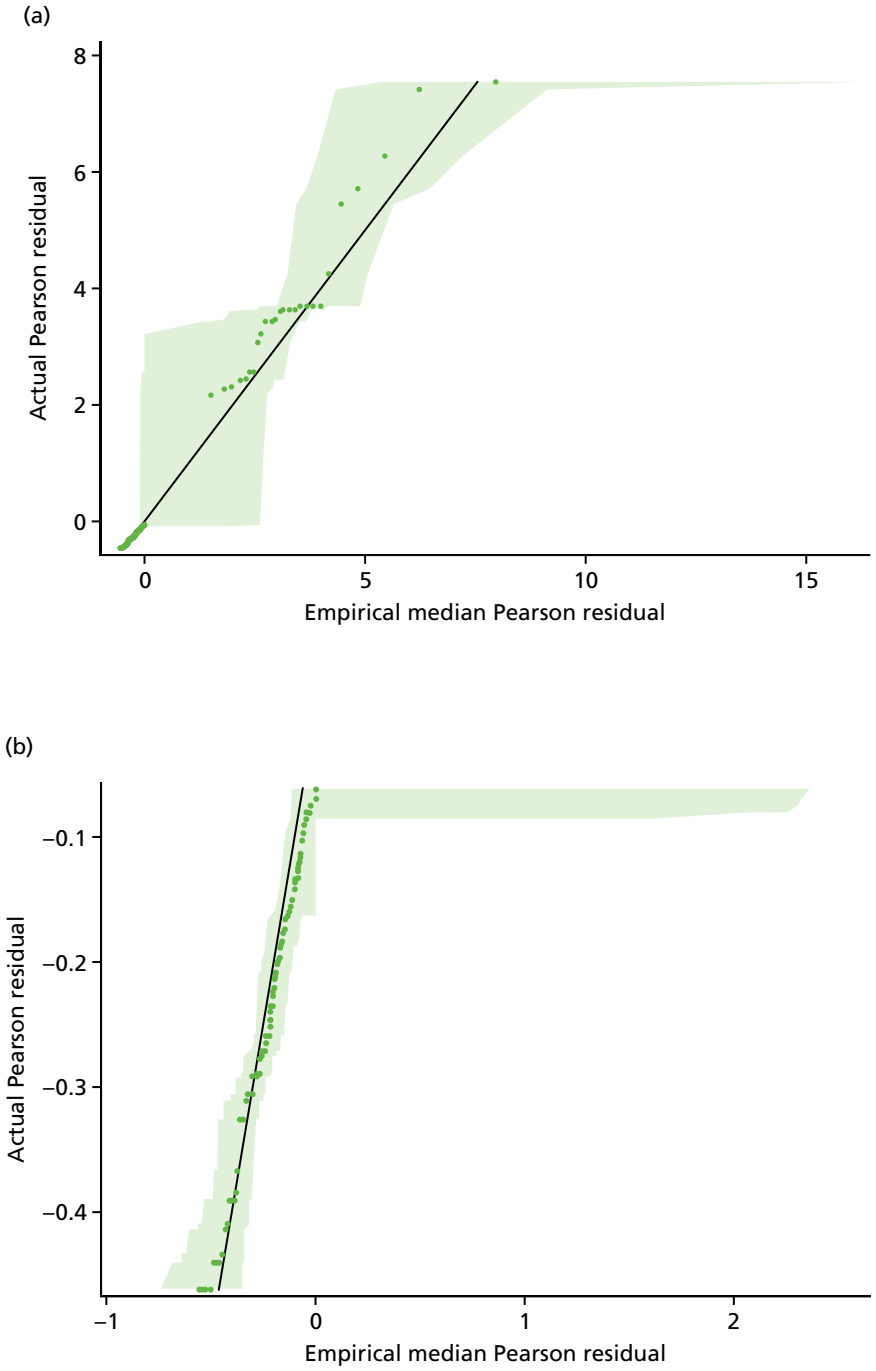


FIGURE 17 Circumferential resection margin positivity: empirical probability plot (including simulated envelope of 2.5th–97.5th percentile empirical Pearson residual) of raw residuals of the random intercept model.

There are two plots:

1. empirical probability plot with confidence region
2. same as plot 1 except restricted to only negative residuals (this has been added because the negative residuals are difficult to read on the other plot).

Delta-betas

Figure 18 presents the plot of exponentiated delta-betas (*y*-axis) versus patient ID. Exponentiated delta-betas further from 1 indicate greater influence of the observation on the estimated treatment effect. The CRM+ observations demonstrably have greater influence on the estimated treatment effect, which is expected given that CRM+ was a rare event. There are no clear overly influential observations.

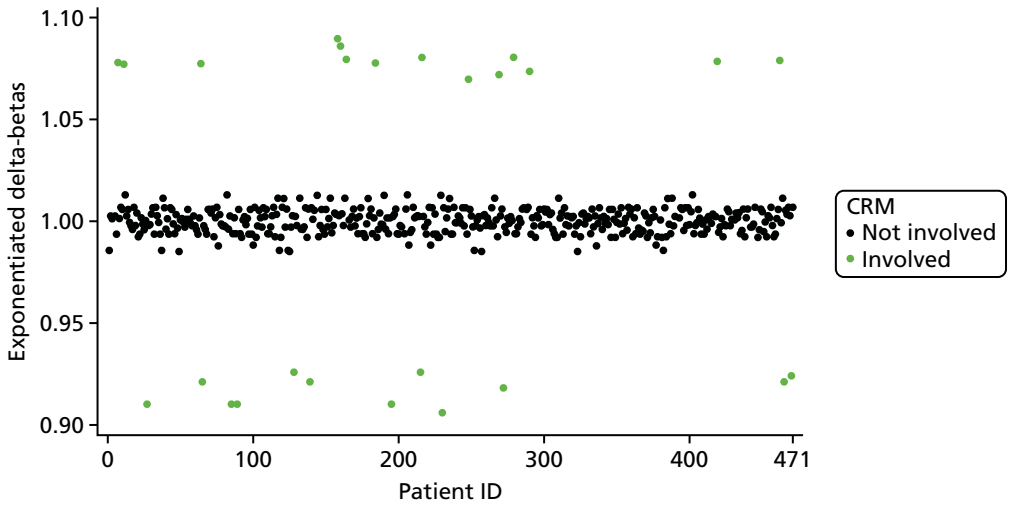


FIGURE 18 Circumferential resection margin positivity: plot of exponentiated delta-betas from the random intercept model. Scatterplot: random intercept model delta-betas by observed outcome.

Subgroup analyses

TABLE 59 Estimated treatment effect ORs for CRM+ subgroup analysis by sex

Effect	Treatment group [number of CRM+/number of patients (%)]		Risk difference (unadjusted 95% CI)	OR (adjusted 95% CI)[a]	*p*-value	
	Laparoscopic surgery	Robotic surgery				
Treatment in males: robotic surgery (vs. laparoscopic)	10/151 (6.6)	12/160 (7.5)	–0.9 (–6.6 to 4.8)	1.118 (0.462 to 2.705)	0.9961	0.9961[b]
Treatment in females: robotic surgery (vs. laparoscopic)	4/73 (5.5)	0/75 (0.0)	5.5 (0.3 to 10.7)	< 0.001 (0 to ∞)	1.000	

a Adjusted for BMI class, preoperative radiotherapy, intended procedure and operating surgeon.
b *p*-value for the treatment effect is referring to a test of heterogeneity of treatment effect between the subgroups.

TABLE 60 Estimated treatment effect ORs for CRM+ subgroup analysis by WHO obesity classification

Effect	Treatment group [number of CRM+/number of patients (%)]		Risk difference (unadjusted 95% CI)	OR (adjusted 95% CI)[a]	*p*-value	
	Laparoscopic surgery	Robotic surgery				
Treatment in obese patients: robotic-assisted surgery (vs. laparoscopic)	1/52 (1.9)	1/53 (1.9)	0.0 (–5.2 to 5.3)	0.989 (0.302 to 3.244)	0.9493	0.5673[b]
Treatment in overweight patients: robotic-assisted surgery (vs. laparoscopic)	6/87 (6.9)	6/89 (6.7)	0.2 (–7.3 to 7.6)	0.914 (0.055 to 15.293)	0.9856	0.7909[b]
Treatment in underweight and normal patients: robotic-assisted surgery (vs. laparoscopic)	7/85 (8.2)	5/93 (5.4)	2.9 (–4.6 to, 10.3)	0.604 (0.180 to 2.021)	0.4123	

a Adjusted for sex, preoperative radiotherapy, intended procedure and operating surgeon.
b *p*-value for the treatment effect is referring to a (pairwise) test of heterogeneity of treatment effect between the subgroups. For example, the second *p*-value in the 'Treatment in obese patients' row refers to a test of heterogeneity of treatment effect between obese patients and underweight/normal patients.

TABLE 61 Estimated treatment effect ORs for CRM+ subgroup analysis by T-stage

Effect	Treatment group [number of CRM+/number of patients (%)]		Risk difference and (unadjusted 95% CI)
	Laparoscopic surgery	Robotic surgery	
Treatment in T0 patients: robotic-assisted surgery (vs. laparoscopic)	0/25 (0.0)	0/26 (0.0)	.
Treatment in T1 patients: robotic-assisted surgery (vs. laparoscopic)	0/24 (0.0)	0/24 (0.0)	.
Treatment in T2 patients: robotic-assisted surgery (vs. laparoscopic)	1/62 (1.6)	2/68 (2.9)	–1.3 (–6.4 to 3.8)
Treatment in T3 patients: robotic-assisted surgery (vs. laparoscopic)	11/106 (10.4)	8/111 (7.2)	3.2 (–4.4 to 10.7)
Treatment in T4 patients: robotic-assisted surgery (vs. laparoscopic)	2/7 (28.6)	2/5 (40.0)	–11.4 (–65.9 to 43.0)

Appendix 3 Key secondary end point: 3-year local recurrence – further details

Local recurrences and censorings, including the reason for censoring, are summarised in *Table 62*.

The methods of confirmation are summarised in *Table 63*.

Table 64 shows the estimated cumulative incidence of local recurrence by treatment group at several time points (1–5 years post randomisation).

TABLE 62 Nature of the end of follow-up for local recurrence analysis, by treatment group

Nature of the end of follow-up for local recurrence analysis (by treatment group)	Treatment group, *n* (%)		
	Standard laparoscopic surgery (*N* = 234)	Robotic-assisted laparoscopic surgery (*N* = 237)	Total (*N* = 471), *n* (%)
Event: local recurrence	14 (6.0)	16 (6.8)	30 (6.4)
Censor: last known to be alive	195 (83.3)	199 (84.0)	394 (83.7)
Censor[a]: death	16 (6.8)	15 (6.3)	31 (6.6)
Censor: withdrawal from further data collection	5 (2.1)	4 (1.7)	9 (1.9)
Censor: non-standard circumstance[b]	4 (1.7)	3 (1.3)	7 (1.5)

a Considered a competing risk event, rather than a censored observation, in the evaluation of the cumulative incidence function.
b Three patients had benign disease (two laparoscopic, one robotic), three patients had a non-curative surgery outcome (one laparoscopic, two robotic) and one patient did not undergo surgery (laparoscopic).

TABLE 63 Method of confirmation of local recurrences, by treatment group

Method of confirmation	Treatment group, *n* (%)		
	Standard laparoscopic surgery (*N* = 14)	Robotic-assisted laparoscopic surgery (*N* = 16)	Total (*N* = 30), *n* (%)
Clinical	2 (14.3)	2 (12.5)	4 (13.3)
Radiological	5 (35.7)	6 (37.5)	11 (36.7)
Pathological	7 (50.0)	8 (50.0)	15 (50.0)

TABLE 64 Estimated cumulative incidence of local recurrence, by treatment group

Time post randomisation (years)	Treatment group			
	Laparoscopic surgery		Robotic surgery	
	Probability of local recurrence	95% CI	Probability of local recurrence	95% CI
1	0.022	0.003 to 0.041	0.034	0.011 to 0.058
2	0.049	0.040 to 0.058	0.052	0.023 to 0.080
3	0.058	0.028 to 0.089	0.061	0.030 to 0.091
4	0.065	0.055 to 0.076	0.078	0.039 to 0.116
5	0.065	0.032 to 0.099	0.078	0.039 to 0.116

Model diagnostics

Deviance residuals

As a result of heavy censoring, the deviance residuals form a bimodal distribution, as seen in *Figure 19*. There are no clear outliers.

Figure 20 shows the standardised Schoenfeld residuals versus time. If the proportional hazards assumption was being violated, then we would expect the relationship between these residuals and time to deviate from a flat line in at least one of the treatment groups, but it does not. This is also reflected in *Figure 21* (standardised score process), as the observed path lies within the simulated paths. Finally, the supremum test of the PH assumption returned a *p*-value of 0.265 relating to the treatment effect. All of this suggests that the proportional hazards assumption is viable for the treatment groups.

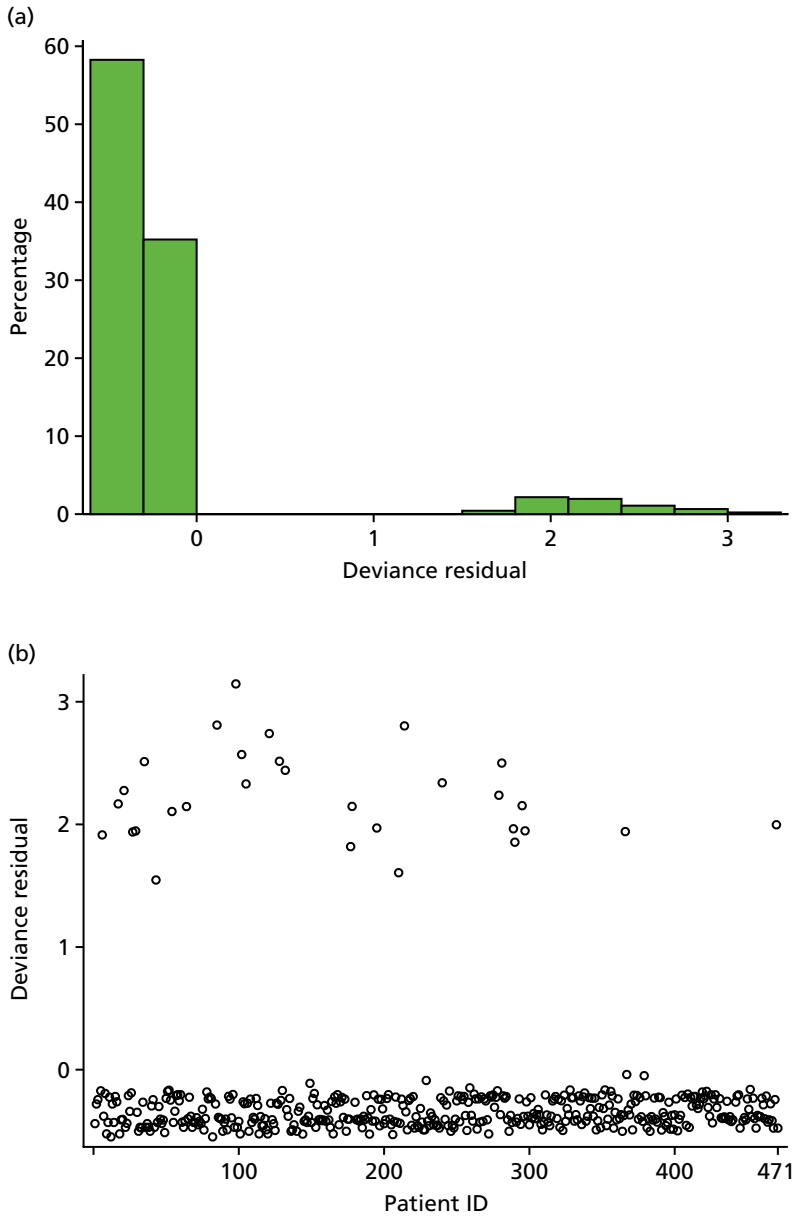


FIGURE 19 Three-year local recurrence. (a) Histogram of deviance residuals; and (b) scatterplot of deviance residuals.

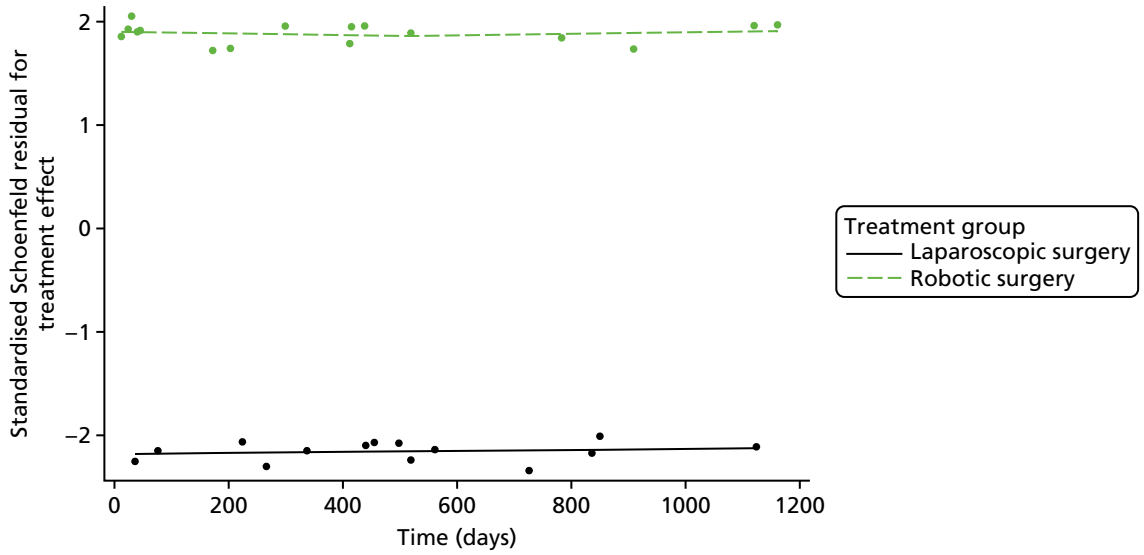


FIGURE 20 Three-year local recurrence: Loess plot of standardised Schoenfeld residuals (by treatment group) for the shared frailty model.

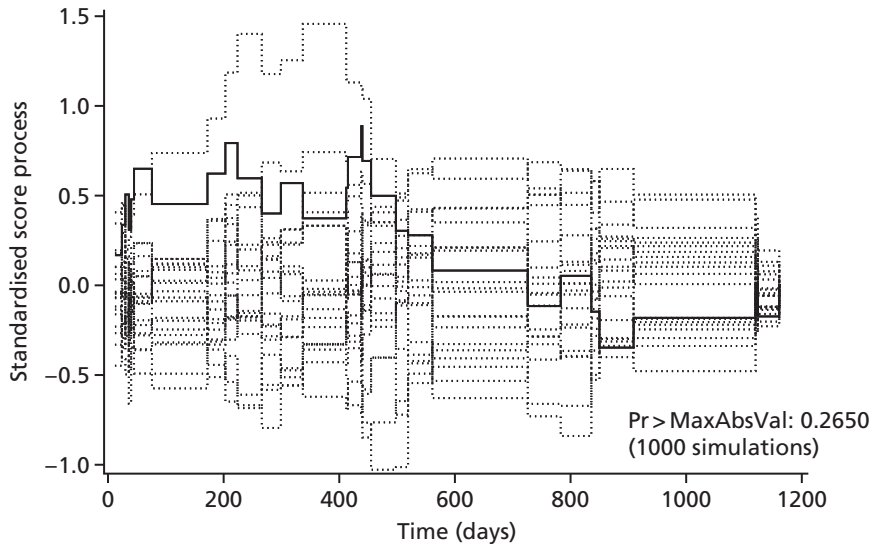


FIGURE 21 Three-year local recurrence: random sample of standardised score process simulated paths.

Delta-betas

Figure 22 presents the exponentiated delta-betas. As one might expect, given the low incidence of local recurrence, the influence of all observed events of local recurrence is notably greater than the influence of censored observations. There is, however, no clear overly influential observations.

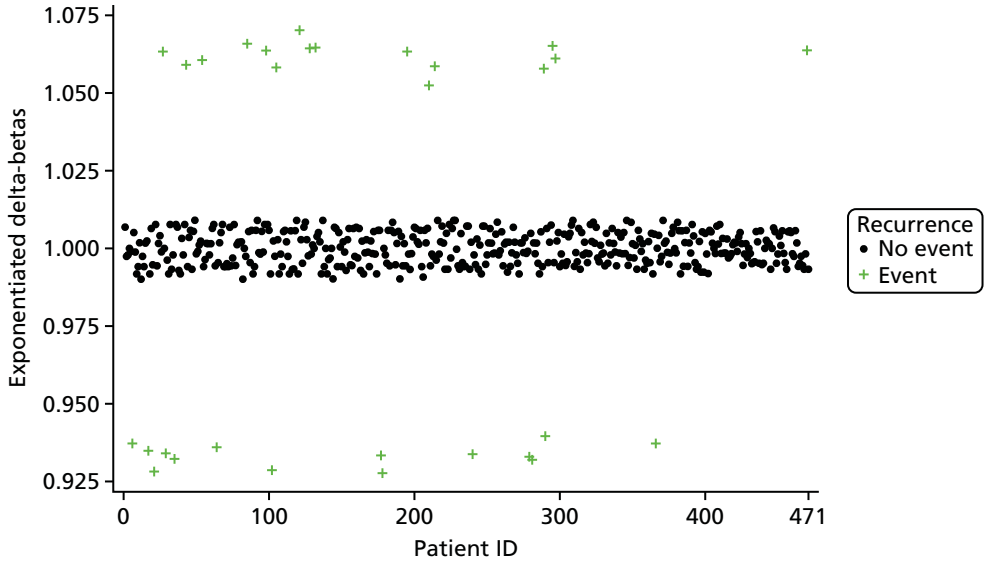


FIGURE 22 Three-year local recurrence: plot of exponentiated delta-betas from the random intercept model.

Subgroup analyses

Odds ratios presented in *Tables 65* and *66* are derived from the linear combination of the estimated treatment (main effect) and treatment-by-subgroup interaction terms on the logit scale. The *p*-values are presented for the test of the treatment effect within each subgroup; this is the first column of *p*-values. For example, in *Table 68* the test that the treatment effect is null (OR = 1) within the female subgroup is 1.000. The *p*-values are also presented for the test of heterogeneity of treatment effect across subgroups, the details of which are given in the footnotes of the tables. Note that a full model was not fitted to test T-stage subgroup analyses, because the small sample sizes and event rates within T-stage groups caused model convergence issues and so crude summaries are given.

Figures 23–26 display the Kaplan–Meier graphs for the effect of neo-adjuvant therapy, operation type, T-stage and sex, on 3-year local recurrence.

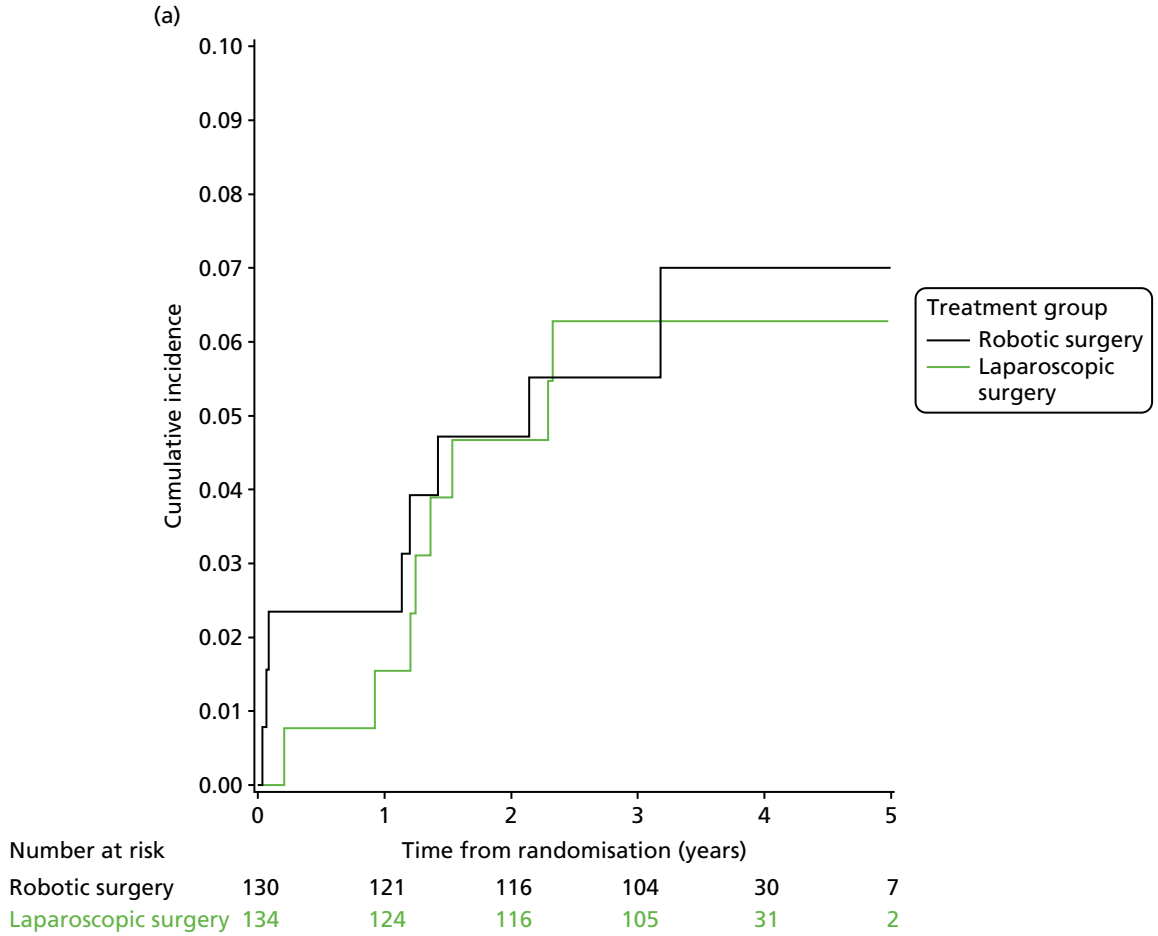


FIGURE 23 Three-year local recurrence by neo-adjuvant therapy. (a) No neo-adjuvant therapy; and (b) neo-adjuvant therapy. (*continued*)

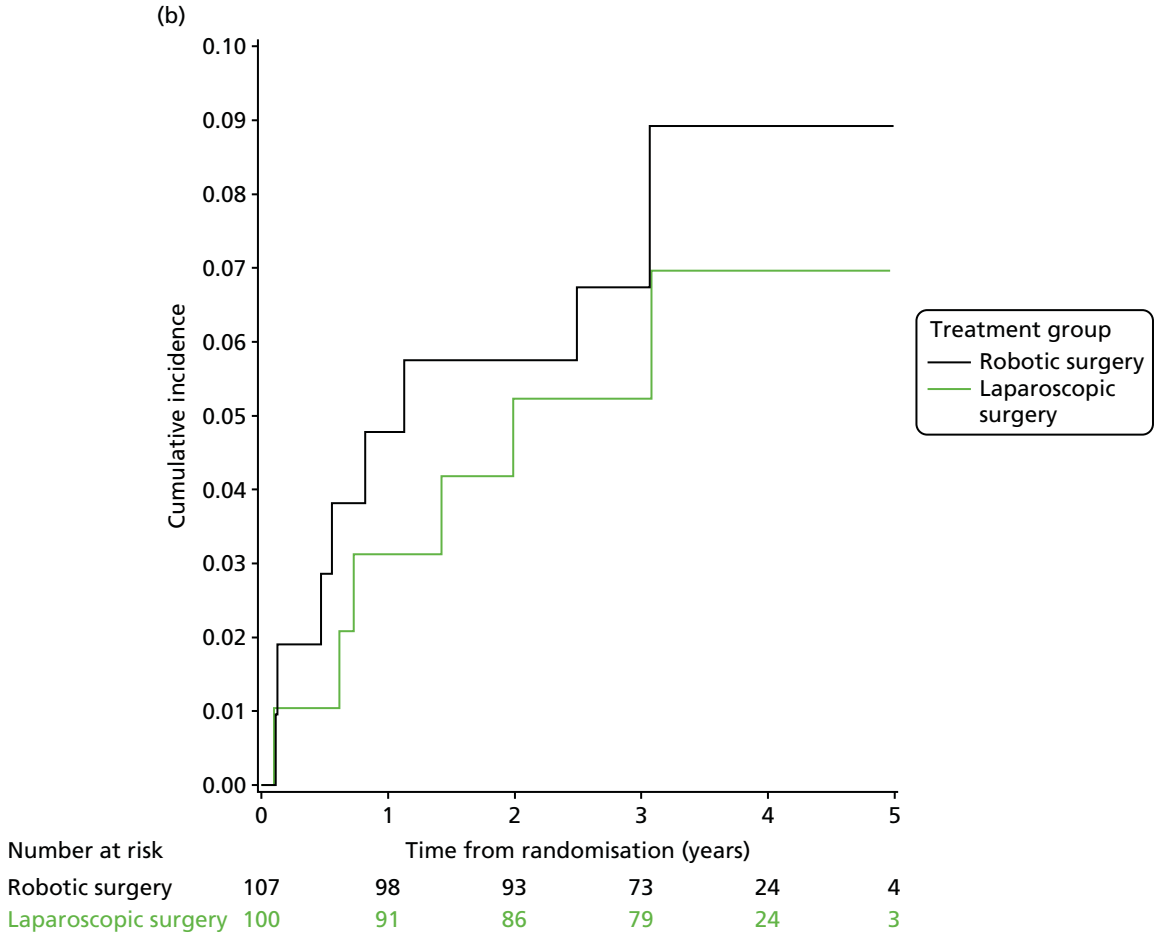


FIGURE 23 Three-year local recurrence by neo-adjuvant therapy. (a) No neo-adjuvant therapy; and (b) neo-adjuvant therapy.

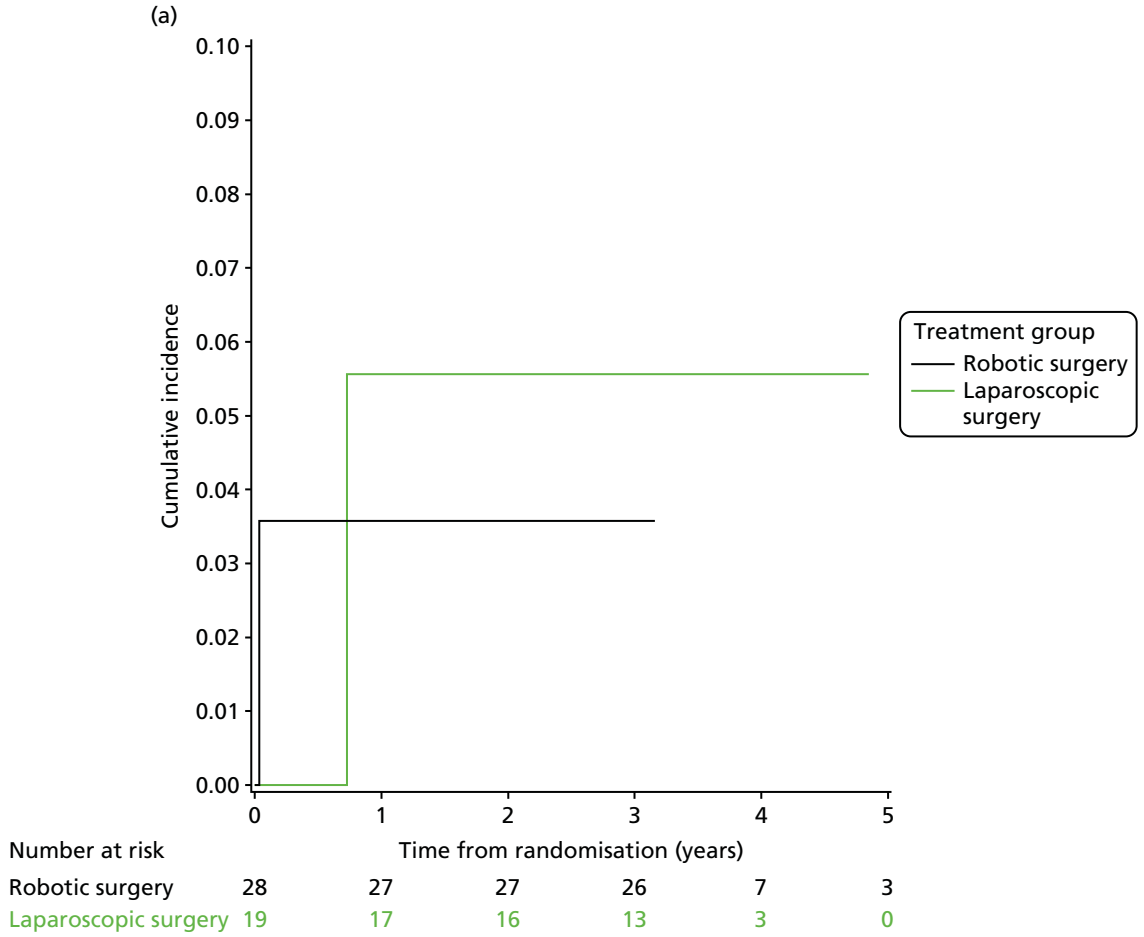


FIGURE 24 Three-year local recurrence by operation type. (a) Cumulative incidence of local recurrence (operation = HAR); (b) cumulative incidence of local recurrence (operation = LAR); and (c) cumulative incidence of local recurrence (operation = APR). (*continued*)

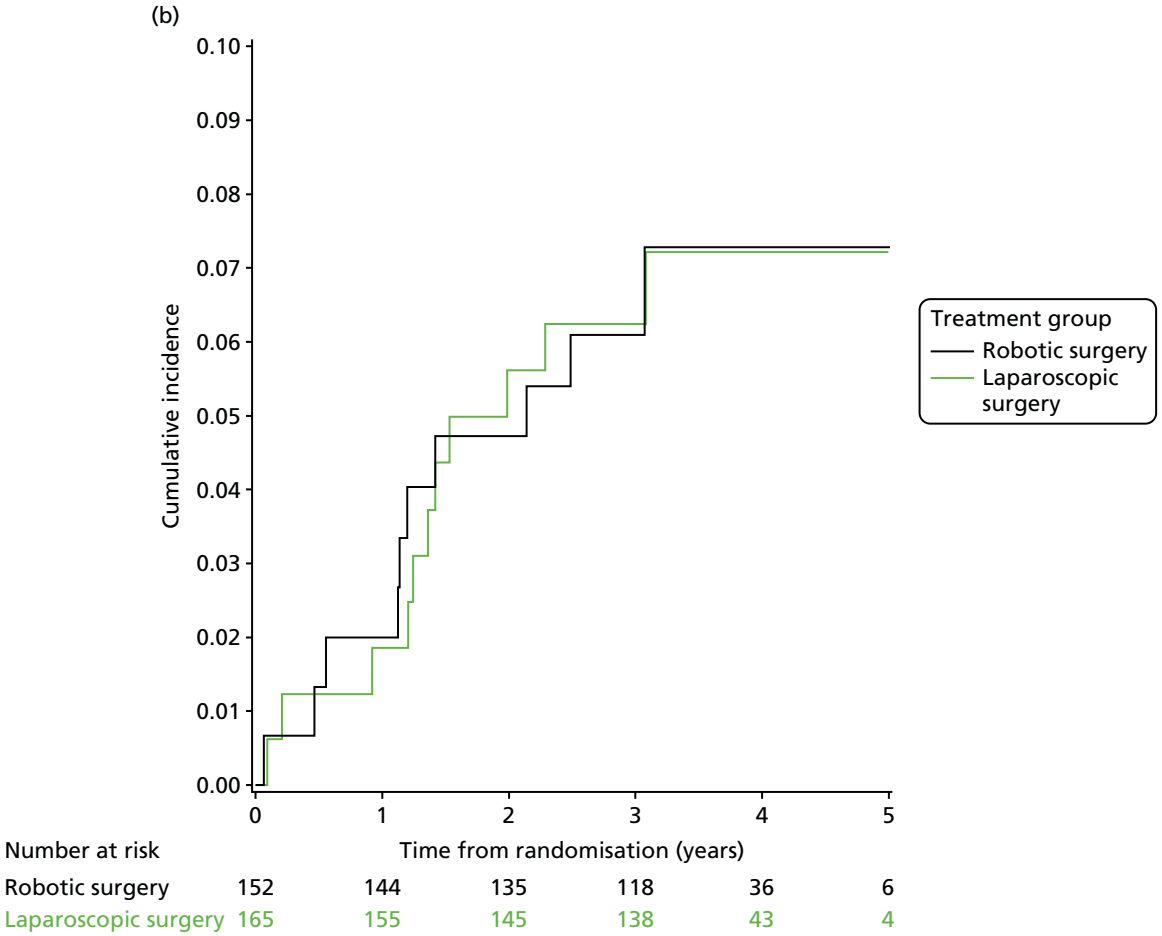


FIGURE 24 Three-year local recurrence by operation type. (a) Cumulative incidence of local recurrence (operation = HAR); (b) cumulative incidence of local recurrence (operation = LAR); and (c) cumulative incidence of local recurrence (operation = APR). (*continued*)

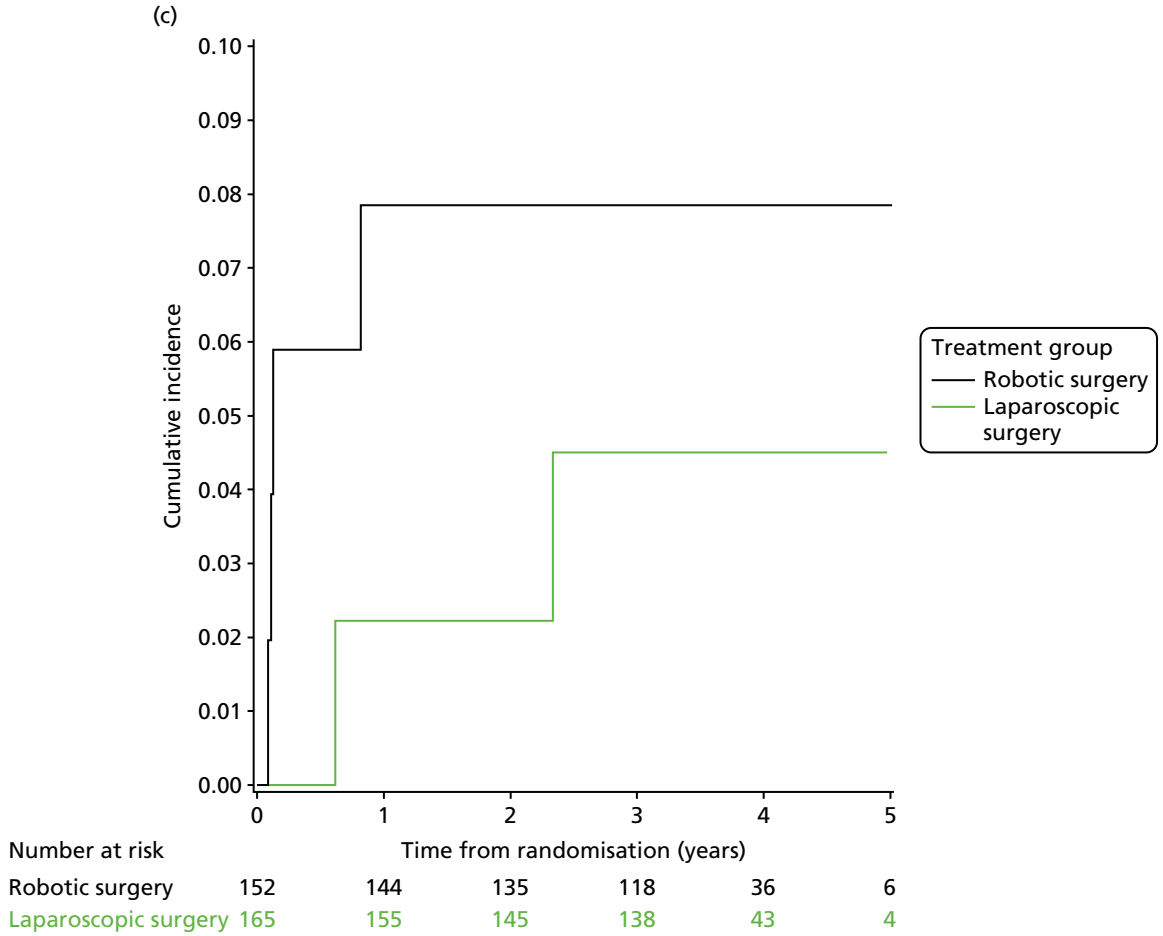


FIGURE 24 Three-year local recurrence by operation type. (a) Cumulative incidence of local recurrence (operation = HAR); (b) cumulative incidence of local recurrence (operation = LAR); and (c) cumulative incidence of local recurrence (operation = APR).

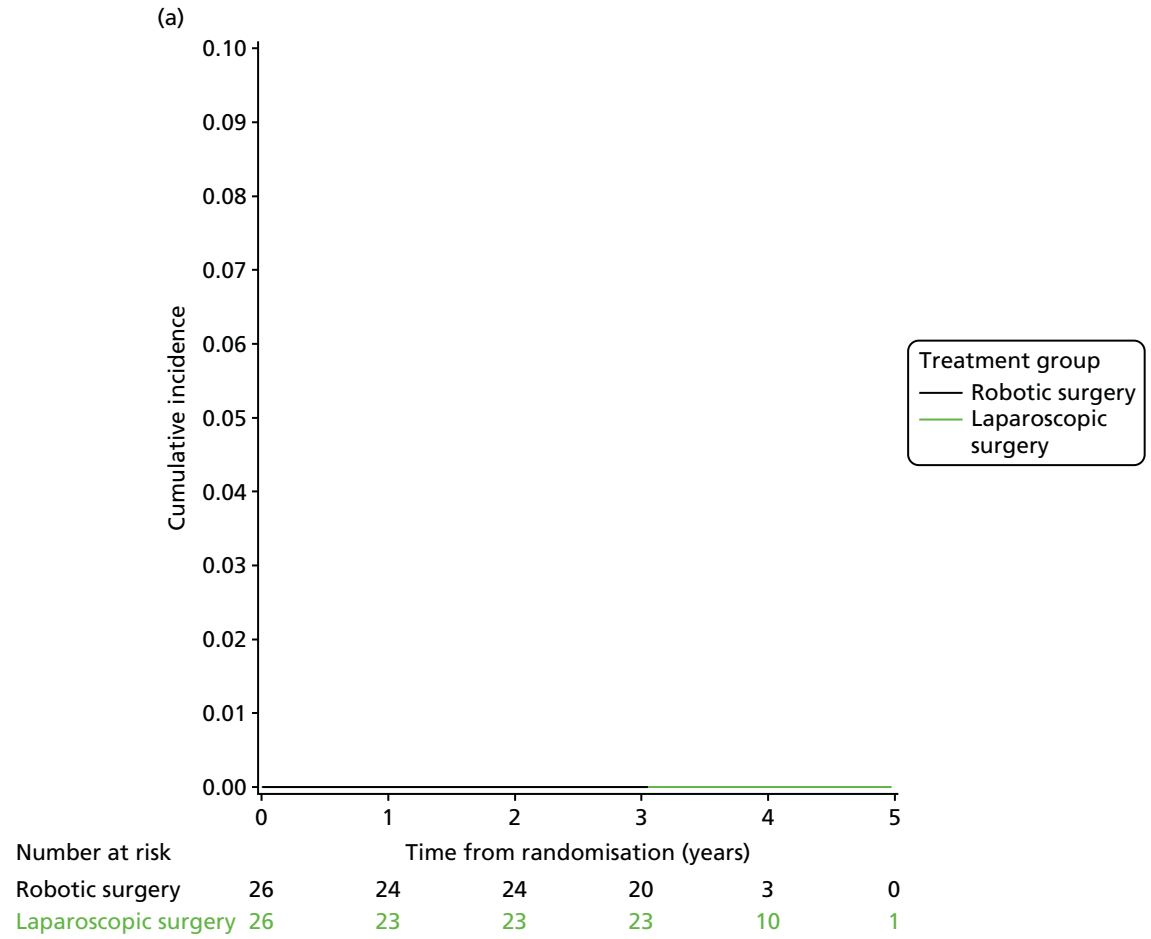


FIGURE 25 Three-year local recurrence by T-stage. (a) T1; (b) T2; and (c) T3. (*continued*)

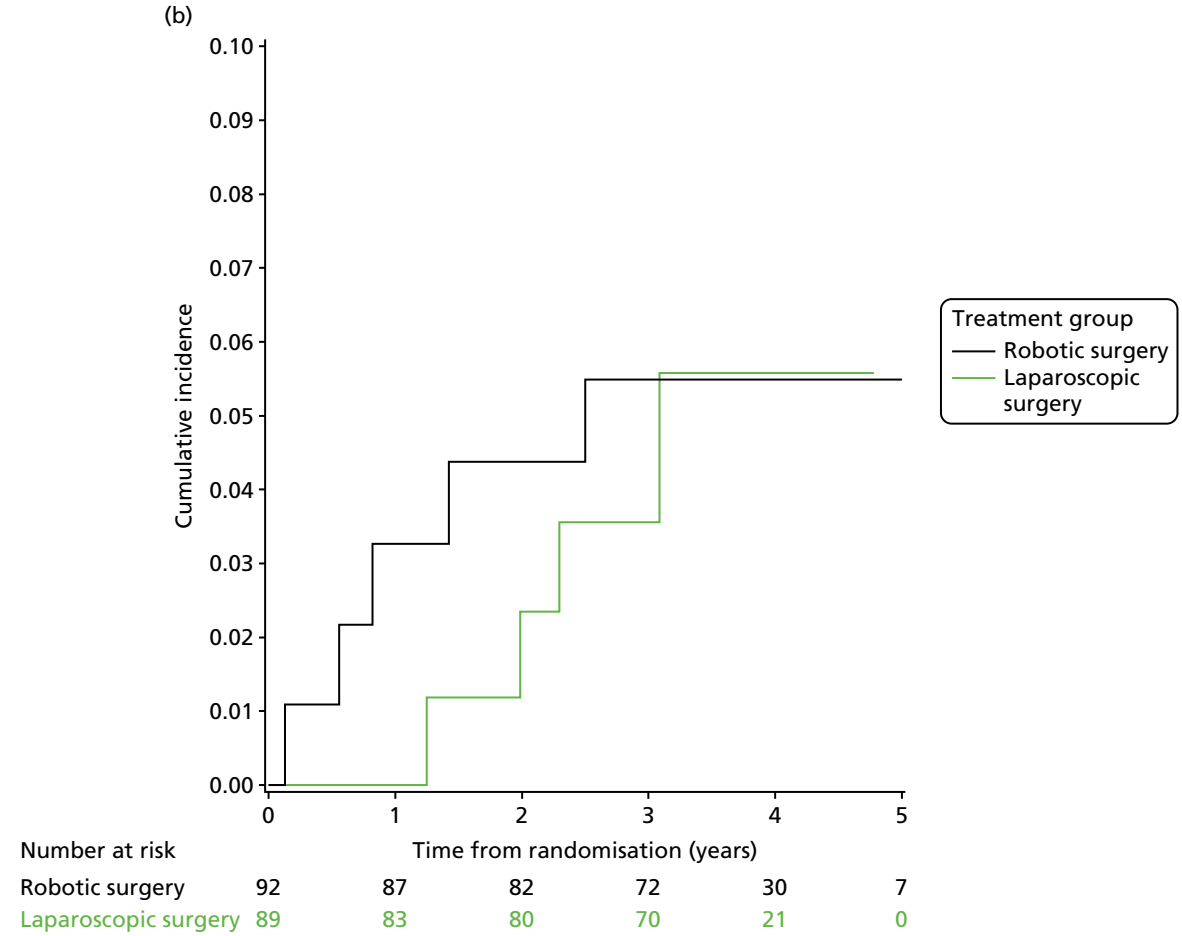


FIGURE 25 Three-year local recurrence by T-stage. (a) T1; (b) T2; and (c) T3. (*continued*)

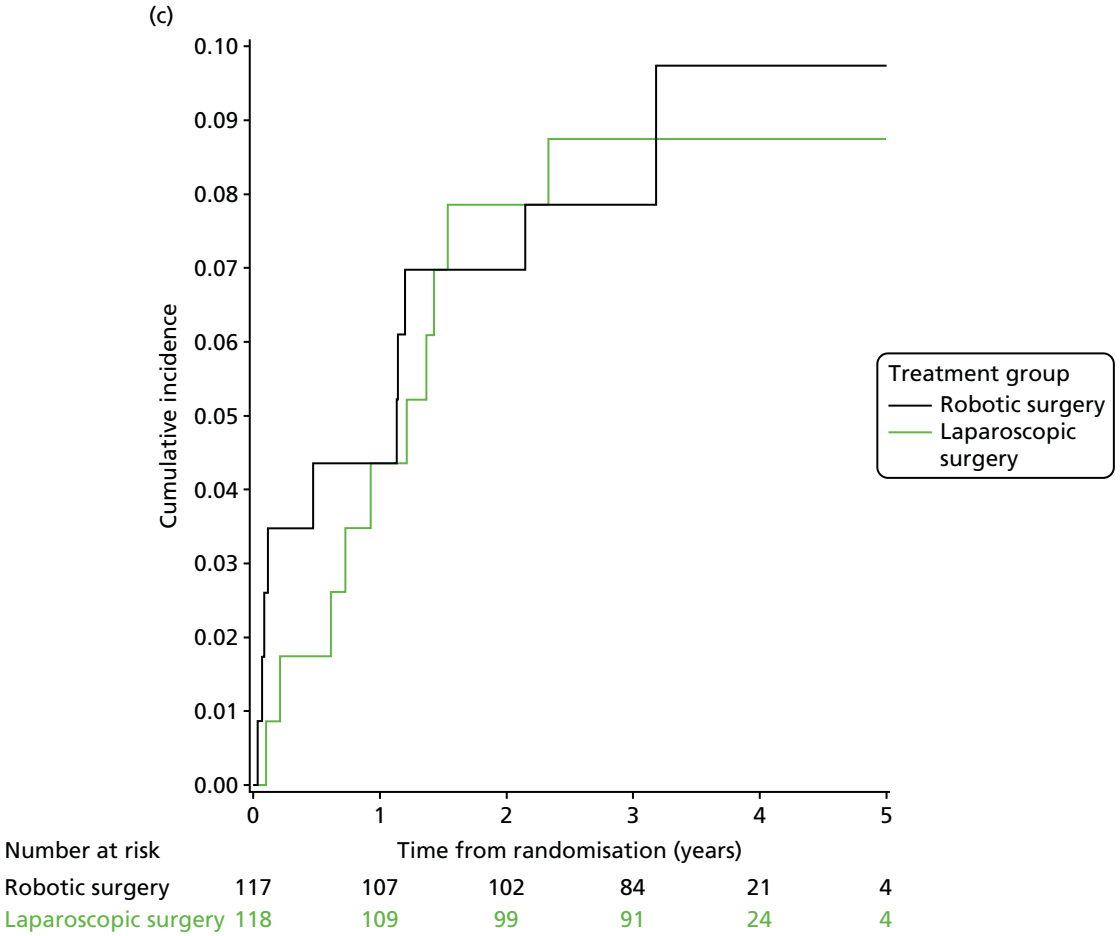


FIGURE 25 Three-year local recurrence by T-stage. (a) T1; (b) T2; and (c) T3.

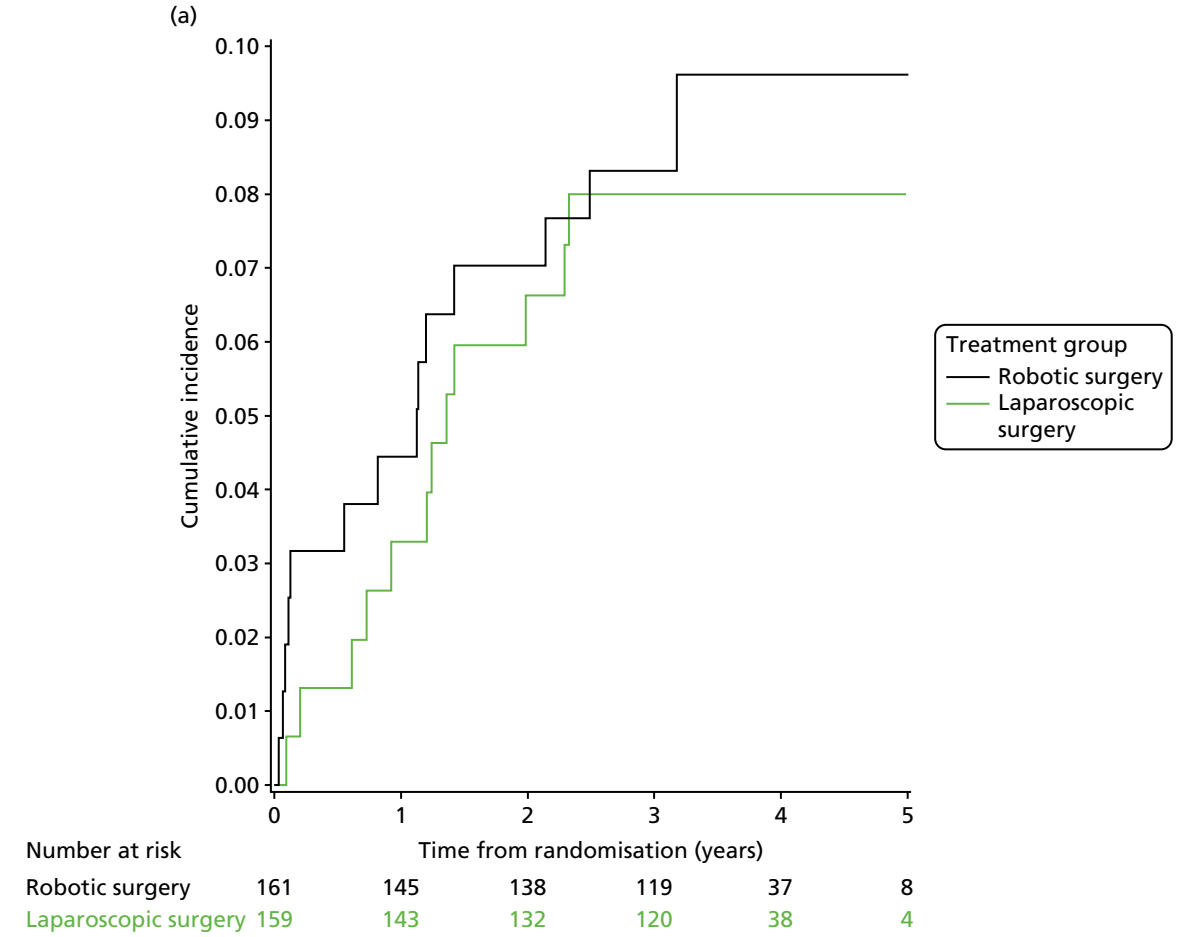


FIGURE 26 Three-year local recurrence by sex. (a) Male; and (b) female. (*continued*)

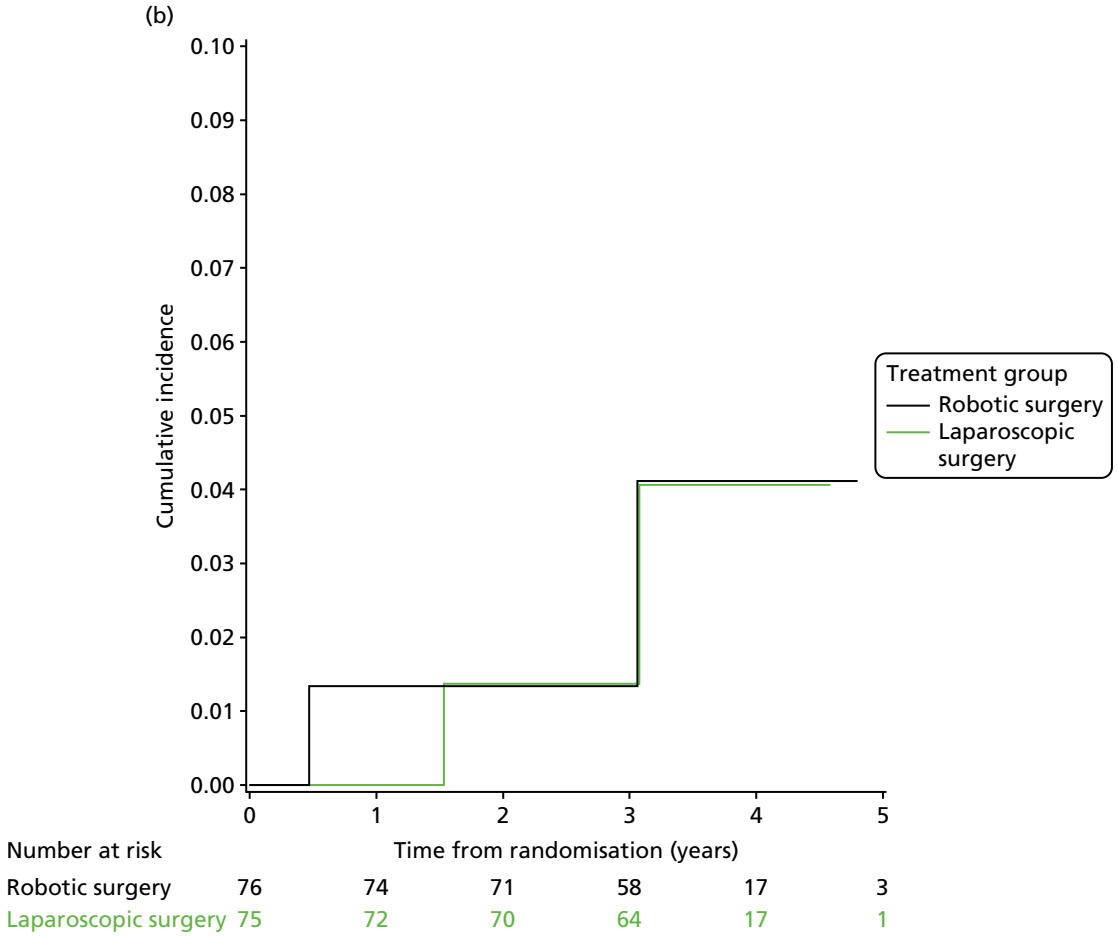


FIGURE 26 Three-year local recurrence by sex. (a) Male; and (b) female.

TABLE 65 Estimated treatment effect HRs for neo-adjuvant therapy

Effect	HR (adjusted 95% CI)[a]	*p*-value	
Treatment in patients who underwent neo-adjuvant therapy: robotic surgery (vs. laparoscopic)	1.233 (0.426 to 3.566)	0.6990	0.8390
Treatment in patients who did not undergo neo-adjuvant therapy: robotic surgery (vs. laparoscopic)	1.061 (0.397 to 2.838)	0.9062	

a Adjusted for BMI class, preoperative radiotherapy, intended procedure and operating surgeon.
Hazard ratios were derived from the treatment term and treatment-by-neo-adjuvant therapy interaction term.

Type of operation

TABLE 66 Estimated treatment effect HRs by operation type

Effect	HR (adjusted 95% CI)[a]	*p*-value	
Treatment in patients who underwent HAR: robotic surgery (vs. laparoscopic)	1.344 (0.121 to 14.957)	0.8089	0.7737
Treatment in patients who underwent LAR: robotic surgery (vs. laparoscopic)	0.975 (0.413 to 2.301)	0.9536	
Treatment in patients who underwent APR: robotic surgery (vs. laparoscopic)	1.924 (0.351 to 10.558)	0.4518	

a Adjusted for BMI class, preoperative radiotherapy, intended procedure and operating surgeon.
Hazard ratios were derived from the treatment term and treatment-by-operation interaction term.

T-stage

TABLE 67 Estimated treatment effect HRs by T-stage

Effect	HR (adjusted 95% CI)[a]	*p*-value	
Treatment in T0 patients: robotic surgery (vs. laparoscopic)[b]			0.9832
Treatment in T1 and T2 patients: robotic surgery (vs. laparoscopic)	1.128 (0.301 to 4.224)	0.8576	
Treatment in T3 and T4 patients: robotic surgery (vs. laparoscopic)	1.019 (0.423 to 2.453)	0.9670	

a Adjusted for BMI class, preoperative radiotherapy, intended procedure and operating surgeon.
b Only one event in T0 patients (in the robotic group) was observed. Within-group comparison between treatment groups was therefore not plausible.
Hazard ratios were derived from the treatment term and treatment-by-T-stage interaction term.

Sex

TABLE 68 Estimated treatment effect HRs by sex

Effect	HR (adjusted 95% CI)[a]	*p*-value	
Treatment in males: robotic surgery (vs. laparoscopic)	1.166 (0.535 to 2.542)	0.6999	0.8794
Treatment in females: robotic surgery (vs. laparoscopic)	0.991 (0.142 to 6.935)	0.9927	

a adjusted for BMI class, preoperative radiotherapy, intended procedure and operating surgeon.
Hazard ratios were derived from the treatment term and treatment-by-sex interaction term.

Appendix 4 Patient self-reported bladder function: further information

Table 69 presents the baseline characteristics of the 351 out of 471 (74.5%) patients who returned questionnaires with sufficient data to derive an I-PSS score. Stratification factors of these 351 patients with complete data are well balanced between the groups, with distributions very similar to the group of patients with missing data (*Table 70*). The I-PSS score at baseline was similar between the groups in these patients, although there was a slightly more positive skew in the robotic group, with a marginally higher mean score indicating slightly higher severity of symptoms at baseline on average.

TABLE 69 Baseline characteristics for complete-case patients with sufficient data to derive I-PSS score by treatment group

Characteristics	Treatment group, *n* (%)		Total (*N* = 351), *n* (%)
	Standard laparoscopic surgery (*N* = 176)	Robotic-assisted laparoscopic surgery (*N* = 175)	
Sex			
Male	116 (65.9)	121 (69.1)	237 (67.5)
Female	60 (34.1)	54 (30.9)	114 (32.5)
BMI classification			
Underweight/normal	69 (39.2)	74 (42.3)	143 (40.7)
Overweight	67 (38.1)	67 (38.3)	134 (38.2)
Obese	40 (22.7)	34 (19.4)	74 (21.1)
Neo-adjuvant therapy			
Yes	81 (46.0)	76 (43.4)	157 (44.7)
No	95 (54.0)	99 (56.6)	194 (55.3)
Intended procedure			
HAR	26 (14.8)	25 (14.3)	51 (14.5)
LAR	119 (67.6)	117 (66.9)	236 (67.2)
APR	31 (17.6)	33 (18.9)	64 (18.2)
Total I-PSS score (baseline)			
Mean (SD)	6.9 (6.91)	8.5 (7.28)	7.7 (7.13)
Median (range)	5.0 (0.0–33.0)	6.0 (0.0–32.0)	6.0 (0.0–33.0)
(Q1, Q3)	(2.0, 9.0)	(3.0, 12.0)	(2.0, 11.0)
Missing	0	0	0
Categorical I-PSS score (baseline)			
Mild	112 (63.6)	101 (57.7)	213 (60.7)
Moderate	50 (28.4)	58 (33.1)	108 (30.8)
Severe	14 (8.0)	16 (9.1)	30 (8.5)

Q1, first interquartile; Q3, third interquartile.

TABLE 70 Baseline characteristics for patients excluded and included in I-PSS analysis

Characteristics	Complete-case analysis, *n* (%)		Total (*N* = 471), *n* (%)
	Excluded (*N* = 120)	Included (*N* = 351)	
Sex			
Male	83 (69.2)	237 (67.5)	320 (67.9)
Female	37 (30.8)	114 (32.5)	151 (32.1)
BMI classification			
Underweight/normal	37 (30.8)	143 (40.7)	180 (38.2)
Overweight	49 (40.8)	134 (38.2)	183 (38.9)
Obese	34 (28.3)	74 (21.1)	108 (22.9)
Neo-adjuvant therapy			
Yes	50 (41.7)	157 (44.7)	207 (43.9)
No	70 (58.3)	194 (55.3)	264 (56.1)
Intended procedure			
HAR	17 (14.2)	51 (14.5)	68 (14.4)
LAR	80 (66.7)	236 (67.2)	316 (67.1)
APR	23 (19.2)	64 (18.2)	87 (18.5)

Appendix 5 Patient self-reported sexual function: males

Table 71 presents the baseline characteristics of the 181 out of 320 (56.6%) male patients who returned questionnaires with sufficient data to derive an IIEF score. Stratification factors of these 181 patients with complete data are well balanced between the groups, with distributions very similar to the group of male patients with missing data (*Table 72*). The IIEF score at baseline was similar between the groups in these patients, although there was a slightly more positive skew in the robotic group, with a marginally higher mean score indicating slightly higher severity of symptoms at baseline on average.

TABLE 71 Patient baseline characteristics for male complete-case patients with sufficient data to derive IIEF score by treatment group

Characteristics	Treatment group, *n* (%)		Total (*N* = 181), *n* (%)
	Standard laparoscopic surgery (*N* = 84)	Robotic-assisted laparoscopic surgery (*N* = 97)	
BMI classification			
Underweight/normal	34 (40.5)	34 (35.1)	68 (37.6)
Overweight	30 (35.7)	43 (44.3)	73 (40.3)
Obese	20 (23.8)	20 (20.6)	40 (22.1)
Neo-adjuvant therapy			
Yes	33 (39.3)	45 (46.4)	78 (43.1)
No	51 (60.7)	52 (53.6)	103 (56.9)
Intended procedure			
HAR	15 (17.9)	11 (11.3)	26 (14.4)
LAR	57 (67.9)	68 (70.1)	125 (69.1)
APR	12 (14.3)	18 (18.6)	30 (16.6)
Total IIEF score (baseline)			
Mean (SD)	37.7 (23.85)	40.1 (24.93)	39.0 (24.40)
Median (range)	37.0 (5.0–72.0)	45.0 (5.0–72.0)	43.0 (5.0–72.0)
(Q1, Q3)	(13.0, 62.0)	(13.0, 65.0)	(13.0, 63.0)

Q1, first interquartile; Q3, third interquartile.

TABLE 72 Baseline characteristics for male patients excluded and included in IIEF analysis

Characteristics	Complete-case analysis, *n* (%)		Total (*N* = 20), *n* (%)
	Excluded (*N* = 139)	Included (*N* = 181)	
BMI classification			
Underweight/normal	53 (38.1)	68 (37.6)	121 (37.8)
Overweight	51 (36.7)	73 (40.3)	124 (38.8)
Obese	35 (25.2)	40 (22.1)	75 (23.4)
Neo-adjuvant therapy			
Yes	62 (44.6)	78 (43.1)	140 (43.8)
No	77 (55.4)	103 (56.9)	180 (56.3)
Intended procedure			
HAR	20 (14.4)	26 (14.4)	46 (14.4)
LAR	87 (62.6)	125 (69.1)	212 (66.3)
APR	32 (23.0)	30 (16.6)	62 (19.4)

TABLE 73 The I-PSS analysis by treatment group at 6 months

	Treatment group, *n* (%)		Total (*N* = 351), *n* (%)
	Standard laparoscopic surgery (*N* = 176)	Robotic-assisted laparoscopic surgery (*N* = 175)	
Total I-PSS score (6 months)			
Mean (SD)	8.0 (7.69)	8.0 (6.81)	8.0 (7.26)
Median (range)	5.0 (0.0–34.0)	7.0 (0.0–34.0)	6.0 (0.0–34.0)
(Q1, Q3)	(2.0, 12.5)	(3.0, 12.0)	(2.0, 12.0)
Categorical I-PSS score (6 months)			
Mild	109 (61.9)	99 (56.6)	208 (59.3)
Moderate	50 (28.4)	67 (38.3)	117 (33.3)
Severe	17 (9.7)	9 (5.1)	26 (7.4)

Q1, first interquartile; Q3, third interquartile.

Appendix 6 Patient self-reported sexual function: females

Table 74 presents the baseline characteristics of the 54 out of 151 (35.8%) female patients who returned questionnaires with sufficient data to derive a FSFI score. Stratification factors of these 54 patients with complete data are well balanced between the groups, with distributions similar to the group of female patients with missing data, with the exception of the higher rate of neo-adjuvant therapy in the patients with complete data and the higher proportion of underweight/normal patients (*Table 73*). The FSFI score at baseline was marginally lower in the robotic group, indicating slightly worse function at baseline for female patients in the robotic group who were included in the analysis.

TABLE 74 Baseline characteristics for female complete-case patients with sufficient data to derive FSFI score by treatment group

Charateristics	Treatment group, *n* (%)		Total (*N* = 54), *n* (%)
	Standard laparoscopic surgery (*N* = 29)	Robotic-assisted laparoscopic surgery (*N* = 25)	
BMI classification			
Underweight/normal	15 (51.7)	12 (48.0)	27 (50.0)
Overweight	10 (34.5)	8 (32.0)	18 (33.3)
Obese	4 (13.8)	5 (20.0)	9 (16.7)
Neo-adjuvant therapy			
Yes	15 (51.7)	16 (64.0)	31 (57.4)
No	14 (48.3)	9 (36.0)	23 (42.6)
Intended procedure			
HAR	6 (20.7)	5 (20.0)	11 (20.4)
LAR	18 (62.1)	15 (60.0)	33 (61.1)
APR	5 (17.2)	5 (20.0)	10 (18.5)
FSFI (baseline)			
Mean (SD)	16.7 (11.74)	14.8 (9.96)	15.8 (10.90)
Median (range)	19.1 (2.0–34.2)	14.8 (2.8–30.1)	16.5 (2.0–34.2)
(Q1, Q3)	(4.4, 28.2)	(5.4, 22.7)	(4.7, 27.3)
Missing	0	0	0

Q1, first interquartile; Q3, third interquartile.

TABLE 75 Baseline characteristics for female patients excluded and included in the complete-case analysis

	Complete-case analysis, *n* (%)		
Charateristics	Excluded (*N* = 97)	Included (*N* = 54)	Total (*N* = 54), *n* (%)
BMI classification			
Underweight/normal	32 (33.0)	27 (50.0)	59 (39.1)
Overweight	41 (42.3)	18 (33.3)	59 (39.1)
Obese	24 (24.7)	9 (16.7)	33 (21.9)
Neo-adjuvant therapy			
Yes	36 (37.1)	31 (57.4)	67 (44.4)
No	61 (62.9)	23 (42.6)	84 (55.6)
Intended procedure			
HAR	11 (11.3)	11 (20.4)	22 (14.6)
LAR	71 (73.2)	33 (61.1)	104 (68.9)
APR	15 (15.5)	10 (18.5)	25 (16.6)

TABLE 76 Female Sexual Function Index at 6 months for patients included in analysis, by treatment group

	Treatment group, *n* (%)		
FSFI (6 months)	Standard laparoscopic surgery (*N* = 29)	Robotic- assisted laparoscopic surgery (*N* = 25)	Total (*N* = 54), *n* (%)
Mean (SD)	16.7 (11.25)	14.2 (10.34)	15.5 (10.81)
Median (range)	16.7 (2.0–33.6)	8.5 (2.9–31.4)	16.1 (2.0–33.6)
(Q1, Q3)	(4.5, 28.3)	(5.2, 24.4)	(5.1, 25.9)
Missing	0	0	0

Q1, first interquartile; Q3, third interquartile.

Appendix 7 Patient-reported generic health

Table 77 presents the baseline characteristics of the 459 out of 471 (97.5%) patients who returned at least one questionnaire with sufficient data to derive a PCS/MCS score. Stratification factors and ASA grades of these 459 patients are well balanced between the groups.

Tables 78 and *79* show the PCS/MCS at baseline, at 30 days post surgery and at 6 months post surgery in the two groups. The baseline PCS and MCS were similar in the two treatment groups.

TABLE 77 Patient baseline characteristics for patients with sufficient data to derive a PCS/MCS score

Characteristics	Treatment group, *n* (%)		Total (*N* = 459), *n* (%)
	Robotic-assisted laparoscopic surgery (*N* = 232)	Standard laparoscopic surgery (*N* = 227)	
Sex			
Male	157 (67.7)	153 (67.4)	310 (67.5)
Female	75 (32.3)	74 (32.6)	149 (32.5)
Neo-adjuvant therapy			
Yes	106 (45.7)	100 (44.1)	206 (44.9)
No	126 (54.3)	127 (55.9)	253 (55.1)
Intended procedure			
HAR	34 (14.7)	33 (14.5)	67 (14.6)
LAR	155 (66.8)	154 (67.8)	309 (67.3)
APR	43 (18.5)	40 (17.6)	83 (18.1)
BMI classification			
Underweight/normal	92 (39.7)	85 (37.4)	177 (38.6)
Overweight	89 (38.4)	88 (38.8)	177 (38.6)
Obese	51 (22.0)	54 (23.8)	105 (22.9)
ASA classification			
A normal healthy patient	39 (16.8)	51 (22.5)	90 (19.6)
A patient with mild systemic disease	147 (63.4)	123 (54.2)	270 (58.8)
A patient with severe systemic disease	46 (19.8)	52 (22.9)	98 (21.4)
A patient with severe systemic disease that is a constant threat to life	0 (0.0)	1 (0.4)	1 (0.2)

TABLE 78 Physical component score by treatment group at baseline, at 30 days and at 6 months

PCS	Treatment group, *n* (%)		
	Robotic-assisted laparoscopic surgery (*N* = 232)	Standard laparoscopic surgery (*N* = 227)	Total (*N* = 459), *n* (%)
Baseline			
Mean (SD)	51.4 (8.90)	51.6 (8.79)	51.5 (8.84)
Median (range)	53.7 (24.8–67.4)	53.7 (24.2–67.4)	53.7 (24.2–67.4)
Missing	6	6	12
Number	226	221	447
30 days			
Mean (SD)	42.4 (8.55)	42.0 (8.42)	42.2 (8.48)
Median (range)	42.3 (22.8–61.7)	42.1 (24.3–63.3)	42.2 (22.8–63.3)
Missing	19	29	48
Number	213	198	411
6 months			
Mean (SD)	48.7 (7.95)	48.3 (8.90)	48.5 (8.43)
Median (range)	49.7 (27.4–61.2)	50.2 (18.9–63.2)	50.0 (18.9–63.2)
Missing	33	32	65
Number	199	195	394

TABLE 79 Mental component score by treatment group at baseline, at 30 days and at 6 months

MCS	Treatment group, *n* (%)		
	Robotic-assisted laparoscopic surgery (*N* = 232)	Standard laparoscopic surgery (*N* = 227)	Total (*N* = 459), *n* (%)
Baseline			
Mean (SD)	47.3 (11.82)	48.1 (11.48)	47.7 (11.65)
Median (range)	50.5 (13.6–67.0)	50.9 (10.3–65.8)	50.8 (10.3–67.0)
Missing	6	6	12
Number	226	221	447
30 days			
Mean (SD)	45.6 (11.73)	44.1 (12.86)	44.8 (12.30)
Median (range)	46.4 (12.9–64.5)	46.3 (7.2–68.0)	46.4 (7.2–68.0)
Missing	19	29	48
Number	213	198	411
6 months			
Mean (SD)	48.9 (11.62)	49.6 (10.04)	49.2 (10.85)
Median (range)	52.3 (10.8–66.3)	51.9 (20.1–67.2)	52.1 (10.8–67.2)
Missing	33	31	64
Number	199	196	395

TABLE 80 Patient baseline characteristics for those included and not included in PCS/MCS analysis

Characteristics	Inclusion, *n* (%)		Total (*N* = 471), *n* (%)
	Not included (*N* = 12)	Included (*N* = 459)	
Sex			
Male	10 (83.3)	310 (67.5)	320 (67.9)
Female	2 (16.7)	149 (32.5)	151 (32.1)
Neo-adjuvant therapy			
Yes	6 (50.0)	206 (44.9)	212 (45.0)
No	6 (50.0)	253 (55.1)	259 (55.0)
Intended procedure			
HAR	2 (16.7)	67 (14.6)	69 (14.6)
LAR	8 (66.7)	309 (67.3)	317 (67.3)
APR	2 (16.7)	83 (18.1)	85 (18.0)
BMI classification			
Underweight/normal	3 (25.0)	177 (38.6)	180 (38.2)
Overweight	6 (50.0)	177 (38.6)	183 (38.9)
Obese	3 (25.0)	105 (22.9)	108 (22.9)
ASA classification			
A normal healthy patient	1 (8.3)	90 (19.6)	91 (19.3)
A patient with mild systemic disease	4 (33.3)	270 (58.8)	274 (58.2)
A patient with severe systemic disease	0 (0.0)	98 (21.4)	98 (20.8)
A patient with severe systemic disease that is a constant threat to life	0 (0.0)	1 (0.2)	1 (0.2)
Missing	7[a] (58.3)	0 (0.0)	7 (1.5)

a The remaining five patients were not included because they did not have a PCS/MCS score (the same patients are missing PCS and MCS scores).

Appendix 8 Patient self-reported fatigue

Table 81 presents the baseline characteristics of the 440 out of 471 (93.4%) patients who returned at least one questionnaire with sufficient data to derive a score for at least one of the scales. Stratification factors and ASA grades for these 440 patients are well balanced between the groups.

Tables 82–86 show each of the scales at baseline, at 30 days post surgery and at 6 months post surgery in the two groups. The baseline scores were similar between the two treatment groups for all five scales.

Tables 88–92 show the fitted model estimates for each of the five scales.

TABLE 81 Baseline characteristics for patients included in fatigue analysis

Characteristics	Treatment group, *n* (%)		Total (*N* = 440), *n* (%)
	Robotic-assisted laparoscopic surgery (*N* = 222)	Standard laparoscopic surgery (*N* = 218)	
Sex			
Male	150 (67.6)	146 (67.0)	296 (67.3)
Female	72 (32.4)	72 (33.0)	144 (32.7)
Neo-adjuvant therapy			
Yes	104 (46.8)	94 (43.1)	198 (45.0)
No	118 (53.2)	124 (56.9)	242 (55.0)
Intended procedure			
HAR	34 (15.3)	33 (15.1)	67 (15.2)
LAR	145 (65.3)	145 (66.5)	290 (65.9)
APR	43 (19.4)	40 (18.3)	83 (18.9)
BMI classification			
Underweight/normal	85 (38.3)	77 (35.3)	162 (36.8)
Overweight	86 (38.7)	87 (39.9)	173 (39.3)
Obese	51 (23.0)	54 (24.8)	105 (23.9)
ASA classification			
A normal healthy patient	36 (16.2)	47 (21.6)	83 (18.9)
A patient with mild systemic disease	143 (64.4)	121 (55.5)	264 (60.0)
A patient with severe systemic disease	43 (19.4)	49 (22.5)	92 (20.9)
A patient with severe systemic disease that is a constant threat to life	0 (0.0)	1 (0.5)	1 (0.2)

TABLE 82 General fatigue by treatment group at baseline, at 30 days and at 6 months

General fatigue	Treatment group		Total (*N* = 440)
	Robotic-assisted laparoscopic surgery (*N* = 222)	Standard laparoscopic surgery (*N* = 218)	
Baseline			
Mean (SD)	10.4 (4.73)	10.1 (4.49)	10.3 (4.61)
Median (range)	10.0 (4.0–20.0)	10.0 (4.0–20.0)	10.0 (4.0–20.0)
Missing	13	11	24
Number	209	207	416
30 days			
Mean (SD)	12.5 (4.31)	12.6 (4.36)	12.6 (4.33)
Median (range)	12.0 (4.0–20.0)	13.0 (4.0–20.0)	13.0 (4.0–20.0)
Missing	31	37	68
Number	191	181	372
6 months			
Mean (SD)	11.0 (4.62)	10.8 (4.41)	10.9 (4.51)
Median (range)	11.0 (4.0–20.0)	11.0 (4.0–20.0)	11.0 (4.0–20.0)
Missing	36	41	77
Number	186	177	363

TABLE 83 Physical fatigue by treatment group at baseline, at 30 days and at 6 months

Physical fatigue	Treatment group		Total (*N* = 440)
	Robotic-assisted laparoscopic surgery (*N* = 222)	Standard laparoscopic surgery (*N* = 218)	
Baseline			
Mean (SD)	10.1 (4.64)	9.5 (4.52)	9.8 (4.58)
Median (range)	9.0 (4.0–20.0)	9.0 (4.0–20.0)	9.0 (4.0–20.0)
Missing	12	12	24
Number	210	206	416
30 days			
Mean (SD)	12.5 (4.42)	13.1 (4.44)	12.8 (4.43)
Median (range)	12.5 (4.0–20.0)	13.0 (4.0–20.0)	13.0 (4.0–20.0)
Missing	28	31	59
Number	194	187	381
6 months			
Mean (SD)	10.7 (4.09)	10.9 (4.55)	10.8 (4.32)
Median (range)	11.0 (4.0–20.0)	10.5 (4.0–20.0)	11.0 (4.0–20.0)
Missing	39	36	75
Number	183	182	365

TABLE 84 Reduced activity by treatment group at baseline, at 30 days and at 6 months

Reduced activity	Treatment group		Total (*N* = 440)
	Robotic-assisted laparoscopic surgery (*N* = 222)	Standard laparoscopic surgery (*N* = 218)	
Baseline			
Mean (SD)	9.9 (4.44)	9.9 (4.40)	9.9 (4.42)
Median (range)	9.5 (4.0–20.0)	9.0 (4.0–20.0)	9.0 (4.0–20.0)
Missing	14	17	31
Number	208	201	409
30 days			
Mean (SD)	12.7 (4.31)	13.1 (4.33)	12.9 (4.32)
Median (range)	13.0 (4.0–20.0)	13.0 (4.0–20.0)	13.0 (4.0–20.0)
Missing	30	37	67
Number	192	181	373
6 months			
Mean (SD)	10.6 (4.23)	10.5 (4.20)	10.6 (4.21)
Median (range)	10.0 (4.0–20.0)	11.0 (4.0–20.0)	10.0 (4.0–20.0)
Missing	38	32	70
Number	184	186	370

TABLE 85 Reduced motivation by treatment group at baseline, at 30 days and at 6 months

Reduced motivation	Treatment group		Total (*N* = 440)
	Robotic-assisted laparoscopic surgery (*N* = 222)	Standard laparoscopic surgery (*N* = 218)	
Baseline			
Mean (SD)	8.5 (3.56)	8.5 (3.58)	8.5 (3.56)
Median (range)	8.0 (4.0–18.0)	8.0 (4.0–20.0)	8.0 (4.0–20.0)
Missing	14	14	28
Number	208	204	412
30 days			
Mean (SD)	9.7 (3.89)	10.2 (3.75)	9.9 (3.82)
Median (range)	9.0 (4.0–20.0)	10.0 (4.0–20.0)	10.0 (4.0–20.0)
Missing	33	34	67
Number	189	184	373
6 months			
Mean (SD)	8.5 (3.39)	8.7 (3.47)	8.6 (3.43)
Median (range)	8.0 (4.0–18.0)	9.0 (4.0–20.0)	8.0 (4.0–20.0)
Missing	45	40	85
Number	177	178	355

TABLE 86 Mental fatigue by treatment group at baseline, at 30 days and at 6 months

Mental fatigue	Treatment group		Total (*N* = 440)
	Robotic-assisted laparoscopic surgery (*N* = 222)	Standard laparoscopic surgery (*N* = 218)	
Baseline			
Mean (SD)	8.5 (4.12)	8.4 (4.38)	8.4 (4.25)
Median (range)	8.0 (4.0–20.0)	7.0 (4.0–20.0)	8.0 (4.0–20.0)
Missing	14	12	26
Number	208	206	414
30 days			
Mean (SD)	9.0 (4.33)	9.4 (4.37)	9.2 (4.35)
Median (range)	9.0 (4.0–20.0)	9.0 (4.0–20.0)	9.0 (4.0–20.0)
Missing	30	37	67
Number	192	181	373
6 months			
Mean (SD)	8.3 (3.90)	8.5 (3.84)	8.4 (3.87)
Median (range)	8.0 (4.0–20.0)	8.0 (4.0–19.0)	8.0 (4.0–20.0)
Missing	37	34	71
Number	185	184	369

TABLE 87 Baseline characteristics for patients not included and included in the fatigue analysis

Characteristics	Inclusion, *n* (%)		
	Not included (*N* = 31)	Included (*N* = 440)	Total (*N* = 471), *n* (%)
Sex			
Male	24 (77.4)	296 (67.3)	320 (67.9)
Female	7 (22.6)	144 (32.7)	151 (32.1)
Neo-adjuvant therapy			
Yes	14 (45.2)	198 (45.0)	212 (45.0)
No	17 (54.8)	242 (55.0)	259 (55.0)
Intended procedure			
HAR	2 (6.5)	67 (15.2)	69 (14.6)
LAR	27 (87.1)	290 (65.9)	317 (67.3)
APR	2 (6.5)	83 (18.9)	85 (18.0)
BMI classification			
Underweight/normal	18 (58.1)	162 (36.8)	180 (38.2)
Overweight	10 (32.3)	173 (39.3)	183 (38.9)
Obese	3 (9.7)	105 (23.9)	108 (22.9)
ASA classification			
A normal healthy patient	8 (25.8)	83 (18.9)	91 (19.3)
A patient with mild systemic disease	10 (32.3)	264 (60.0)	274 (58.2)
A patient with severe systemic disease	6 (19.4)	92 (20.9)	98 (20.8)
A patient with severe systemic disease that is a constant threat to life	0 (0.0)	1 (0.2)	1 (0.2)
Missing	7 (22.6)	0 (0.0)	7 (1.5)

TABLE 88 Results for statistical analysis for general fatigue

Effect	Estimate	Standard error	Pr > |t|	95% CI
Intercept	9.5014	0.8549		
30 days	1.5634	0.7257	0.0316	0.1385 to 2.9882
6 months	–0.2130	0.7108	0.7645	–1.6085 to 1.1824
Treatment: robotic-assisted laparoscopic surgery (vs. standard)	0.2517	0.4320	0.5603	–0.5965 to 1.0999
Sex: female (vs. male)	1.0912	0.3728	0.0035	0.3592 to 1.8232
Neo-adjuvant therapy: no (vs. yes)	–1.0316	0.4638	0.0264	–1.9421 to –0.1210
Intended procedure: APR (vs. HAR)	1.4207	0.6392	0.0266	0.1657 to 2.6758
Intended procedure: LAR (vs. HAR)	0.6583	0.5147	0.2013	–0.3522 to 1.6689
BMI class: obese (vs. underweight/normal)	–0.1142	0.5732	0.8421	–1.2397 to 1.0113
BMI class: overweight (vs. underweight/normal)	–0.5285	0.4972	0.2882	–1.5047 to 0.4477
ASA grade: (II vs. I)	0.09524	0.5837	0.8704	–1.0509 to 1.2413
ASA grade: (III vs. I)	1.4558	0.7238	0.0447	0.03480 to 2.8769
Robotic-assisted laparoscopic surgery and 30-day interaction	–0.3385	0.4715	0.4730	–1.2642 to 0.5872
Robotic-assisted laparoscopic surgery and 6-month interaction	–0.2278	0.4758	0.6322	–1.1619 to 0.7063
ASA grade II and 30-day interaction	–0.2271	0.6349	0.7207	–1.4736 to 1.0194
ASA grade III and 30-day interaction	–0.9397	0.7672	0.2211	–2.4461 to 0.5666
ASA grade II and 6-month interaction	0.5650	0.6239	0.3655	–0.6599 to 1.7899
ASA grade III and 6-month interaction	–1.0492	0.7621	0.1691	–2.5455 to 0.4472
No neo-adjuvant therapy and 30-day interaction	1.4474	0.4810	0.0027	0.5030 to 2.3917
No neo-adjuvant therapy and 6-month interaction	0.8010	0.4859	0.0997	–0.1531 to 1.7551
Obese and 30-day interaction	1.4919	0.6187	0.0162	0.2771 to 2.7067
Obese and 6-month interaction	0.9778	0.6335	0.1232	–0.2661 to 2.2216
Overweight and 30-day interaction	0.4738	0.5378	0.3787	–0.5822 to 1.5297
Overweight and 6-month interaction	0.8162	0.5383	0.1299	–0.2408 to 1.8731

TABLE 89 Results of statistical analysis for physical fatigue

Effect	Estimate	Standard error	Pr > |t|	95% CI
Intercept	7.8866	0.8527		
30 days	3.8465	0.7294	< 0001	2.4146 to 5.2785
6 months	0.6665	0.7218	0.3561	–0.7505 to 2.0836
Treatment: robotic-assisted laparoscopic surgery (vs. standard)	0.3964	0.4262	0.3527	–0.4404 to 1.2332
Sex: female (vs. male)	0.7560	0.3635	0.0379	0.04242 to 1.4696
Neo-adjuvant therapy: no (vs. yes)	–0.5301	0.4614	0.2510	–1.4359 to 0.3758
Intended procedure: APR (vs. HAR)	1.6339	0.6251	0.0091	0.4066 to 2.8611
Intended procedure: LAR (vs. HAR)	1.1461	0.5021	0.0228	0.1603 to 2.1320
BMI class: obese (vs. underweight/normal)	–0.2618	0.5675	0.6447	–1.3761 to 0.8524
BMI class: overweight (vs. underweight/normal)	–0.8362	0.4923	0.0898	–1.8027 to 0.1302
ASA grade: (II vs. I)	0.9504	0.5846	0.1045	–0.1974 to 2.0983
ASA grade: (III vs. I)	2.0824	0.7252	0.0042	0.6586 to 3.5062
Robotic-assisted laparoscopic surgery and 30-day interaction	–0.9425	0.4749	0.0476	–1.8749 to –0.01009
Robotic-assisted laparoscopic surgery and 6-month interaction	–0.7351	0.4815	0.1273	–1.6804 to 0.2102
ASA grade II and 30-day interaction	–1.7434	0.6408	0.0067	–3.0015 to –0.4853
ASA grade III and 30-day interaction	0.003981	0.6358	0.9950	–1.2443 to 1.2522
ASA grade II and 6-month interaction	–1.6193	0.7721	0.0363	–3.1353 to –0.1034
ASA grade III and 6-month interaction	–1.1335	0.7782	0.1457	–2.6613 to 0.3943
Obese and 30-day interaction	0.9738	0.6242	0.1192	–0.2517 to 2.1993
Obese and 6-month interaction	1.4096	0.6396	0.0279	0.1539 to 2.6654
Overweight and 30-day interaction	0.6666	0.5417	0.2189	–0.3970 to 1.7302
Overweight and 6-month interaction	0.8577	0.5480	0.1180	–0.2182 to 1.9336
No neo-adjuvant therapy and 30-day interaction	1.0634	0.4833	0.0281	0.1146 to 2.0123
No neo-adjuvant therapy and 6-month interaction	0.7799	0.4919	0.1133	–0.1858 to 1.7455

TABLE 90 Results of statistical analysis for reduced activity

Effect	Estimate	Standard error	Pr > |t|	95% CI
Intercept	8.4917	0.8260		
30 days	3.0550	0.7251	< 0001	1.6314 to 4.4786
6 months	–0.4520	0.7075	0.5231	–1.8410 to 0.9370
Treatment: robotic-assisted laparoscopic surgery (vs. standard)	–0.1634	0.4148	0.6938	–0.9777 to 0.6510
Sex: female (vs. male)	0.7244	0.3482	0.0378	0.04081 to 1.4080
Neo-adjuvant therapy: no (vs. yes)	–0.7325	0.4477	0.1022	–1.6115 to 0.1465
Intended procedure: APR (vs. HAR)	2.0437	0.6001	0.0007	0.8654 to 3.2221
Intended procedure: LAR (vs. HAR)	1.0258	0.4829	0.0340	0.07763 to 1.9739
BMI class: obese (vs. underweight/normal)	–0.4832	0.5515	0.3813	–1.5659 to 0.5996
BMI class: overweight (vs. underweight/normal)	–0.5767	0.4777	0.2277	–1.5146 to 0.3611
ASA grade: (II vs. I)	0.8395	0.5703	0.1415	–0.2802 to 1.9591
ASA grade: (III vs. I)	1.7121	0.7019	0.0150	0.3341 to 3.0901
Robotic-assisted laparoscopic surgery and 30-day interaction	–0.3080	0.4725	0.5147	–1.2356 to 0.6196
Robotic-assisted laparoscopic surgery and 6-month interaction	0.08269	0.4743	0.8616	–0.8485 to 1.0139
ASA grade II and 30-day interaction	–1.0403	0.6383	0.1036	–2.2935 to 0.2130
ASA grade III and 30-day interaction	0.3085	0.6239	0.6211	–0.9164 to 1.5334
ASA grade II and 6-month interaction	–0.7610	0.7664	0.3211	–2.2658 to 0.7437
ASA grade III and 6-month interaction	–0.4147	0.7528	0.5819	–1.8927 to 1.0633
Obese and 30-day interaction	1.5630	0.6212	0.0121	0.3434 to 2.7825
Obese and 6-month interaction	1.1287	0.6275	0.0725	–0.1033 to 2.3608
Overweight and 30-day interaction	0.4547	0.5363	0.3969	–0.5983 to 1.5077
Overweight and 6-month interaction	0.4901	0.5386	0.3631	–0.5673 to 1.5476
No neo-adjuvant therapy and 30-day interaction	0.9029	0.4786	0.0596	–0.03670 to 1.8424
No neo-adjuvant therapy and 6-month interaction	1.1007	0.4822	0.0228	0.1539 to 2.0475

TABLE 91 Results of statistical analysis for reduced motivation

Effect	Estimate	Standard error	Pr > |t|	95% CI
Intercept	8.1637	0.6844		
30 days	0.9350	0.5761	0.1051	–0.1961 to 2.0662
6 months	–0.3237	0.5658	0.5675	–1.4347 to 0.7873
Treatment: robotic-assisted laparoscopic surgery (vs. standard)	–0.03917	0.3531	0.9117	–0.7324 to 0.6540
Sex: female (vs. male)	0.7819	0.3053	0.0107	0.1824 to 1.3813
Neo-adjuvant therapy: no (vs. yes)	–0.4394	0.3716	0.2374	–1.1689 to 0.2902
Intended procedure: APR (vs. HAR)	1.0636	0.5227	0.0422	0.03734 to 2.0898
Intended procedure: LAR (vs. HAR)	0.4422	0.4242	0.2976	–0.3907 to 1.2750
BMI class: obese (vs. underweight/normal)	–0.03125	0.4627	0.9462	–0.9398 to 0.8773
BMI class: overweight (vs. underweight/normal)	–0.1524	0.4059	0.7075	–0.9493 to 0.6445
ASA grade: (II vs. I)	–0.1479	0.4712	0.7536	–1.0731 to 0.7773
ASA grade: (III vs. I)	0.03344	0.5745	0.9536	–1.0945 to 1.1614
Robotic-assisted laparoscopic surgery and 30-day interaction	–0.4805	0.3794	0.2057	–1.2253 to 0.2644
Robotic-assisted laparoscopic surgery and 6-month interaction	–0.3817	0.3853	0.3222	–1.1381 to 0.3748
ASA grade II and 30-day interaction	–0.08389	0.5096	0.8693	–1.0845 to 0.9167
ASA grade II and 6-month interaction	0.8070	0.5026	0.1088	–0.1799 to 1.7939
ASA grade III and 30-day interaction	0.3101	0.6122	0.6127	–0.8920 to 1.5121
ASA grade III and 6-month interaction	0.5378	0.6112	0.3792	–0.6622 to 1.7379
Obese and 30-day interaction	0.9059	0.4994	0.0701	–0.07469 to 1.8864
Obese and 6-month interaction	0.4964	0.5084	0.3292	–0.5018 to 1.4945
Overweight and 30-day interaction	0.1694	0.4309	0.6944	–0.6768 to 1.0155
Overweight and 6-month interaction	–0.09029	0.4392	0.8372	–0.9526 to 0.7720
No neo-adjuvant therapy and 30-day interaction	0.8648	0.3857	0.0253	0.1075 to 1.6221
No neo-adjuvant therapy and 6-month interaction	0.1111	0.3935	0.7777	–0.6614 to 0.8836

TABLE 92 Results of statistical analysis for mental fatigue

Effect	Estimate	Standard error	Pr > |t|	95% CI
Intercept	8.3964	0.7944		
30 days	0.8814	0.5758	0.1263	–0.2491 to 2.0118
6 months	–0.7408	0.5652	0.1904	–1.8506 to 0.3689
Treatment: robotic-assisted laparoscopic surgery (vs. standard)	0.1374	0.4075	0.7360	–0.6626 to 0.9374
Sex: female (vs. male)	0.8899	0.3613	0.0140	0.1807 to 1.5992
Neo-adjuvant therapy: no (vs. yes)	–0.1726	0.3656	0.6370	–0.8905 to 0.5453
Intended procedure: APR (vs. HAR)	1.1727	0.6192	0.0586	–0.04292 to 2.3884
Intended procedure: LAR (vs. HAR)	0.6569	0.5025	0.1915	–0.3296 to 1.6435
BMI class: obese (vs. underweight/normal)	–0.6885	0.5392	0.2021	–1.7471 to 0.3702
BMI class: overweight (vs. underweight/normal)	–0.1299	0.4685	0.7817	–1.0497 to 0.7899
ASA grade: (II vs. I)	–0.8621	0.5458	0.1147	–1.9336 to 0.2095
ASA grade: (III vs. I)	–0.5135	0.6657	0.4407	–1.8205 to 0.7935
Robotic-assisted laparoscopic surgery and 30-day interaction	–0.4997	0.4152	0.2291	–1.3148 to 0.3154
Robotic-assisted laparoscopic surgery and 6-month interaction	–0.5284	0.4171	0.2056	–1.3473 to 0.2905
ASA grade II and 30-day interaction	–0.06827	0.5569	0.9025	–1.1615 to 1.0250
ASA grade II and 6-month interaction	0.9765	0.5473	0.0748	–0.09799 to 2.0510
ASA grade III and 30-day interaction	–0.4344	0.6625	0.5123	–1.7351 to 0.8664
ASA grade III and 6-month interaction	0.1242	0.6547	0.8496	–1.1612 to 1.4096
Obese and 30-day interaction	1.0461	0.5460	0.0558	–0.02591 to 2.1181
Obese and 6-month interaction	1.1519	0.5494	0.0364	0.07317 to 2.2307
Overweight and 30-day interaction	0.1229	0.4735	0.7952	–0.8068 to 1.0527
Overweight and 6-month interaction	0.5304	0.4753	0.2648	–0.4027 to 1.4636

Appendix 9 Disease-free survival: further information

Disease-free survival events and censorings, including reason for censoring, are summarised in *Table 93*.

Types of events (i.e. what event contributed to the DFS event) are summarised in *Table 94*.

The methods of confirmation are summarised in *Table 95*.

Table 96 shows Kaplan–Meier estimates of DFS by treatment group at several time points (1–5 years post randomisation).

Subgroup analyses: further information

Figures 27–30 display the Kaplan-Meier graphs for the effect of neo-adjuvant therapy, operation type, T-stage and sex, on 3-year DFS.

TABLE 93 Disease-free survival events and censoring, by treatment group

Nature of the end of follow-up for DFS analysis	Treatment group, *n* (%)		Total (*N* = 471), *n* (%)
	Standard laparoscopic surgery (*N* = 234)	Robotic-assisted laparoscopic surgery (*N* = 237)	
Event	56 (23.9)	58 (24.5)	114 (24.2)
Censor: last known to be alive and disease-free	169 (72.2)	172 (72.6)	341 (72.4)
Censor: withdrawal from further data collection	5 (2.1)	4 (1.7)	9 (1.9)
Censor: non-standard circumstance[a]	4 (1.7)	3 (1.3)	7 (1.5)

a Three patients had benign disease, three patients had a non-curative surgery outcome and one patient did not undergo surgery.

TABLE 94 Explanation for DFS events, by treatment group

Type of (first) recurrence or death (by treatment group)	Treatment group, *n* (%)		Total (*N* = 114), *n* (%)
	Standard laparoscopic surgery (*N* = 56)	Robotic-assisted laparoscopic surgery (*N* = 58)	
Locoregional spread	13 (23.2)	14 (24.1)	27 (23.7)
Liver metastasis	15 (26.8)	6 (10.3)	21 (18.4)
Lung metastasis	10 (17.9)	15 (25.9)	25 (21.9)
New primary cancer	8 (14.3)	10 (17.2)	18 (15.8)
Death	9 (16.1)	8 (13.8)	17 (14.9)
Other[a]	1 (1.8)	5 (8.6)	6 (5.3)

a 'Other' recurrences were: peritoneal carcinomatosis, early peritoneal disease, extramural vascular invasion, bone metastasis, skeleton and adrenal gland.

Note

This is a summary of just the first recurrence event for patients who had a recurrence (i.e. they are the events that contributed to the analysis of DFS). These patients may have had additional recurrences that are not included in this table.

TABLE 95 Methods of confirmation of DFS event, by treatment group

Method of confirmation of recurrences	Treatment group, *n* (%)		Total (*N* = 73), *n* (%)
	Standard laparoscopic surgery (*N* = 38)	Robotic-assisted laparoscopic surgery (*N* = 35)	
Clinical	1 (2.6)	3 (8.6)	4 (5.5)
Radiological	26 (68.4)	24 (68.6)	50 (68.5)
Pathological	10 (26.3)	8 (22.9)	18 (24.7)
Positron emission tomography scan	1 (2.6)	0 (0.0)	1 (1.4)

TABLE 96 Kaplan–Meier estimates of DFS, by treatment group

Time (years)	Treatment group			
	Laparoscopic surgery		Robotic surgery	
	Probability of a DFS event	95% CI	Probability of a DFS event	95% CI
1	0.084	0.048 to 0.121	0.116	0.075 to 0.157
2	0.183	0.132 to 0.234	0.186	0.136 to 0.236
3	0.220	0.165 to 0.274	0.225	0.171 to 0.279
4	0.259	0.193 to 0.324	0.274	0.210 to 0.338
5	0.390	0.192 to 0.588	0.274	0.210 to 0.338

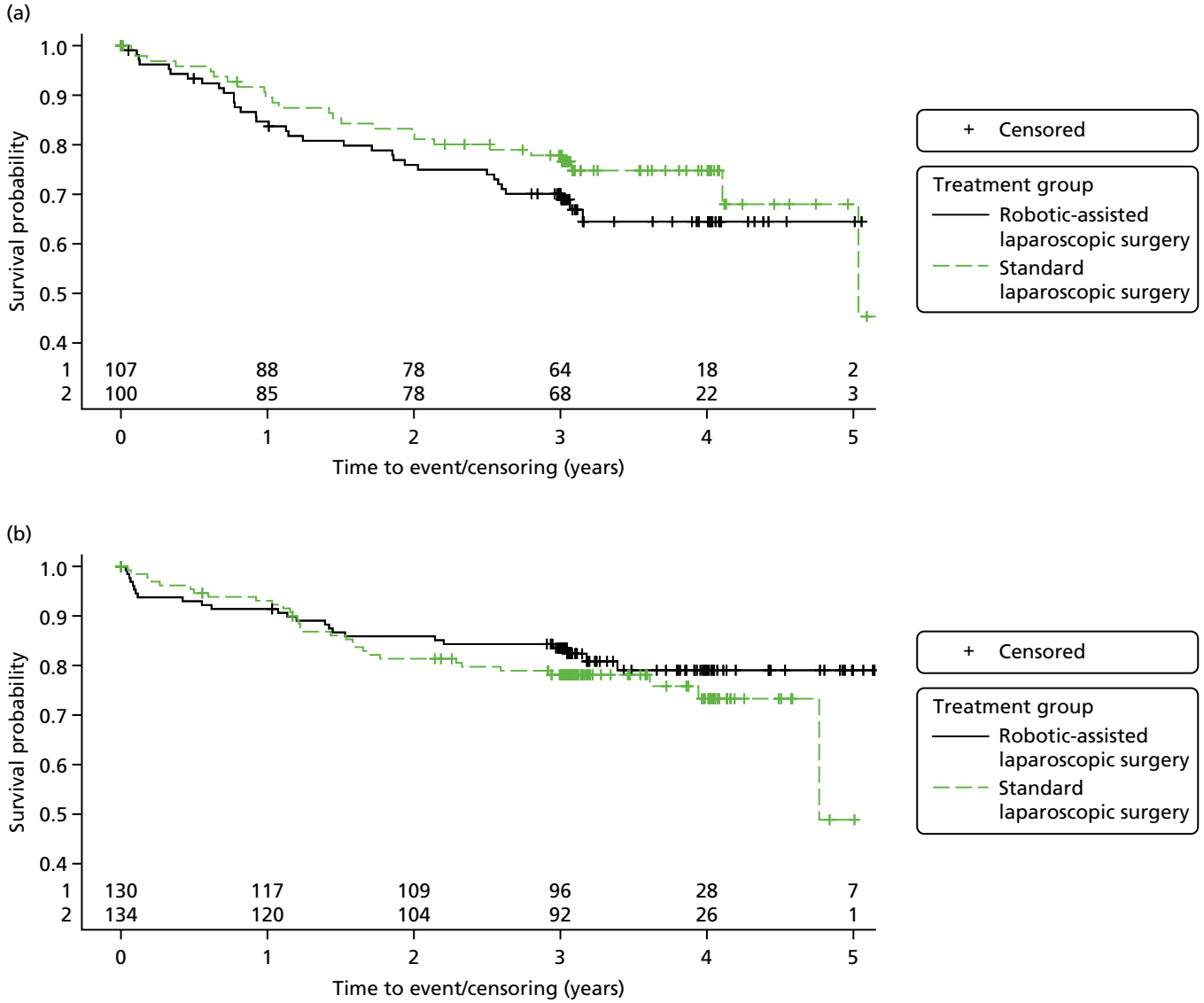


FIGURE 27 Disease-free survival by neo-adjuvant therapy. Product-limit survival estimates with number of patients at risk. (a) No neo-adjuvant therapy; and (b) neo-adjuvant therapy.

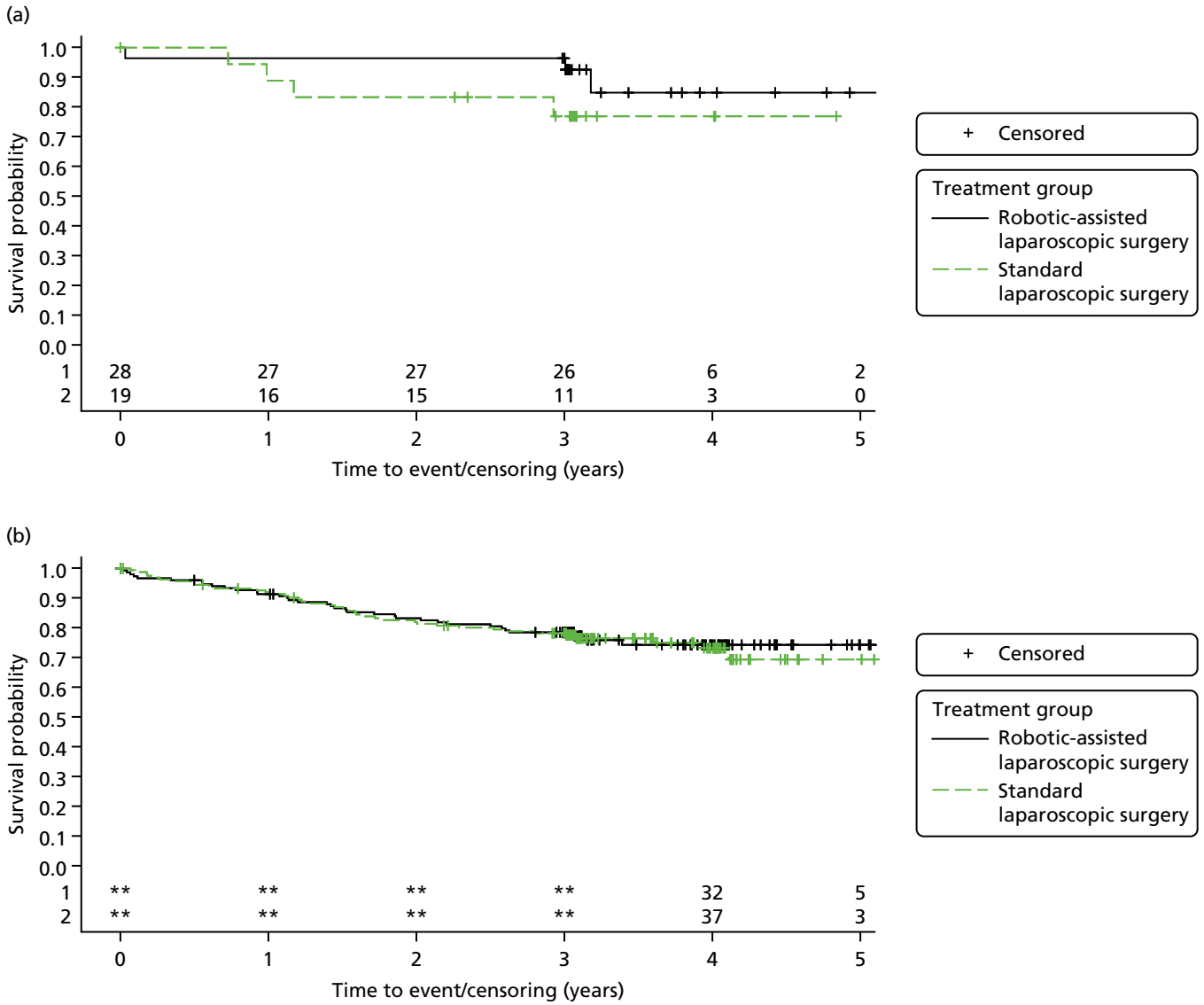


FIGURE 28 Disease-free survival by operation type. (a) HAR; (b) LAR; and (c) APR. Product-limit survival estimates with number of patients at risk. (*continued*)

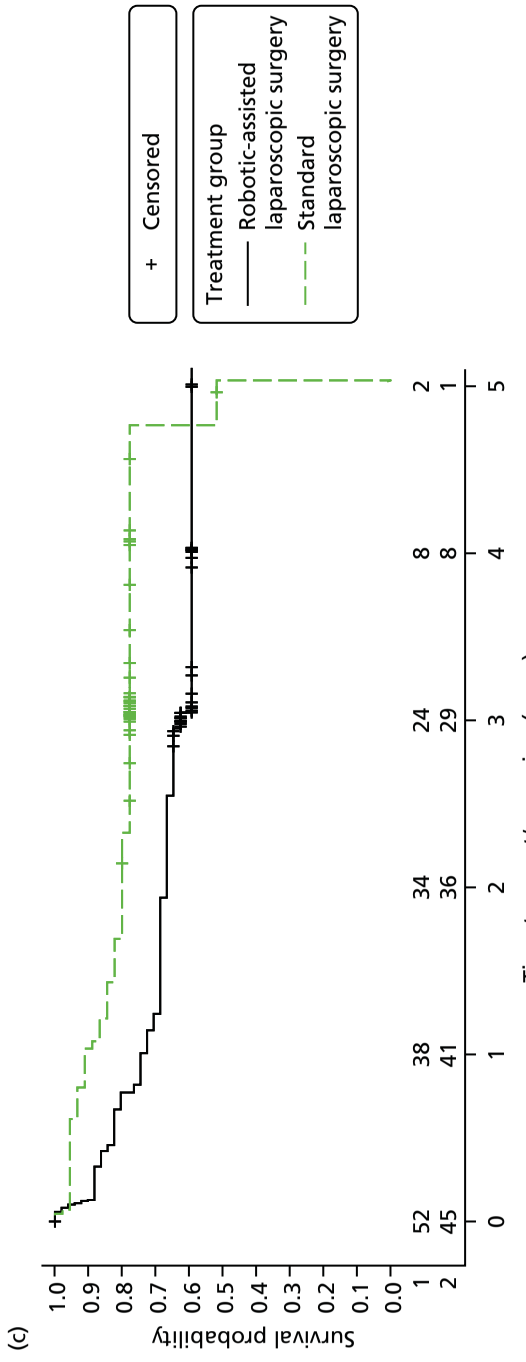


FIGURE 28 Disease-free survival by operation type. (a) HAR; (b) LAR; and (c) APR. Product-limit survival estimates with number of patients at risk.

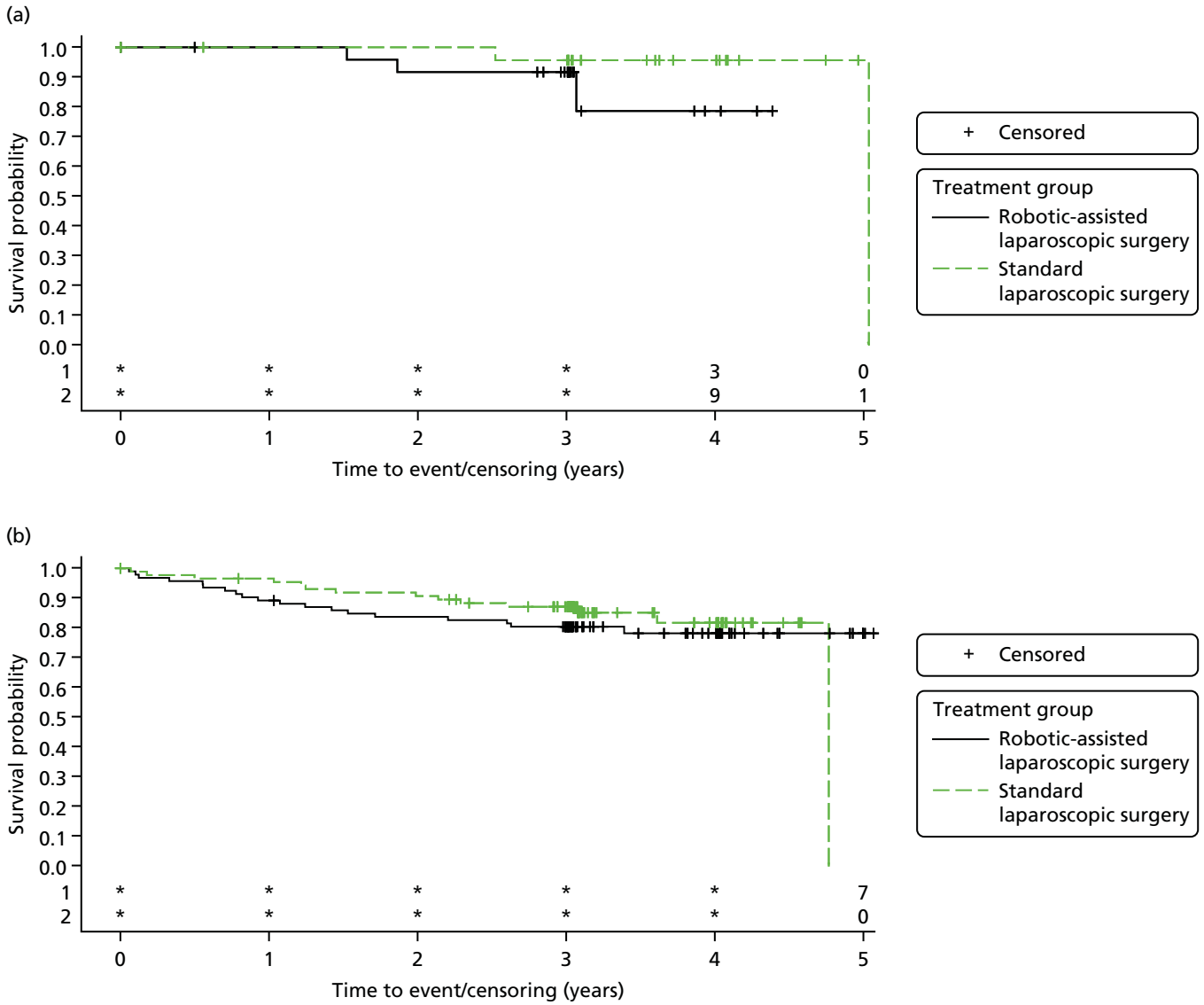


FIGURE 29 Disease-free survival by T-stage. (a) T0; (b) T1 and T2; and (c) T3 and T4. Product-limit survival estimates with number of patients at risk. (*continued*)

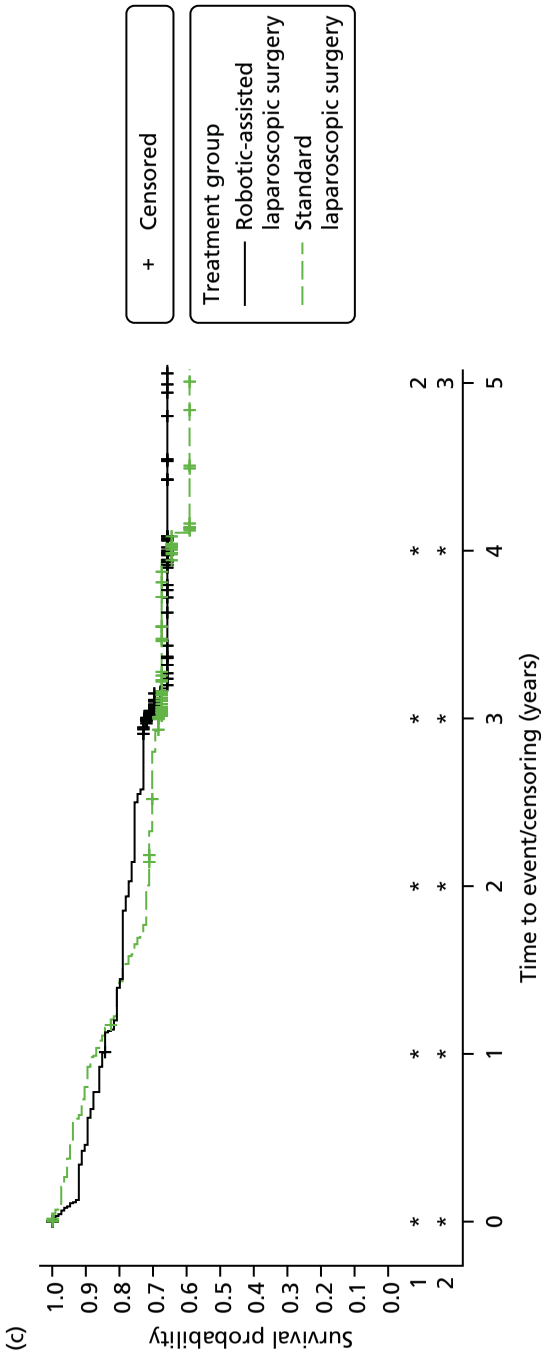


FIGURE 29 Disease-free survival by T-stage. (a) T0; (b) T1 and T2; and (c) T3 and T4. Product-limit survival estimates with number of patients at risk.

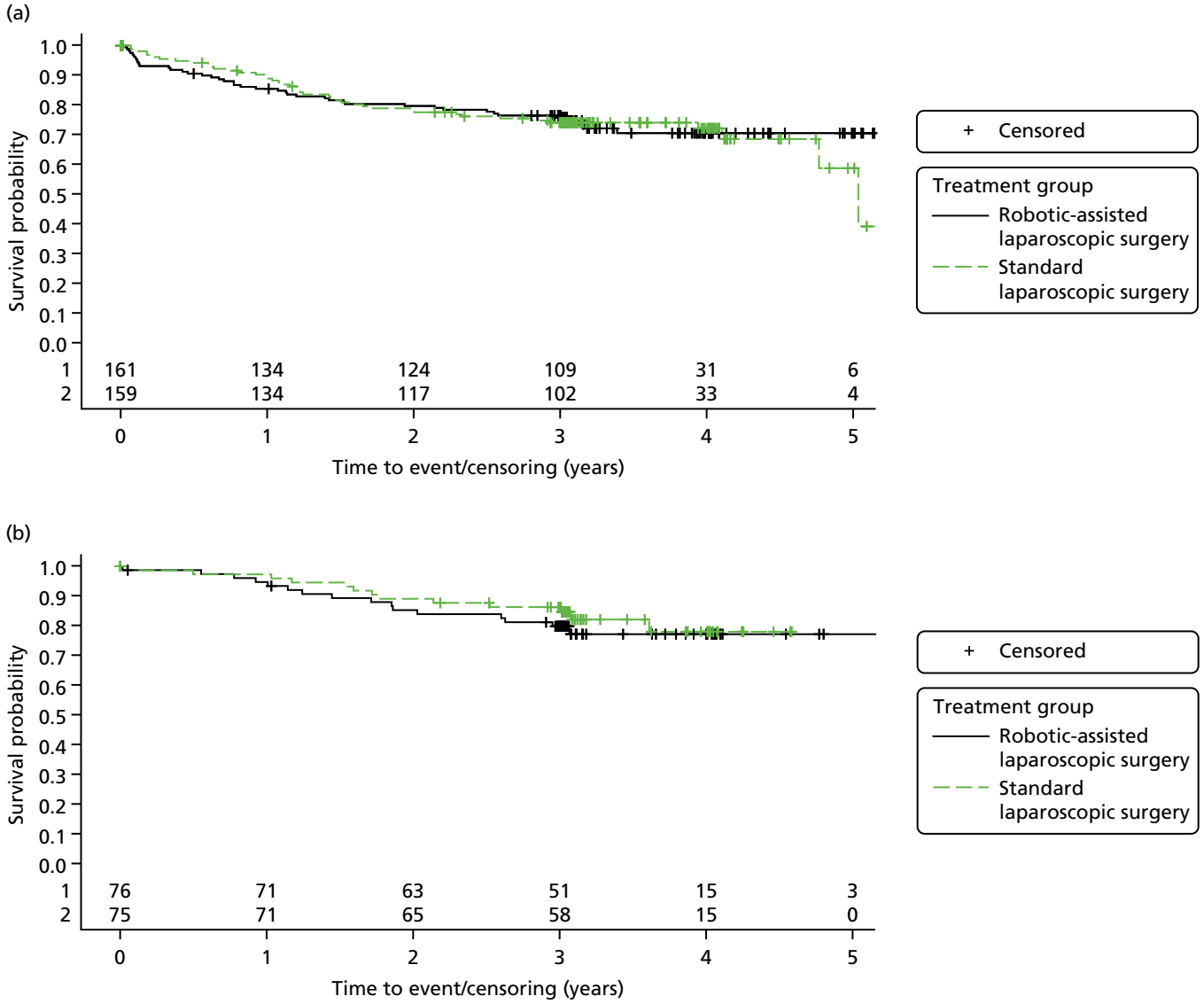


FIGURE 30 Disease-free survival by sex. (a) Males; and (b) females. Product-limit survival estimates with number of patients at risk.

Neo-adjuvant therapy

TABLE 97 Disease-free survival: subgroup analysis for neo-adjuvant therapy

Effect	HR (adjusted 95% CI)[a]	*p*-value	
Treatment in patients who underwent neo-adjuvant therapy: robotic surgery (vs. laparoscopic)	1.338 (0.795 to 2.251)	0.2728	0.1653
Treatment in patients who did not undergo neo-adjuvant therapy: robotic surgery (vs. laparoscopic)	0.787 (0.459 to 1.350)	0.3843	

a Adjusted for BMI class, preoperative radiotherapy, intended procedure and operating surgeon.
Hazard ratios derived from the treatment term and treatment-by-neo-adjuvant therapy interaction term.

Type of operation

TABLE 98 Disease-free survival: subgroup analysis for type of operation

Effect	HR (adjusted 95% CI)[a]	*p*-value	
Treatment in patients who underwent HAR: robotic surgery (vs. laparoscopic)	0.437 (0.097 to 1.957)	0.2825	0.1818
Treatment in patients who underwent LAR: robotic surgery (vs. laparoscopic)	0.914 (0.580 to 1.440)	0.6985	
Treatment in patients who underwent APR: robotic surgery (vs. laparoscopic)	1.703 (0.832 to 3.487)	0.1450	

a Adjusted for BMI class, preoperative radiotherapy, intended procedure and operating surgeon.
Hazard ratios derived from the treatment term and treatment–by-operation interaction term.

T-stage

TABLE 99 Disease-free survival: subgroup analysis by T-stage

Effect	HR (adjusted 95% CI)[a]	*p*-value	
Treatment in T0 patients: robotic surgery (vs. laparoscopic)	1.878 (0.309 to 11.391)	0.4934	0.6226
Treatment in T1 and T2 patients: robotic surgery (vs. laparoscopic)	1.252 (0.623 to 2.516)	0.5287	
Treatment in T3 and T4 patients: robotic surgery (vs. laparoscopic)	0.925 (0.585 to 1.463)	0.7402	

a Adjusted for BMI class, preoperative radiotherapy, intended procedure and operating surgeon.
Hazard ratios derived from the treatment term and treatment-by-T stage interaction term.

Sex

TABLE 100 Disease-free survival: subgroup analysis by sex

Effect	HR (adjusted 95% CI)[a]	*p*-value	
Treatment in males: robotic surgery (vs. laparoscopic)	0.971 (0.633 to 1.490)	0.8925	0.5570
Treatment in females: robotic surgery (vs. laparoscopic)	1.251 (0.599 to 2.613)	0.5520	

a Adjusted for BMI class, preoperative radiotherapy, intended procedure and operating surgeon. Hazard ratios derived from the treatment term and treatment–by-sex interaction term.

Appendix 10 Overall survival: further information

Overall survival events and censorings, including reason for censoring, are summarised in *Table 101*.

Table 102 shows Kaplan–Meier estimates of DFS by treatment group at several time points (1–5 years post randomisation).

Subgroup analyses: further information

Figures 31–34 display the Kaplan-Meier graphs for the effect of neo-adjuvant therapy, operation type, T-stage and sex, on 3-year OS.

TABLE 101 Overall survival: deaths and censoring, by treatment group

Nature of the end of follow-up	Treatment group, *n* (%)		Total (*N* = 471), *n* (%)
	Standard laparoscopic surgery (*N* = 234)	Robotic-assisted laparoscopic surgery (*N* = 237)	
Event	23 (9.8)	23 (9.7)	46 (9.8)
Censor: last known to be alive	205 (87.6)	210 (88.6)	415 (88.1)
Censor: withdrawal from further data collection	6 (2.6)	4 (1.7)	10 (2.1)

TABLE 102 Kaplan–Meier estimate of OS, by treatment group

	Treatment group			
	Laparoscopic surgery		Robotic surgery	
Time (years)	Probability of survival	95% CI	Probability of survival	95% CI
1	0.970	0.947 to 0.992	0.970	0.949 to 0.992
2	0.939	0.908 to 0.970	0.940	0.909 to 0.971
3	0.925	0.891 to 0.959	0.914	0.878 to 0.950
4	0.891	0.844 to 0.939	0.887	0.840 to 0.933
5	0.786	0.626 to 0.945	0.887	0.840 to 0.933

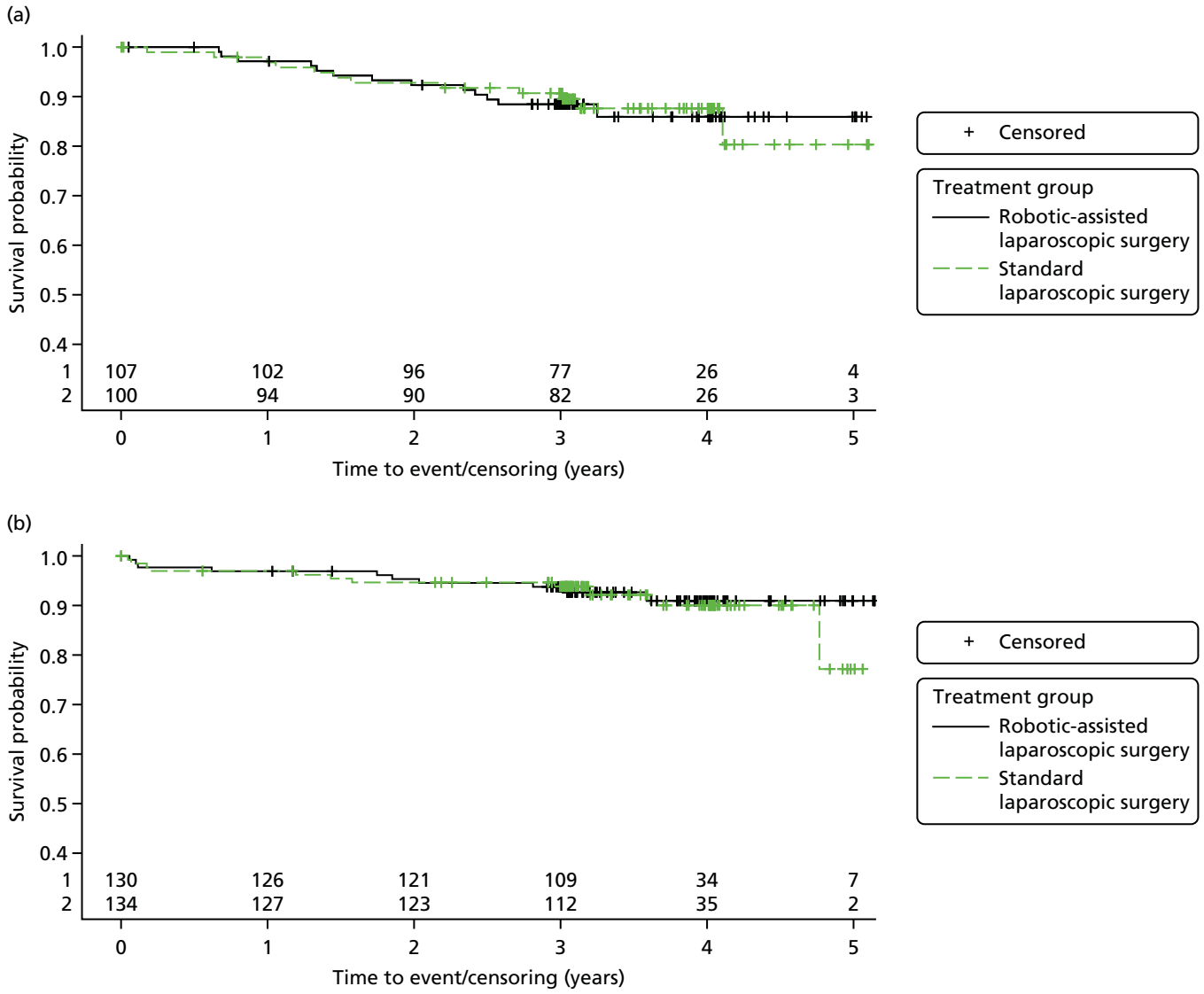


FIGURE 31 Overall survival by neo-adjuvant therapy. (a) Neo-adjuvant therapy; and (b) no neo-adjuvant therapy.

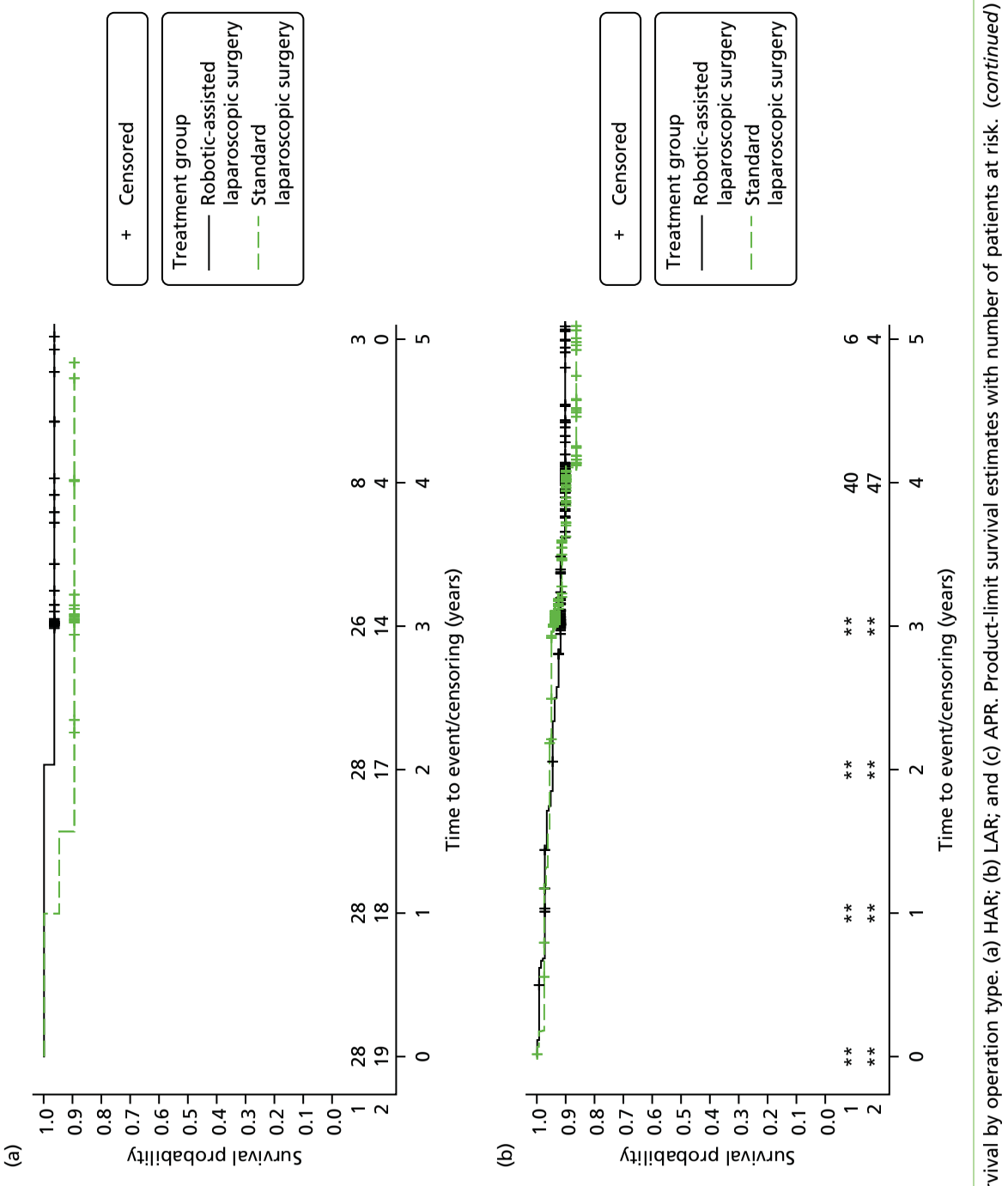


FIGURE 32 Overall survival by operation type. (a) HAR; (b) LAR; and (c) APR. Product-limit survival estimates with number of patients at risk. (*continued*)

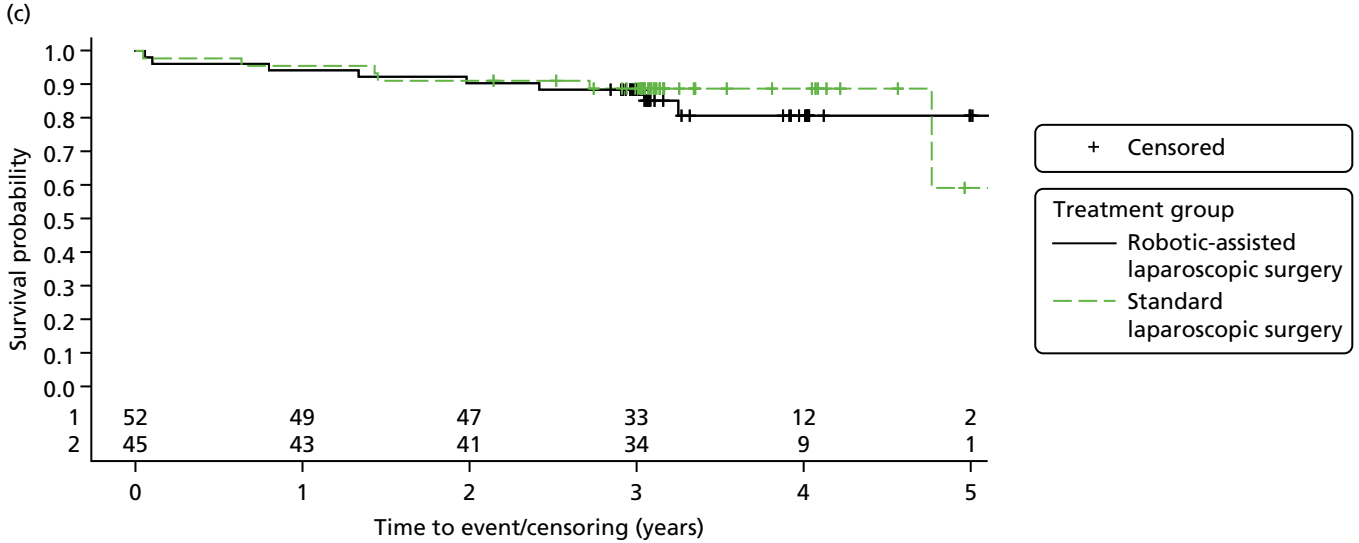


FIGURE 32 Overall survival by operation type. (a) HAR; (b) LAR; and (c) APR. Product-limit survival estimates with number of patients at risk.

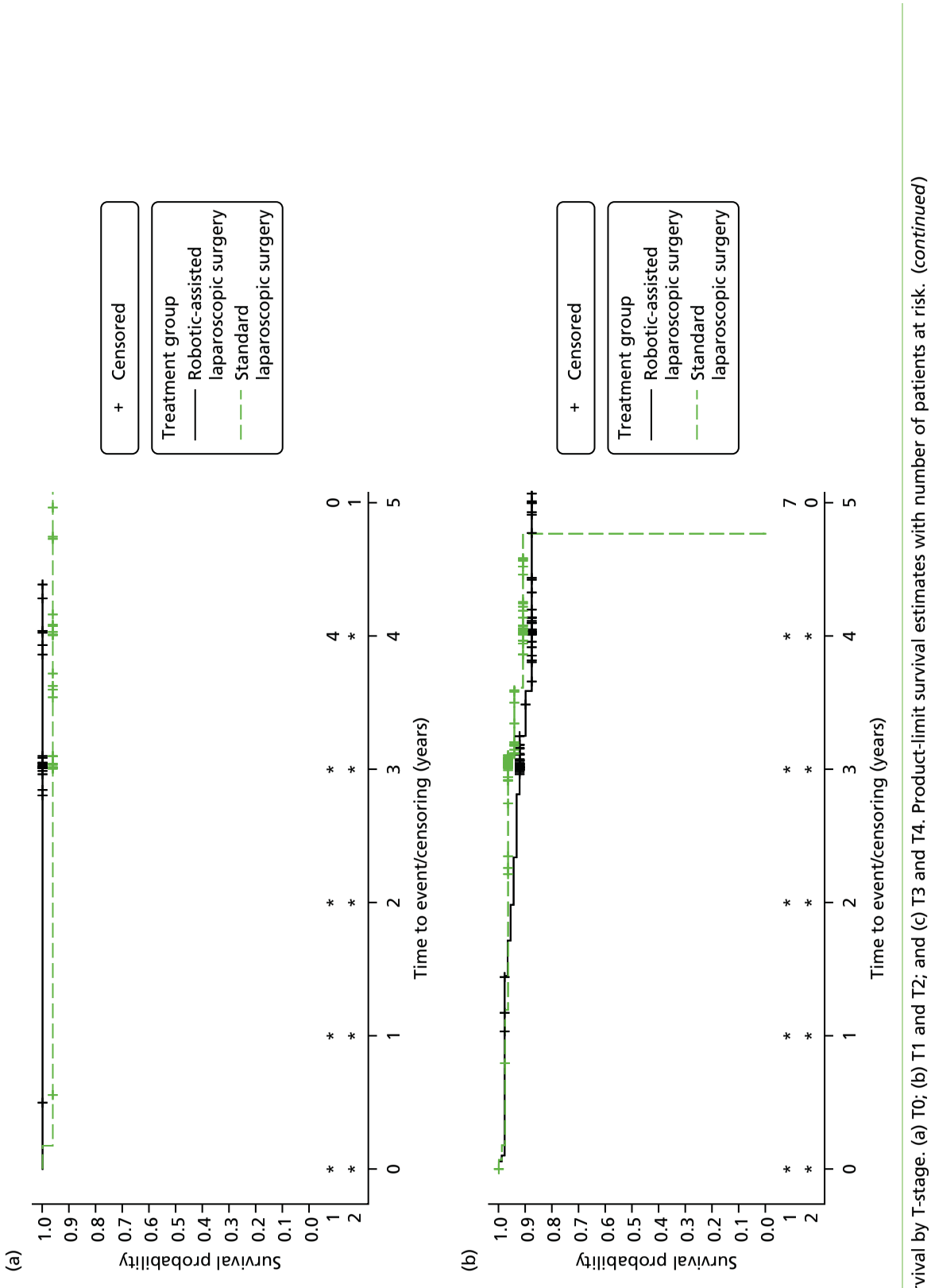


FIGURE 33 Overall survival by T-stage. (a) T0; (b) T1 and T2; and (c) T3 and T4. Product-limit survival estimates with number of patients at risk. (*continued*)

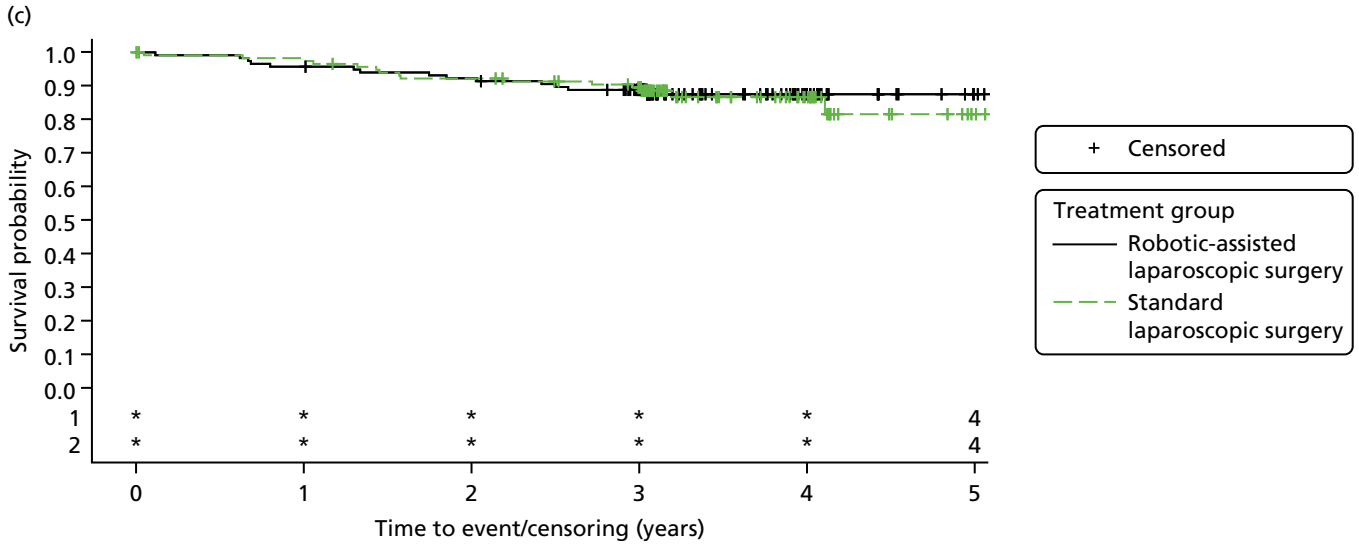


FIGURE 33 Overall survival by T-stage. (a) T0; (b) T1 and T2; and (c) T3 and T4. Product-limit survival estimates with number of patients at risk.

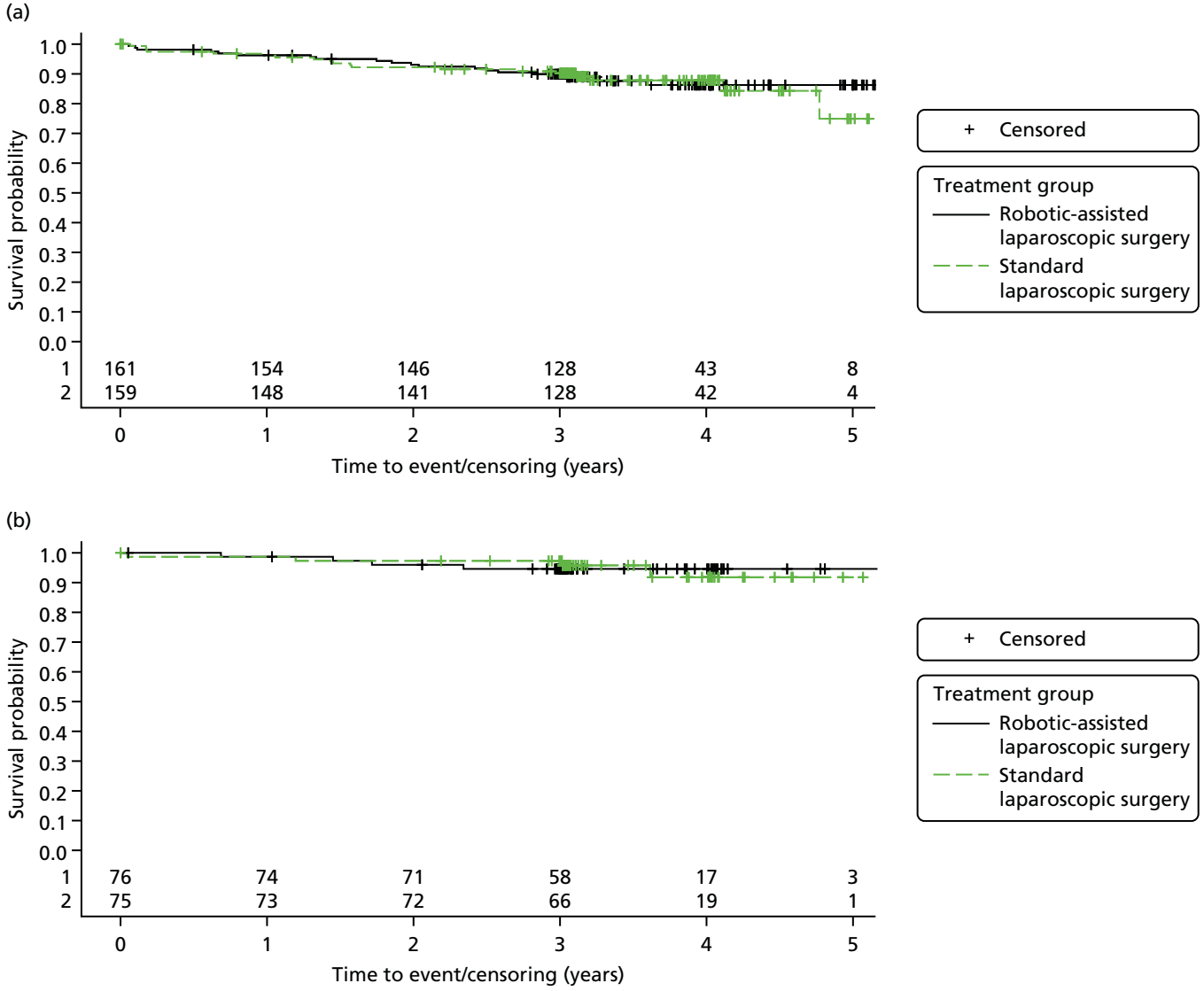


FIGURE 34 Overall survival by sex. (a) Males; and (b) females. Product-limit survival estimates with number of patients at risk.

Neo-adjuvant therapy

TABLE 103 Overall survival: subgroup analysis by neo-adjuvant therapy

Effect	HR (adjusted 95% CI)[a]	*p*-value	
Treatment in patients who underwent neo-adjuvant therapy: robotic surgery (vs. laparoscopic)	0.977 (0.448 to 2.133)	0.9544	0.9018
Treatment in patients who did not undergo neo-adjuvant therapy: robotic surgery (vs. laparoscopic)	0.908 (0.379 to 2.172)	0.8279	

a Adjusted for BMI class, preoperative radiotherapy, intended procedure and operating surgeon.
Hazard ratios derived from the treatment term and treatment-by-neo-adjuvant therapy interaction term.

Type of operation

TABLE 104 Overall survival: subgroup analysis by type of operation

Effect	HR (adjusted 95% CI)[a]	*p*-value	
Treatment in patients who underwent HAR: robotic surgery (vs. laparoscopic)	0.275 (0.025 to 3.053)	0.2933	0.5713
Treatment in patients who underwent LAR: robotic surgery (vs. laparoscopic)	0.980 (0.460 to 2.088)	0.9591	
Treatment in patients who underwent APR: robotic surgery (vs. laparoscopic)	1.122 (0.389 to 3.238)	0.8312	

a Adjusted for BMI class, preoperative radiotherapy, intended procedure and operating surgeon.
Hazard ratios derived from the treatment term and treatment–by-operation interaction term.

T-stage

TABLE 105 Overall survival: subgroup analysis by T-stage

Effect	HR (adjusted 95% CI)[a]	*p*-value	
Treatment in T0 patients: robotic surgery (vs. laparoscopic)[b]			0.8240
Treatment in T1 and T2 patients: robotic surgery (vs. laparoscopic)	1.310 (0.464 to 3.701)	0.6106	
Treatment in T3 and T4 patients: robotic surgery (vs. laparoscopic)	0.875 (0.421 to 1.819)	0.7215	

a Adjusted for BMI class, preoperative radiotherapy, intended procedure and operating surgeon.
b Only one event in T0 patients (in laparoscopic group). Within-group comparison between treatment groups not plausible.
Hazard ratios derived from the treatment term and treatment-by-T-stage interaction term.

Sex

TABLE 106 Overall survival: subgroup analysis by sex

Effect	HR (adjusted 95% CI)[a]	*p*-value	
Treatment in males: robotic surgery (vs. laparoscopic)	0.930 (0.491 to 1.763)	0.8248	0.9092
Treatment in females: robotic surgery (vs. laparoscopic)	1.018 (0.251 to 4.128)	0.9801	

a Adjusted for BMI class, preoperative radiotherapy, intended procedure and operating surgeon. Hazard ratios derived from the treatment term and treatment–by-sex interaction term.

Appendix 11 Pathology central review

The agreement of the local pathology fields with the central review was assessed for pathology fields that fed into analyses. Summaries of agreement are presented below.

Circumferential resection margin positivity (CRM+)

A total of 359 out of 471 (76.2%) patients had both local pathology data and central review data for the CRM+ field, allowing for the evaluation of agreement. Agreement was non-evaluable for the remaining 112 patients, with reasons summarised in *Table 107*.

There was agreement between local and central pathology in 343 out of 359 (95.5%) cases. The local and central evaluation of CRM+ is cross-tabulated in *Table 108*. In these 359 patients, the central review yields a CRM+ rate of 29 out of 359 (8.1%), which is greater than the rate of 21 out of 359 (5.8%) yielded by local pathology.

TABLE 107 Reasons for non-evaluable CRM+ agreement between local and central pathology review, by treatment group

Reasons	Treatment group, *n* (%)		Total (*N* = 112), *n* (%)
	Standard laparoscopic surgery (*N* = 60)	Robotic-assisted laparoscopic surgery (*N* = 52)	
Central review data for CRM+ missing or non-evaluable	50 (83.3)	50 (96.2)	100 (89.3)
Local pathology data for CRM+ missing	1 (1.7)	1 (1.9)	2 (1.8)
Local pathology data and central review data for CRM+ missing or non-evaluable	9 (15.0)	1 (1.9)	10 (8.9)

TABLE 108 Agreement between central and local review of CRM+. Cross-tabulation of central (rows) and local (columns) for CRM+

CRM+ involvement	Agreement, *n* (%)		Total (*N* = 359), *n* (%)
	Yes (*N* = 21)	No (*N* = 338)	
Yes	17 (81.0)	12 (3.6)	29 (8.1)
No	4 (19.0)	326 (96.4)	330 (91.9)

Plane of surgery

A total of 420 out of 471 (89.2%) patients had both local pathology data and central review data for the plane of surgery field, allowing for the evaluation of agreement. Agreement was non-evaluable for the remaining 51 patients, with reasons summarised in *Table 109*.

There was agreement between local and central pathology in 262 out of 420 (62.4%) cases. The local and central evaluation of plane of surgery is cross-tabulated in *Table 110*. In these 420 patients, the central review yields a rate of mesorectal fascial plane of 179 out of 420 (42.6%), which is notably less than the rate of 319 out of 420 (76.0%) yielded by local pathology. The large majority of the disagreement between local and central pathology is seen when local pathology considered the plane of surgery to be mesorectal fascial, which is then downgraded to intramesorectal or muscularis propria plane by the central review.

TABLE 109 Reasons for non-evaluable agreement for plane of surgery, by treatment group

Reasons	Treatment group, *n* (%)		Total (*N* = 51), *n* (%)
	Standard laparoscopic surgery (*N* = 26)	Robotic-assisted laparoscopic surgery (*N* = 25)	
Central review data for plane of surgery missing or non-evaluable	15 (57.7)	21 (84.0)	36 (70.6)
Local pathology data for plane of surgery missing	3 (11.5)	2 (8.0)	5 (9.8)
Local pathology data and central review data for plane of surgery missing or non-evaluable	8 (30.8)	2 (8.0)	10 (19.6)

TABLE 110 Agreement between local and central review of plane of surgery. Cross-tabulation of central (rows) and local (columns) for mesorectum plane

Mesorectum plane	Mesorectum plane, *n* (%)			
	Mesorectal fascial plane (*N* = 319)	Intramesorectal plane (*N* = 67)	Muscularis propria plane (*N* = 34)	Total (*N* = 420), *n* (%)
Mesorectal fascial plane	174 (54.5)	4 (6.0)	1 (2.9)	179 (42.6)
Intramesorectal plane	124 (38.9)	58 (86.6)	3 (8.8)	185 (44.0)
Muscularis propria plane	21 (6.6)	5 (7.5)	30 (88.2)	56 (13.3)

Pathological T-stage

A total of 456 out of 471 (96.8%) patients had both local pathology data and central review data for the pT-stage field, allowing for the evaluation of agreement. Agreement was non-evaluable for the remaining 15 patients, with reasons summarised in *Table 111*.

There was agreement between local and central pathology in 424 out of 456 (93.0%) cases. The local and central evaluation of pT-stage is cross-tabulated in *Table 112*.

TABLE 111 Reasons for non-evaluable agreement for T-stage, by treatment group

Reasons for non-evaluable agreement for pT-stage (by group)	Treatment group, *n* (%)		Total (*N* = 15), *n* (%)
	Standard laparoscopic surgery (*N* = 10)	Robotic-assisted laparoscopic surgery (*N* = 5)	
Central review data for T-stage missing or non-evaluable	2 (20.0)	1 (20.0)	3 (20.0)
Local pathology data for T-stage missing	1 (10.0)	2 (40.0)	3 (20.0)
Local pathology data and central review data for T-stage missing or non-evaluable	7 (70.0)	2 (40.0)	9 (60.0)

TABLE 112 Agreement between local and central pathology review for T-stage. Cross-tabulation of central (rows) and local (columns) for pT stage

pT-stage	pT stage, *n* (%)					Total (*N* = 456), *n* (%)
	0 (*N* = 52)	1 (*N* = 46)	2 (*N* = 129)	3 (*N* = 217)	4 (*N* = 12)	
0	46 (88.5)	0 (0.0)	0 (0.0)	1 (0.5)	0 (0.0)	47 (10.3)
1	4 (7.7)	46 (100.0)	2 (1.6)	1 (0.5)	0 (0.0)	53 (11.6)
2	2 (3.8)	0 (0.0)	116 (89.9)	6 (2.8)	0 (0.0)	124 (27.2)
3	0 (0.0)	0 (0.0)	11 (8.5)	205 (94.5)	1 (8.3)	217 (47.6)
4	0 (0.0)	0 (0.0)	0 (0.0)	4 (1.8)	11 (91.7)	15 (3.3)

Appendix 12 Summary of protocol changes

Version and date	Summary of changes
V1.0 dated 17 February 2010	N/A – original protocol submitted for ethical review
V2.0 dated 25 March 2010	Changes were required to the PIS/ICF document following ethical review, protocol was upversioned to match revised version of PIS/ICF (no changes made to the protocol)
V3.0 dated 2 August 2010	• Contents and references updated • Clarification of eligibility criteria • Surgeon eligibility: it was initially stipulated that surgeons must have performed at least 15 rectal cancer resections per annum and have prior experience of at least 10 robotic-assisted rectal cancer resections. It was felt that to ensure that the 'learning curve' effect did not bias the trial data, only surgeons who had performed at least 30 rectal cancer resections, with a minimum of 10 of these to be standard laparoscopic procedures, and 10 of these to be robotic-assisted procedures, should be included in the trial to ensure surgeon competency in both arms of the trial • BMI added as a stratification factor due to recent publications that increased BMI may be associated with an increased risk of conversion to open surgery • Schedule of Events modified for clarity • Pathology section updated, including collection of extra tumour samples for tissue banking (explicit consent was obtained) • Update of Study Organisational Structure diagram • Clarification of pathology appendix • Minor administrative changes
V4.0 dated 1 March 2011	• Funder (EME) requested changes to reference to EME and removal of logos on front cover • Contacts and table of contents updated • Section 2 and 19: addition of ROLARR protocol publication • Section 4 Eligibility: clarifications to procedures following feedback from the international trial launch meetings • Section 5.2 Randomisation: clarification of timings following feedback from the international launch meetings • Table 1: expedited safety reporting timeline revised to 30 days (correction to previous version) • Clarification added that all procedures will be video'd and the CTRU will inform sites which procedures to submit • Pregnancy statement added • Section 7.5 and Appendix 1: clarifications were made to pathology processes following consultation with a trial pathologist following the analysis of the first trial specimen. Also procedures for submitting the slides for trial purposes and extra tumour/normal tissue blocks as an optional separate study were updated to ensure HTA compliance and clarity in procedures • Section 7.7: clarification of annual follow-up timing and addition of stoma details • Section 7.9 Pregnancy: section added to ensure patients who may become pregnant on trial are handled correctly • Section 7.12: end of study definition included as it was omitted in previous version of protocol • Section 8: safety updates following feedback from the international launch meetings. Clarification of timing for expedited safety reports • Sections 8.5 and 15.1: sections regarding procedures/responsibilities for Serious Breaches of GCP added in line with the latest CTRU policies • Section 16 indemnity: updated following discussion with the insurers • Section 17.2 Responsibilities: LIMM responsibilities added regarding the central pathology assessment and optional separate tissue block study • Section 18: clarification to publication policy as journal author restrictions may be in place • Section 23 Abbreviations: updated
V5.0 dated 19 March 2014	• Contacts updated • Removal of secondary endpoint: Global Operative Assessment of Laparoscopic Skills tool (GOALS). It was planned that an assessment of surgical skills would be carried out using the GOALS assessment. Videos were to be taken of the complete mesorectal dissection from all cases inclusive of both laparoscopic and robotic operations however this proved to be unfeasible due to the large size of the files

Version and date	Summary of changes
	• Clarification of inclusion criteria surrounding diagnosis of rectal cancer amenable to curative surgery. The inclusion criteria stated that a T-staging of 1–3 was a component of a patient being 'amenable to curative surgery' for the purpose of this trial. This was discussed by the Trial Management Group who agreed that the decision of the team to perform surgery acts as a sufficient indication that the patient is amenable to curative surgery. Therefore T-staging of 1–3 is a guide only; to reflect this, 'i.e.' was removed and replaced with 'for example' • Specified that histopathology reports are only to be collected if reported in English • Clarification to assessment of Unexpected Serious Complications: the chief investigator can upgrade or downgrade assessment in the event of disagreement between local assessment in line with CTRU standard guidance • Increase in sample size. Following a successful extension request, the trial recruited to target ahead of the revised milestones, and the opportunity to recruit further patients within the revised timelines and budget was taken to maximise study power. The sample size was amended from 400 to a maximum of 520 participants, to increase the power of the study from 80% to a maximum of 90% power • Removal of South-East Asian Spoke to reflect actual spoke arrangements. The South-East Asian Spoke were unable to secure additional funding for them to deliver the CTU spoke function in Singapore, so were unable to act as a spoke. CTRU, Leeds (i.e. the Hub) therefore co-ordinated centres in South-East Asia and sites across the rest of the world (with the exception of sites in the USA which were coordinated by the North American Spoke) • Neo-adjuvant therapy: clarification that eligibility should be reassessed on completion of neo-adjuvant therapy and guidance added on timing of consent • Analyses on a surgeon basis: additional wording added to expand on planned analyses (wording omitted in error from previous protocols) • Expected complications list expanded and grouped into relevant categories • Minor administrative changes
V6.0 dated 1 July 2015	This amendment to the protocol included an additional one-off questionnaire as a supplementary study to the ROLARR trial to determine the incidence and severity of Low Anterior Resection Syndrome (LARS) within participants of the ROLARR trial. It was felt that the results would have important consequences when counselling future patients with rectal cancer on the likely functional outcomes of surgery. The protocol was amended to include an additional appendix (*Appendix 4*) to cover the Low Anterior Resection Syndrome (LARS) supplementary study. Eligible ROLARR participants from Denmark, Germany, Italy, the USA and the UK were invited to complete a one-off postal survey

CTU, Clinical Trials Unit; ICF, informed consent form; GCP, Good Clinical Practice; LIMM, Leeds Institute of Molecular Medicine; N/A, not applicable; PIS: patient information sheet.

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library