Clinical effectiveness and cost-effectiveness of emergency surgery for adult emergency hospital admissions with common acute gastrointestinal conditions: the ESORT study

Richard Grieve, Andrew Hutchings, Silvia Moler Zapata, Stephen O'Neill, David G Lugo-Palacios, Richard Silverwood, David Cromwell, Tommaso Kircheis, Elizabeth Silver, Claire Snowdon, Paul Charlton, Geoff Bellingan, Ramani Moonesinghe, Luke Keele, Neil Smart and Robert Hinchliffe
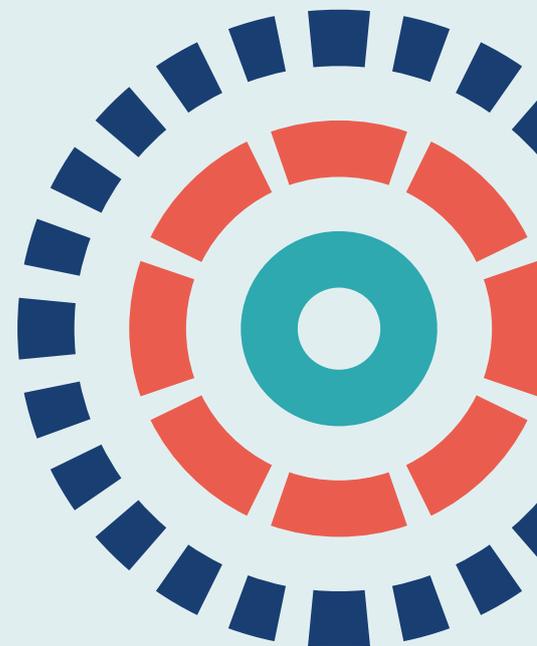
Clinical effectiveness and cost-effectiveness of emergency surgery for adult emergency hospital admissions with common acute gastrointestinal conditions: the ESORT study

Richard Grieve,[1]* Andrew Hutchings,[1]
Silvia Moler Zapata,[1] Stephen O'Neill,[1]
David G Lugo-Palacios,[1] Richard Silverwood,[2]
David Cromwell,[1] Tommaso Kircheis,[1]
Elizabeth Silver,[3] Claire Snowdon,[4] Paul Charlton,[5]
Geoff Bellingan,[6,7] Ramani Moonesinghe,[8]
Luke Keele,[9] Neil Smart[10] and Robert Hinchliffe[11]

[1]Department of Health Services Research and Policy, London School of Hygiene &
Tropical Medicine, London, UK
[2]Centre for Longitudinal Studies, University College London, London, UK
[3]Women and Children First, London, UK
[4]Department for Medical Statistics, London School of Hygiene & Tropical Medicine,
London, UK
[5]Patient ambassador, National Institute for Health and Care Research, Southampton, UK
[6]Intensive Care Medicine, University College London, London, UK
[7]NIHR Biomedical Research Centre at University College London Hospitals NHS
Foundation Trust and University College London, London, UK
[8]Centre for Perioperative Medicine, University College London Hospitals, London, UK
[9]Department of Surgery, University of Pennsylvania, Philadelphia, PA, USA
[10]College of Medicine and Health, University of Exeter, Exeter, UK
[11]NIHR Bristol Biomedical Research Centre, University of Bristol, Bristol, UK

*Corresponding author

Declared competing interests of authors

Health and Social Care Delivery Research

Criteria for inclusion in the *Health and Social Care Delivery Research* journal
Reports are published in *Health and Social Care Delivery Research* (HSDR) if (1) they have resulted from work for the HSDR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

HSDR programme

The HSDR programme funds research to produce evidence to impact on the quality, accessibility and organisation of health and social care services. This includes evaluations of how the NHS and social care might improve delivery of services.

For more information about the HSDR programme please visit the website at https://www.nihr.ac.uk/explore-nihr/funding-programmes/health-and-social-care-delivery-research.htm

This report

Abstract

Clinical effectiveness and cost-effectiveness of emergency surgery for adult emergency hospital admissions with common acute gastrointestinal conditions: the ESORT study

Richard Grieve[ID],[1*] Andrew Hutchings[ID],[1] Silvia Moler Zapata[ID],[1]
Stephen O'Neill[ID],[1] David G Lugo-Palacios[ID],[1] Richard Silverwood[ID],[2]
David Cromwell[ID],[1] Tommaso Kircheis[ID],[1] Elizabeth Silver[ID],[3]
Claire Snowdon[ID],[4] Paul Charlton[ID],[5] Geoff Bellingan[ID],[6,7]
Ramani Moonesinghe[ID],[8] Luke Keele[ID],[9] Neil Smart[ID][10]
and Robert Hinchliffe[ID][11]

[1]Department of Health Services Research and Policy, London School of Hygiene & Tropical Medicine, London, UK
[2]Centre for Longitudinal Studies, University College London, London, UK
[3]Women and Children First, London, UK
[4]Department for Medical Statistics, London School of Hygiene & Tropical Medicine, London, UK
[5]Patient ambassador, National Institute for Health and Care Research, Southampton, UK
[6]Intensive Care Medicine, University College London, London, UK
[7]NIHR Biomedical Research Centre at University College London Hospitals NHS Foundation Trust and University College London, London, UK
[8]Centre for Perioperative Medicine, University College London Hospitals, London, UK
[9]Department of Surgery, University of Pennsylvania, Philadelphia, PA, USA
[10]College of Medicine and Health, University of Exeter, Exeter, UK
[11]NIHR Bristol Biomedical Research Centre, University of Bristol, Bristol, UK

*Corresponding author richard.grieve@lshtm.ac.uk

Background: Evidence is required on the clinical effectiveness and cost-effectiveness of emergency surgery compared with non-emergency surgery strategies (including medical management, non-surgical procedures and elective surgery) for patients admitted to hospital with common acute gastrointestinal conditions.

Objectives: We aimed to evaluate the relative (1) clinical effectiveness of two strategies (i.e. emergency surgery vs. non-emergency surgery strategies) for five common acute conditions presenting as emergency admissions; (2) cost-effectiveness for five common acute conditions presenting as emergency admissions; and (3) clinical effectiveness and cost-effectiveness of the alternative strategies for specific patient subgroups.

Methods: The records of adults admitted as emergencies with acute appendicitis, cholelithiasis, diverticular disease, abdominal wall hernia or intestinal obstruction to 175 acute hospitals in England between 1 April 2010 and 31 December 2019 were extracted from Hospital Episode Statistics and linked to mortality data from the Office for National Statistics. Eligibility was determined using *International Statistical Classification of Diseases and Related Health Problems*, Tenth Revision, diagnosis codes, which were agreed by clinical panel consensus. Patients having emergency surgery were identified from Office of Population Censuses and Surveys procedure codes. The study addressed the potential for unmeasured confounding with an instrumental variable design. The instrumental variable was each hospital's propensity to use emergency surgery compared with non-emergency

surgery strategies. The primary outcome was the 'number of days alive and out of hospital' at 90 days. We reported the relative effectiveness of the alternative strategies overall, and for prespecified subgroups (i.e. age, number of comorbidities and frailty level). The cost-effectiveness analyses used resource use and mortality from the linked data to derive estimates of incremental costs, quality-adjusted life-years and incremental net monetary benefits at 1 year.

Results: Cohort sizes were as follows: 268,144 admissions with appendicitis, 240,977 admissions with cholelithiasis, 138,869 admissions with diverticular disease, 106,432 admissions with a hernia and 133,073 admissions with an intestinal obstruction. Overall, at 1 year, the average number of days alive and out of hospitals at 90 days, costs and quality-adjusted life-years were similar following either strategy, after adjusting for confounding. For each of the five conditions, overall, the 95% confidence intervals (CIs) around the incremental net monetary benefit estimates all included zero. For patients with severe frailty, emergency surgery led to a reduced number of days alive and out of hospital and was not cost-effective compared with non-emergency surgery, with incremental net monetary benefit estimates of –£18,727 (95% CI –£23,900 to –£13,600) for appendicitis, –£7700 (95% CI –£13,000 to –£2370) for cholelithiasis, –£9230 (95% CI –£24,300 to £5860) for diverticular disease, –£16,600 (95% CI –£21,100 to –£12,000) for hernias and –£19,300 (95% CI –£25,600 to –£13,000) for intestinal obstructions. For patients who were 'fit', emergency surgery was relatively cost-effective, with estimated incremental net monetary benefit estimates of £5180 (95% CI £684 to £9680) for diverticular disease, £2040 (95% CI £996 to £3090) for hernias, £7850 (95% CI £5020 to £10,700) for intestinal obstructions, £369 (95% CI –£728 to £1460) for appendicitis and £718 (95% CI £294 to £1140) for cholelithiasis. Public and patient involvement translation workshop participants emphasised that these findings should be made widely available to inform future decisions about surgery.

Limitations: The instrumental variable approach did not eliminate the risk of confounding, and the acute hospital perspective excluded costs to other providers.

Conclusions: Neither strategy was more cost-effective overall. For patients with severe frailty, non-emergency surgery strategies were relatively cost-effective. For patients who were fit, emergency surgery was more cost-effective.

Future work: For patients with multiple long-term conditions, further research is required to assess the benefits and costs of emergency surgery.

Study registration: This study is registered as reviewregistry784.

Contents

CONTENTS

x

List of tables

List of figures

xvi

List of boxes

List of supplementary material

Report Supplementary Material 1 Case GRIPP2 form

Report Supplementary Material 2 ESORT PPI panel: minutes of workshops held on 6 and 9 July 2020

Report Supplementary Material 3 ESORT clinical panel: summary note – May 2021

Report Supplementary Material 4 ESORT statistical analysis plan: July 2021

Report Supplementary Material 5 ESORT PPI panel: minutes of workshops held on 15 and 17 September 2021

Supplementary material can be found on the NIHR Journals Library report page (https://doi.org/10.3310/CZFL0619).

Supplementary material has been provided by the authors to support the report and any files provided at submission will have been seen by peer reviewers, but not extensively reviewed. Any supplementary material provided at a later stage in the process may not have been peer reviewed.

List of abbreviations

CEA	cost-effectiveness analysis	NELA	National Emergency Laparotomy Audit
CI	confidence interval	NES	non-emergency surgery
DAOH	days alive and out of hospital	NICE	National Institute for Health and Care Excellence
ES	emergency surgery	NIHR	National Institute for Health and Care Research
ESORT	Emergency Surgery OR noT		
GLM	generalised linear model	ONS	Office for National Statistics
HES	Hospital Episode Statistics	OPCS-4	Office of Population Censuses and Surveys version 4
HRQoL	health-related quality of life	PPI	patient and public involvement
ICD-10	*International Statistical Classification of Diseases and Related Health Problems, Tenth Revision*	QALY	quality-adjusted life-year
		RCS	Royal College of Surgeons
INB	incremental net monetary benefit	RCT	randomised controlled trial
		RMSE	root mean squared error
IQR	interquartile range	SA	sensitivity analysis
IV	instrumental variable	SCARF	Secondary Care Administrative Records Frailty
LIV	local instrumental variable		
LOS	length of stay	TTO	tendency to operate

Plain English summary

Background

Many patients are admitted to NHS hospitals in an emergency with common acute conditions. In England, just over half of these patients have surgery within days of admission, and this is emergency surgery. Other patients have another approach to care, for example drug treatment or surgery at a later date.

Patients are treated differently across England. We do not know whether or not the benefits and costs of emergency surgery are greater than alternative approaches to care.

Methods

We aimed to find out which patients should have emergency surgery, and to understand the benefits, risks and costs involved. We used routinely collected information from hospitals in England for the years 2009–19 to compare:

- outcomes for patients who had emergency surgery or another approach to care
- the costs of emergency surgery and other approaches to care
- how benefits and costs change for different types of patients.

Results

Overall, we found that people tended to have similar outcomes and costs, regardless of whether they had emergency surgery or another approach to care.

However, for some subgroups, we did find the following important differences:

- For patients with intestinal obstruction who were not frail, we found that emergency surgery generally improved outcomes and reduced costs.
- For people who were very frail, other approaches to care led to greater benefits and reduced costs, and emergency surgery could lead to worse outcomes.

Further research to find out whether emergency surgery or later surgery improves outcome for patients with several long-term conditions would be useful.

Patient and public involvement

A patient and public involvement panel initially helped us focus on what outcomes matter for patients, and later helped us plan how to communicate our results to the public.

We worked with people with autism or learning difficulties to produce an easy-read guide that explained the study and how research can use health records.

Scientific summary

Background

Patients with common acute gastrointestinal conditions who require emergency hospital admissions may receive emergency surgery (ES) or non-emergency surgery (NES) strategies, which include medical management, non-surgical procedures (e.g. radiological-guided drainage of abscess) or surgery deferred to the elective setting. For some common acute conditions, such as diverticular disease, there are well-developed NES strategies, and little evidence that ES leads to better outcomes. For other conditions, such as acute cholelithiasis, randomised controlled trials (RCTs) have found that NES strategies may have unintended consequences, with patients having recurrent symptoms and delayed surgery, therefore, leading to further pressure on surgical waiting lists. However, to the best of our knowledge, none of these RCTs have included routine emergency admissions to hospitals. Hence, for many patients with common conditions, the relative benefits and costs of ES compared with NES strategies are unknown, and there is wide variation across NHS hospital trusts in ES rates and outcomes.

To provide evidence to inform future service provision, this observational study evaluates the relative effectiveness and costs of ES compared with NES strategies for common acute conditions.

Aims and objectives

The study's aim was to evaluate the clinical effectiveness and cost-effectiveness of ES compared with NES strategies for adults admitted as emergencies for five common gastrointestinal conditions. Our objectives were to evaluate the:

1. clinical effectiveness of ES compared with NES strategies for the five selected acute gastrointestinal conditions
2. relative cost-effectiveness of ES compared with NES strategies for the five conditions
3. clinical effectiveness and cost-effectiveness of ES compared with NES strategies for specific patient subgroups.

Methods

Overview
The Emergency Surgery OR noT (ESORT) study assessed the clinical effectiveness and cost-effectiveness of ES compared with NES strategies for adults with common acute gastrointestinal conditions who presented as emergency admissions to NHS hospitals in England. The target population included adults admitted as emergencies with acute appendicitis, cholelithiasis, diverticular disease, an abdominal wall hernia or an intestinal obstruction. We included these five acute conditions because there is clinical uncertainty about the relative benefits of ES compared with NES strategies, and wide variation in ES rates across NHS hospitals. We defined the target population and comparator strategies from emergency admissions within a single data source [i.e. Hospital Episode Statistics (HES)] to ensure consistent definitions across the patient cohort. Eligibility was determined using *International Statistical Classification of Diseases and Related Health Problems*, Tenth Revision (ICD-10), diagnosis codes that were agreed by clinical panel consensus. We applied 'target trial' principles to define explicit eligibility criteria and the comparator strategies from HES data.

Ethics

The research was approved by the London School of Hygiene and Tropical Medicine Ethics Committee (reference 21687). The study involved the secondary analyses of existing pseudoanonymised data and did not require UK National Ethics Committee approval.

Data and study design

We collated information on diagnosis (ICD-10), case mix, surgical procedures received, resource use and outcomes, including all-cause mortality from HES data linked to Office for National Statistics (ONS) mortality data. The study addressed confounding by indication with an instrumental variable (IV) design, as this can provide accurate estimates of treatment effectiveness and cost-effectiveness even when there are unmeasured differences between the groups being compared. 'Tendency to operate' (TTO) was used as the IV, which was previously validated for evaluating the clinical effectiveness of ES in the USA. The IV in the ESORT study was defined as each hospital's propensity to use ES and NES strategies for each of the acute conditions. The rationale for the IV was that patients of similar prognosis were more likely to receive ES in some hospitals, and NES strategies in others. The ESORT study assessed whether or not the TTO met the two requisite conditions for an IV, that is (1) that the hospital's TTO was associated with having ES and (2) that the TTO was not related to outcomes, except through the receipt of ES.

Definitions of cohorts, and emergency surgery and non-emergency surgery strategies

The target populations included hospital admissions for patients aged $\geq$ 18 years that (1) occurred between 1 April 2010 and 31 December 2019; (2) included a main diagnosis with a relevant ICD-10 diagnosis code, or for the intestinal obstruction cohort a relevant diagnosis in the second diagnosis field if the main diagnosis was colorectal cancer; (3) was within an emergency admission through the emergency department or from a primary care referral; (4) was under a consultant general surgeon, subspecialty general surgeon or any other surgeon working in the general surgery specialty; and (5) was the first or second episode within the admission. Admissions for which there was a prior emergency admission with a relevant diagnosis in the previous 12 months were excluded.

Emergency surgery was defined from a list of Office of Population Censuses and Surveys codes and had to be within a maximum time window, which was defined by the consensus of the clinical panel. The maximum time windows were within 3 days for a hernia, within 7 days for appendicitis, cholelithiasis or a intestinal obstruction, or any time within the emergency admission for diverticular disease. The NES strategy included antibiotic therapy, non-operative procedures, including interventional radiology, and operative procedures that did not meet the ES criteria.

Data on patient characteristics were extracted from HES and used to define prespecified subgroups, including age (years), number of comorbidities (using the Charlson Comorbidity Index) and frailty [according to the Secondary Care Administrative Records Frailty (SCARF) index]. The SCARF index is based on the concept of frailty as an accumulation of deficits. The SCARF index uses 32 deficits, with each one defined from ICD-10 codes.

Outcomes

The primary outcome measure was the number of days alive and out of hospital (DAOH) at 90 days. The DAOH is a composite measure, which encompasses mortality and total length of hospital stay (LOS), including re-admissions. The DAOH has been previously validated as an outcome following surgery, and its choice was supported by the study's panel of ex-patients and public contributors. The calculation of DAOH used HES data on the total duration of hospitalisation over the 90-day period, and the date of death from linkage to the ONS death record. The sample size for each condition was projected to be sufficient to assess overall differences between the comparison groups in the mean DAOH of at least 1 day, with 80% power and 95% levels of statistical significance. The secondary outcomes were the two aspects of DAOH: (1) 90-day mortality and LOS and (2) any emergency re-admission within 30 days.

Analysis

For each qualifying emergency admission, the TTO was calculated as the proportion of patients admitted as eligible emergency admissions in that specific hospital who received ES in the previous 12 months, therefore, suggesting that the hospital's past preference for ES strongly predicts treatment choice for the current patient. The IV was used to address confounding in estimating the relative effectiveness of ES compared with NES strategies for the overall cohort and for the above prespecified subgroup variables. To further support the conditions underlying the IV, the analysis adjusted for the covariates described above, together with the time period and proxies for the quality of acute care in each hospital.

The sensitivity analyses assessed whether or not the results were robust to alternative assumptions, including different definitions of ES and the quality of acute care, and the use of regression adjustment, which assumed that there were no unmeasured confounders, rather than the IV approach.

Cost-effectiveness analysis

The cost-effectiveness analysis (CEA) took an acute hospital perspective and included all inpatient costs up to 1 year. Information on individual-level resource use was extracted from the HES data and included the procedures received and the total hospital LOS, including critical care and re-admissions. Unit costs were taken from a literature review and the NHS Payment by Results database, and combined with resource use, to report total costs per patient up to 1 year. We reported quality-adjusted life-years (QALYs) up to 1 year by combining survival time with health-related quality-of-life (HRQoL) estimates from the literature.

We estimated the relative cost-effectiveness of ES compared with NES strategies for each of the five acute conditions for the overall cohort and for each of the prespecified subgroups, adjusting for confounding with the IV approach. The cost-effectiveness metric was the incremental net monetary benefit (INB) with incremental QALYs, defined as the difference in QALYs between ES and NES strategies, which were then valued at £20,000 per QALY. We tested whether or not the results were robust to alternative sources for the HRQoL estimates, to different time horizons (i.e. 5 years vs. 1 year) and to adjustment for confounding with regression instead of the IV approach.

Results

Cohorts and instrumental variable validity

Cohort sizes were as follows: 268,144 admissions with appendicitis, 240,977 admissions with cholelithiasis, 138,869 admissions with diverticular disease, 106,432 admissions with a hernia and 133,073 admissions with an intestinal obstruction. The proportions of patients who had ES were 92.3% of admissions with acute appendicitis, 21.6% of admissions with cholelithiasis, 11.4% of admissions with diverticular disease, 58.8% of admissions with an abdominal wall hernia and 30.5% of admissions with an intestinal obstruction. There was wide variation in ES rates across NHS hospitals. The greatest variation was for cholelithiasis, with a median ES rate across hospitals of 8.4% (minimum, 2.3%; maximum, 66.4%), followed by hernia (median, 59.8%; minimum, 30.8%; maximum, 79.2%) appendicitis (median, 93.0%; minimum, 67.5%; maximum, 98.6%), intestinal obstruction (median, 30.0%; minimum, 20.4%; maximum, 51.4%) and diverticular disease (median, 11.2%; minimum, 3.5%; maximum, 21.0%). For each acute condition, the hospital-level TTO was strongly correlated with ES receipt (i.e. IV condition 1). The TTO balanced the observed baseline covariates, which increased confidence that the IV also balanced unmeasured confounders (i.e. IV condition 2).

Overall clinical effectiveness and cost-effectiveness results

Overall, the average numbers of DAOH were similar following ES and NES strategies, with mean differences of −0.73 [95% confidence interval (CI) −2.10 to 0.64] days for appendicitis, 0.60 (95% CI −0.10 to 1.30) days for cholelithiasis, −2.66 (95% CI −15.7 to 10.4) days for diverticular disease,

–0.07 (95% CI –2.40 to 2.25) days for hernias and 3.32 (95% CI –3.13 to 9.76) days for intestinal obstructions, after adjusting for confounding. Overall all-cause mortality and LOS at 90 days, and average costs and QALYs at 1 year, were all similar following ES and NES strategies for each of the five acute conditions. The sensitivity analyses generally reported similar results to the base-case analyses, aside from the regression analyses, the results of which highlighted the importance of adjusting for unmeasured confounding with the IV approach.

Clinical effectiveness and cost-effectiveness by patient subgroup

The relative clinical effectiveness and cost-effectiveness of ES compared with NES strategies differed by subgroup according to levels of frailty and the number of comorbidities. For patients with severe frailty, the mean numbers of DAOH at 90 days were smaller following ES compared with NES strategies, with mean differences of –21.0 (95% CI –27.4 to –14.6) days for appendicitis,–5.72 (95% CI –11.3 to –0.2) days for cholelithiasis,–38.9 (95% CI –63.3 to –14.6) days for diverticular disease, –19.5 (95% CI –26.6 to –12.3) days for hernias and–34.5 (95% CI –46.7 to –22.4) days for intestinal obstructions. For patients with severe frailty, ES was less cost-effective than NES strategies, with INB estimates of –£18,727 (95% CI –£23,900 to –£13,600) for appendicitis, –£7700 (95% CI –£13,000 to –£2370) for cholelithiasis,–£9230 (95% CI –£24,300 to £5860) for diverticular disease, –£16,600 (95% CI –£21,100 to –£12,000) for hernias and–£19,300 (95% CI –£25,600 to –£13,000) for intestinal obstruction.

For patients who were categorised as 'fit' (i.e. without frailty) at presentation, mean numbers of DAOH were larger for ES than for NES strategies for diverticular disease (5.35, 95% CI –2.56 to 13.28 days), hernias (2.26, 95% CI 0.37 to 4.15 days) and intestinal obstruction (18.2, 95% CI 14.8 to 22.47 days), and were similar for appendicitis (–0.18, 95% CI –1.56 to 1.20 days) and cholelithiasis (0.93, 95% CI 0.48 to 1.39 days). ES was relatively cost-effective for patients who were fit, with estimated INBs of £5180 (95% CI £684 to £9,680) for diverticular disease, £2040 (95% CI £996 to £3090) for hernias, £7850 (95% CI £5020 to £10,700) for intestinal obstruction, £369 (95% CI –£728 to £1460) for appendicitis and £718 (95% CI £294 to £1140) for cholelithiasis.

For patients with two or more comorbidities, NES strategies were, on average, more cost-effective than ES for patients with acute appendicitis, cholelithiasis or a hernia, but ES was more cost-effective for patients with multiple comorbidities who had diverticular disease or intestinal obstruction.

Limitations

The IV approach reduced, but could not eliminate, the risk of confounding. HRQoL measures were not available from routine data sources. The costs of outpatient visits or of care home stays were not considered.

Conclusions

Implications for health care

- For patients presenting as emergency hospital admissions with common acute gastrointestinal conditions, the ES and NES strategies for the overall cohort led to similar average numbers of DAOH at 90 days, and similar average hospitalisation costs and QALYs at 1 year, after addressing confounding with the IV approach. The CEA did not provide strong evidence that either strategy was more cost-effective in the overall populations of patients with each of the five acute conditions.
- For patients with severe levels of frailty, NES strategies were relatively cost-effective for each of the five acute conditions. For patients who were fit, ES was, on average, more cost-effective than NES strategies for each condition.

- The evidence from the ESORT study supports ongoing national initiatives to encourage frailty assessment for patients with acute conditions as part of any preoperative assessment following emergency admission. Frailty assessment can help identify those patients who would benefit more from a NES strategy. For some patients, a NES strategy may include later planned surgery, which can enable a patient's long-term conditions to be optimised.
- The evidence from the ESORT study can help inform service providers, patient and carers about the relative benefits and risks of ES compared with NES strategies, recognising that these benefits and risks differ according to factors beyond the patients chronological age, including their frailty level and number of comorbidities.

Recommendations for further research

The following recommendations are listed in priority order:

- The ESORT study highlighted the importance of the appropriate choice of ES or NES strategy for patients with comorbidities, which includes patients with multiple long-term conditions. For patients with multiple long-term conditions, further research is required to assess the benefits and costs of prompt intervention with ES, compared with later surgery. Delaying surgery could help to 'optimise' the management of these patients by, for example, modifying concomitant medications.
- Future RCTs of ES compared with NES strategies are warranted, and may be most feasible and useful for some patients with an abdominal wall hernia, as this is the population for whom there appears clear equipoise between ES and NES strategies.
- Further research is required on patients' HRQoL following ES and NES strategies.
- Following the COVID-19 pandemic, in England, NHS waiting lists are projected to reach 13 million by 2025, with implications for hospitals' capacity for elective surgery and ES. Further research in other conditions would, therefore, be useful to identify which patient groups ES or NES strategies are clearly cost-effective. The approach taken in the ESORT study, that is, combining large-scale routine data with advanced analytical approaches to address confounding, could be extended to other acute conditions. This approach could identify patient groups for whom either 'early' or 'later' intervention is more cost-effective.

Study registration

This study is registered as reviewregistry784.

Funding

Chapter 1 Introduction

Background and rationale

Emergency surgery (ES) poses a considerable global burden to health systems and is associated with high morbidity, mortality and resource utilisation.[1] ES accounts for approximately 750,000 admissions per year in England alone,[2] with surgical procedures representing approximately 10% of the annual NHS budget.[3] Following the COVID-19 pandemic, in England, NHS waiting lists are projected to reach 13 million by 2025, with implications for hospitals' capacity for elective surgery and ES.[4] For common acute conditions that present as emergency admissions, an important clinical decision is which patients should receive ES and which patients should receive non-emergency surgery (NES) strategies, which can include medical management, non-surgical procedures (e.g. radiological-guided drainage of abscess) or surgery deferred to the elective (planned) setting.

In England, although there are approximately 4000 NHS consultant general surgeons who spend, on average, 50% of their time on emergency general surgery admissions, there is insufficient capacity to provide ES 24 hours a day, 7 days a week.[2] In 2016, there were 697,314 emergency general surgical admissions to NHS trusts in England, of which 305,507 (43.8%) did not receive an operative procedure.[5] The Getting it Right First Time report[2] for emergency general surgery found wide variation across NHS trusts in care quality and outcomes after ES, which may reflect local logistical and resource constraints, but also clinical uncertainty.[5]

Emergency surgery rates for patients with acute gastrointestinal conditions have declined over the last 20 years,[6] and protocols for NES strategies have been developed and implemented as part of randomised controlled trials (RCTs).[7–12] However, relatively few RCTs in the emergency setting have compared ES and NES strategies for patients with acute conditions and, in general, these trials have included an insufficient number and range of patients to inform routine service provision. For patients with uncomplicated acute appendicitis, several RCTs and ensuing meta-analyses have compared ES with NES strategies, but the evidence is equivocal.[7,8,13] For patients with acute cholecystitis, published RCTs have reported reduced complications following 'early' and 'delayed' surgery, but have not considered non-operative strategies. For patients with acute diverticular disease, NES strategies are well developed, but published RCTs have considered different forms of ES, rather than comparing ES with NES strategies.[14,15] For other acute conditions presenting as emergency admissions, such as abdominal wall hernia and intestinal obstruction, no published RCTs of ES versus NES strategies exist, amid ethical concerns about randomisation.[16]

In observational studies comparing ES with NES strategies, the major concern is confounding by indication (i.e. patients who receive ES may be sicker).[17] Traditional risk adjustment methods are unable to fully allow for prognostic differences between the patients receiving NES strategies because information, for example from radiological investigation, is not available within routine data sources. As these unmeasured variables may predict both ES receipt and outcome, these studies are liable to provide biased estimates of the effectiveness of ES.[17]

The appeal of an instrumental variable (IV) design is that it can provide accurate estimates of treatment effectiveness even when there are unmeasured differences between the comparison group.[18,19] Keele et al.[20] developed an IV design to address confounding when evaluating ES in the USA using claims data. Keele et al.'s study[20] used the surgeon's ES rate across preceding emergency admissions as its IV, and reported that this variable appeared to meet the criteria required to be a valid IV. Keele et al.[20] found that for some conditions, including diverticular disease, ES led to higher 30-day mortality than non-operative care; however, these results may not apply to the NHS in the UK,

where thresholds for ES may be different. In addition, although the provider's preference for ES appeared to be a valid IV in the context of US claims data, it would require further assessment before it could be used in assessing the effectiveness of ES from routine data in the UK.

A further important gap in the evidence required to inform ES provision is that few previous studies have evaluated the cost-effectiveness of ES compared with NES strategies for these common acute conditions. For patients with appendicitis, previous cost-effectiveness analysis (CEA), like the precedent comparative effectiveness literature, provide equivocal results. Some studies have reported that NES strategies result in improved outcomes and lower costs than ES.[21-23] By contrast, other studies report that ES leads to improved outcomes, at either additional[8] or reduced[24] average costs. For patients with acute cholecystitis, studies have provided conflicting evidence, with some studies[25,26] reporting that ES is more cost-effective than delayed surgery and other studies[27] reporting that outcomes are better and costs are lower if surgery is delayed. None of these studies have considered NES strategies other than delayed surgery and, similarly, for patients with other common acute conditions, including acute diverticular disease, abdominal wall hernia and intestinal obstruction, there is little available evidence on the cost-effectiveness of ES compared with NES strategies.

A final concern is that previous studies evaluating ES strategies for patients with common acute conditions presenting as emergency admissions have failed to include patients from sufficiently broad populations to report relative clinical effectiveness and cost-effectiveness across population subgroups.[7-17,21-27] Reports by the Royal College of Surgeons (RCS) of England (London, UK) have generally found lower levels of ES for patients aged > 75 years than for patients aged 65–74 years, notably for patients with acute cholelithiasis.[28,29] The RCS guidelines have emphasised that research is required that simultaneously considers which factors, beyond age, may modify the relative clinical effectiveness and cost-effectiveness of ES. Such studies should consider factors such as the patients' frailty levels at emergency admission and whether or not patients have multimorbidity. Although such factors may be related to age, the factors may still modify the clinical effectiveness and cost-effectiveness of ES compared with NES strategies after allowing for biological age.

Aims and objectives

The aim of the Emergency Surgery OR noT (ESORT) study was to evaluate the clinical effectiveness and cost-effectiveness of ES compared with NES strategies for patients with common acute conditions presenting as emergency admissions to NHS hospitals in England. The specific objectives were to evaluate the:

- relative effectiveness of ES compared with NES strategies for five common acute conditions presenting as emergency admissions
- relative cost-effectiveness of ES compared with NES strategies for five common acute conditions presenting as emergency admissions
- clinical effectiveness and cost-effectiveness of ES compared with NES strategies for specific patient subgroups.

Changes to the research proposed

Reduction from seven to five acute conditions

The original proposal was for the target population to include seven acute conditions (i.e. acute appendicitis, cholelithiasis, diverticular disease, abdominal wall hernia, intestinal obstruction, acute intestinal ischaemia and acute complicated peptic ulcer disease) that present as emergency hospital admissions. As proposed during the first 6 months of the research, we applied and received Hospital Episode Statistics (HES) data for all seven conditions, and used these data to refine the specific

definitions for each subpopulation. We then presented summary information and discussed the suitability of including each of the seven conditions. We judged suitability according to the following three considerations:

1. From the *International Statistical Classification of Diseases and Related Health Problems*, Tenth Revision (ICD-10) codes available in the HES data, was it possible to define a homogenous population or subpopulation for which there was equipoise between ES and NES strategies?
2. Was it possible to use Office of Population Censuses and Surveys version 4 (OPCS-4) procedure codes and information on the timing of procedures to provide a clear definition of ES and NES strategies?
3. Was it anticipated that the study could provide evidence on the relative clinical effectiveness and cost-effectiveness that would be useful for clinical and health service decision-making?

Following careful consideration of each condition against these criteria, we decided that it would be inappropriate to include two of the conditions (i.e. acute intestinal ischaemia and acute complicated peptic ulcer disease) for the following reasons. First, our initial investigation using HES data suggested that given the different aetiologies that patients present with for both acute intestinal ischaemia and acute complicated peptic ulcer disease, there would be considerable heterogeneity in the patient population, and it would not be possible to define homogeneous subpopulations, for which there was equipoise between the comparator strategies from the ICD-10 codes available in the HES data. Second, the view of the clinical co-applicants and advisors was that the patient heterogeneity would make providing a clear definition of the ES intervention and comparator very challenging for these two conditions. Third, it was judged unlikely that the study would, therefore, be able to provide estimates of relative clinical effectiveness and cost-effectiveness that could inform practice.

For the other five conditions, we judged that the requisite criteria were met. It was recognised that the priority was to refine the definition of the inclusion criteria and of ES for these five conditions, and this would warrant careful consideration by a wider clinical panel (see *Chapter 2, Methods, Clinical panel criteria*).

Additional years of data

The original proposal was to study the period up to 31 March 2018. Owing to delays in accessing the data, and the onset of the COVID-19 pandemic in early 2020, the team decided to amend the study protocol and to apply for data up to 30 June 2020. By extending the time period, this increased the available sample size and enabled the study to provide more precise estimates of the clinical effectiveness and cost-effectiveness of ES. This enabled cohorts of patients to be included who had an emergency admission up to 31 December 2019.

All panels virtual

A further change to the study was in response to the impact of the COVID-19 pandemic. To continue to solicit the input of two key groups [i.e. clinicians and patient and public involvement (PPI) members], it was necessary to host the panels and workshops online, and this meant that the first clinical panel was conducted entirely remotely, with an initial meeting carried out over Zoom (Zoom Video Communications, San Jose, CA, USA), followed up by submissions sent and received via e-mail.

The two PPI workshops at the design stage were also carried out over Zoom, with a detailed pre-brief provided in advance via the Sway platform (Microsoft Corporation, Redmond, WA, USA).

Although this change was a necessity because of the situation with the pandemic, we found that the virtual offering may have improved participation, particularly with the PPI workshops, given that it did not preclude individuals taking part who would have faced challenges attending in person due to, for example, mobility issues or caring responsibilities.

Public and patient involvement

The PPI work in the ESORT study sought to develop a clearer understanding of what is important for patients, their families and the public in general when someone arrives at a hospital in an emergency. Patient perspectives were embedded into the ESORT study from the very start, with close involvement from two patient advocates in the application to National Institute for Health and Care Research (NIHR), PPI panels held at the design and translation stages of the project, and review by the North Thames Applied Research Collaboration Research Advisory Panel at key stages.

We followed the GRIPP2 (Guidance for Reporting Involvement of Patients and the Public 2) methodology in reporting PPI processes, and the table is available as *Report Supplementary Material 1* and on our study website.[30]

Public and patient involvement strategy

The PPI strategy was developed with co-applicant Paul Charlton (a NIHR Patient Research Ambassador with a broad public contributor involvement in health research) and Stephen Harkins (who had experienced ES for a common condition). Both Paul Charlton and Stephen Harkins took part in conference calls to critique the design before each stage of the funding application. Paul Charlton and Stephen Harkins felt that as well as the study needing input from patients who had experienced ES, it was important to include patients who had not. In addition, Paul Charlton and Stephen Harkins felt that the current situation in the NHS, with some patients with the same condition having ES and others not, could result in adverse consequences for patients. Both contributors felt that generating better evidence about which patients should have ES was an important issue for patients and the public.

Stephen Harkins shared his experience of decision-making at the time of hospitalisation, and how the emergency context shaped events. Both PPI contributors helped agree a detailed plan for the PPI design workshop, which was run separately to the clinician panel, which focused on clinical issues. Both workshops fed into the study design, but running them separately enabled the study to draw on the different skill sets required to meet the respective aims. The PPI design workshops benefited from preparatory training material, which helped panel members understand data sources, the complexities of hospital-based decision-making and the key terminologies used.

Public and patient involvement panels

We held two (virtual) sets of panels, two at the design stage (July 2020) and two at the translation stage (September 2021). We invited participants who:

- had experienced coming into hospital as an emergency due to appendicitis, diverticulitis or gallstones
- were a family member/carer of someone who has had this experience
- had not had this experience but wanted to help the study.

The PPI design stage workshops were held in July 2020, and the findings of the workshops helped inform the design of the ESORT study (see *Report Supplementary Material 2*). Full details are provided on the ESORT website.[30] In brief, each workshop was attended by seven panellists, each of whom had experience of an emergency admission with one of the conditions of interest. In advance of the workshop, participants were provided with bespoke preparatory material via Sway. The workshop participants were asked to consider outcome measures for patients following emergency admission to hospital for acute conditions, and the group agreed that an appropriate measure would capture mortality and the number of days in hospital (see *Chapter 3*). The participants also expressed interest in joining the 'translation workshop' to discuss the study's results (see *Chapter 5*), in helping to co-produce the study's lay summary (see *Plain English summary*) and in providing further details on the ESORT study website.[30]

Report overview

This report details the three interlinked components of the study. *Chapter 2* explains how the patient populations and the ES and NES strategies were defined from the consensus of the clinical panel, and then from the HES data. *Chapter 2* also highlights the wide variation in ES rates across NHS acute trust hospitals for common acute conditions presenting as emergency hospital admissions. This unexplained variation underlies the IV design, which is used to address confounding in assessing relative effectiveness (see *Chapter 3*) and cost-effectiveness (see *Chapter 4*) of ES. In *Chapter 3*, we provide a summary of the methods for assessing clinical effectiveness, and summarise the results both overall (objective 1), and according to prespecified subgroups (objective 3). Similarly, in *Chapter 4*, we summarise the main aspects of the CEA, and provide the results overall (objective 2) and according to subgroups (objective 3).

Each of the three substantive chapters are also supported by additional material that is provided in the appendices, report supplementary material and on the ESORT study website.[30] In *Chapter 5*, we discuss the overall findings from the study, including findings from the PPI translation workshops, and provide recommendations for clinical practice and for further research.

Chapter 2 Cohort description

Introduction

The design of the ESORT study was based on the principles of a 'target trial', that is, an observational study that seeks to emulate a (hypothetical) pragmatic trial.[31,32] Three key issues with applying a target trial framework to routine observational data are: (1) defining the target population according to inclusion and exclusion criteria that would be applied in a hypothetical trial (see *Study eligibility*), (2) defining the intervention and comparator strategies (see *Clinical panel criteria*) and (3) defining the time when the eligibility criteria are met and a decision on treatment strategy made, or time zero [see *Definition of day (time) zero*]. The fourth and fifth key issues (i.e. those of defining the outcomes of interest and of the approach to handle unmeasured confounding in the absence of randomisation) are addressed in *Chapter 3* (see *Chapter 3*, *Outcomes*, *Instrumental variable: the tendency to operate*, *Patient-level covariates* and *Statistical analysis*). Conceptualising the observational study in this way encourages researchers to make explicit decisions about the study design and the assumptions required to identify causal treatment effects.

The ESORT study used routine HES data for England to define cohorts of emergency admissions for five common acute gastrointestinal conditions: (1) acute appendicitis, (2) cholelithiasis, (3) diverticular disease, (4) abdominal wall hernia and (5) intestinal obstruction. This chapter describes (1) the criteria for defining the cohorts, (2) the criteria for identifying patients who had ES, (3) the patient characteristics of the five cohorts and (4) variation in ES rates across hospitals.

Methods

Study data

The study used HES admitted patient care data that were provided under a data-sharing agreement with NHS Digital.[33] Linked HES adult critical care data and Civil Registration date of death data were also provided for deriving resource use and outcomes. The HES admitted patient care data comprised all adult emergency admissions for patients with relevant three-character ICD-10 diagnosis codes for the five conditions, together with patients' historic and subsequent admissions between 1 April 2009 and 30 June 2020. Each admission included one or more finished consultant episodes, which provided data on interventions and procedures (using OPCS-4 procedure codes) carried out while under the consultant's care. The pseudonymised HES admitted patient care data also included information on patients' sociodemographic characteristics (i.e. age, sex, ethnicity and Index of Multiple Deprivation by decile) and other administrative and organisational information, such as dates of admission and discharge and the NHS trust and hospital the patient was admitted to.

The aim was to create a cohort of eligible patient admissions for each condition. The first episode meeting eligibility criteria was defined as the index episode. The admission containing the index episode was defined as the index admission.

The RCS of England's Charlson Comorbidity Index score was derived from data by using ICD-10 diagnosis codes from the index episode and from other episodes in the prior 12 months.[34] The score comprises four categories: (1) no comorbidities, (2) one comorbidity, (3) two comorbidities and (4) three or more comorbidities.

An indicator of frailty was derived using the Secondary Care Administrative Records Frailty (SCARF) index.[35] The SCARF index is based on the cumulative deficits model of frailty and is designed to

capture 32 deficits that cover functional impairment, geriatric syndromes, nutrition problems, cognition and mood, and medical comorbidities. The SCARF index uses ICD-10 diagnostic codes from the index admission and from other episodes up to 2 years previously, with patients categorised as either fit or having mild, moderate or severe frailty.

Study eligibility

General study eligibility criteria for inclusion in the five cohorts were as follows:

- The index episode occurred in a general acute hospital between 1 April 2010 and 31 December 2019 (see *Hospital eligibility*). Data for the year prior to 1 April 2010 were used for deriving measures of comorbidity and frailty, and for calculating the IV (see *Chapter 3*). The end date of 31 December 2019 was chosen to minimise any impact that COVID-19 might have on the 90-day follow-up period used for the study's primary outcome.
- The index episode included a primary diagnosis with an ICD-10 diagnosis code that was judged relevant according to the consensus of a clinical panel. For the intestinal obstruction cohort, a relevant diagnosis was allowed in the second HES diagnosis field if the primary diagnosis was colorectal cancer.
- The index admission was an emergency admission through the emergency department or from a primary care referral. This was to ensure that the cohorts included typical admissions and that transfers from other NHS trusts (e.g. to a more specialist tertiary referral centre) were excluded.
- The index episode was under a consultant general surgeon, subspecialty general surgeon or a surgeon working in the general surgery specialty.
- The index episode was the first or second episode within the admission.

An index episode was deemed ineligible if any of the following criteria were met:

- An emergency admission that included a relevant main diagnosis that had occurred in the 12 months prior to the index episode.
- An ICD-10 diagnosis for a condition deemed as an exclusion criterion according to the consensus of a clinical panel appeared in the index episode or any earlier episode within the index admission.
- The index admission included a transfer between hospitals within a NHS trust before the index episode.
- A procedure defined as ES by consensus of a clinical panel occurred in any elective or emergency admission up to 90 days before the start of the index episode.
- The index admission lacked a final episode that indicated the admission was complete and provided the patient's status at discharge.

Clinical panel criteria

A clinical panel of 11 surgeons and one anaesthetist was convened and met twice in total (once in March 2020 and once in April 2020). The panel had three main purposes: (1) to refine the inclusion and exclusion criteria for defining the study population, (2) to refine the list of procedures within the definition of 'ES' and (3) to define 'the most appropriate' time window that constitutes 'ES'. Results from the panel's first meeting were collated in a summary report and circulated to panellists prior to the second meeting.

Cohort inclusion and exclusion criteria for each condition except hernia were discussed at the first meeting. The panellists' views on criteria for hernia were elicited by e-mail after the meeting and incorporated in the summary report. A list of ICD-10 diagnosis codes and potential reasons for exclusion were compiled and discussed at the second meeting. Panellists privately indicated their agreement, or otherwise, with each potential inclusion or exclusion criterion. Inclusion criteria required the support of at least 75% of the panel and exclusion criteria required the support of at least 25% of the panel.

The panel's assessment of procedures counting as ES followed a Delphi process. A list of OPCS-4 procedure codes was drawn-up and presented at the first meeting. Panellists discussed and then privately rated the list of procedures. Panellists were also asked to indicate the time window, that is, the number of days from time zero (i.e. the first day of the index episode under the care of a general surgeon) within which surgery must occur to be regarded as ES. The time window question was asked separately for surgery in the index admission and surgery within a re-admission to allow for a panellist's opinion to differ according to whether or not surgery in a discharged and re-admitted patient could count as ES, compared with one for whom surgery took place within the index admission. ES was defined as procedures, with the support of at least 50% of the panel, that fell within a time window defined according to the median panel rating. The results from the panel survey were used to define the main (base-case) analyses for the subsequent assessment of clinical effectiveness, with differences of opinion reflected in sensitivity analyses (see *Chapter 3, Sensitivity analyses*).

A full list of panel ratings for all diagnoses and procedures for the five conditions is included *Report Supplementary Material 3*. A summary of panel decisions on inclusion and exclusion criteria is shown in *Table 1*. Decisions of the panel resulted in the cholelithiasis cohort including only calculus of the gall bladder, and the diverticular disease cohort including only diverticular disease of the large intestine. Incisional and parastomal hernias were not included in the hernia cohort. The panel supported exclusions to three cohorts (appendicitis, hernia and intestinal obstruction) for reasons such as pregnancy, ischaemia and specific cancer diagnoses.

A majority of procedures rated by the panel were defined as ES for four of the conditions (*Table 2*). The exception was cholelithiasis, for which less than one-quarter of procedures were defined as ES. Two post hoc changes to the panel's classification of ES procedures were made (see footnotes to *Table 2*).

Hernia surgery had the strictest timing threshold of 3 days for both an ES procedure in the index admission or in a re-admission. The threshold for intestinal obstruction, cholelithiasis and appendicitis was 7 days in either an index admission or a re-admission. For diverticular disease, surgery could be classified as ES if it occurred at any time in an index admission or within 14 days in a re-admission.

TABLE 1 Clinical panel-derived diagnostic inclusion and exclusion criteria summary

	Condition				
Criterion	**Appendicitis**	**Cholelithiasis**	**Diverticular disease**	**Hernia**	**Intestinal obstruction**
Inclusion criteria					
Agreed for inclusion	All included	Calculus of gall bladder	Large intestine	Inguinal; femoral; umbilical; ventral	Intestinal adhesions, Intussusception; volvulus; gallstone ileus; other obstruction
Dropped from inclusion	None	Calculus of bile duct; other cholelithiasis	Small intestine; small and large intestine; unspecified	Incisional; parastomal	Paralytic ileus; other impaction; ileus, unspecified
Exclusion criteria					
Agreed for exclusion	Pregnancy; Appendiceal cancer			Pregnancy; ischaemia; cancer	Colorectal cancer with metastases; gynaecological cancer; ischaemia

TABLE 2 Clinical panel-derived definitions of ES for each condition

Definition	Condition				
	Appendicitis	Cholelithiasis	Diverticular disease	Hernia	Intestinal obstruction
Procedures defined as ES	21 of 33	11 of 48	45 of 57	52 of 59[a]	111 of 140
Common procedures excluded from definition of ES	Unspecified other excision of appendix[b]	Endoscopic sphincterotomy	Image-controlled percutaneous drainage	None	None
Threshold for a procedure in the index admission to be ES	7 days	7 days	Any time	3 days	7 days
Threshold for a procedure in a re-admission to be ES	7 days	7 days	14 days	3 days	7 days

a Six additional umbilical hernia procedures were added to the list of ES procedures (T241, Repair of umbilical hernia using insert of natural material; T248, Other specified primary repair of umbilical hernia; T971, Repair of recurrent umbilical hernia using insert of natural material; T973, Repair of recurrent umbilical hernia using sutures; T978, Other specified repair of recurrent umbilical hernia; T979 Unspecified repair of recurrent umbilical hernia). These additional umbilical hernia procedures were not included in the original list of procedures reviewed by the panel because umbilical hernia was not included in the earlier version of the HES data set used to identify potential ES procedures. The six procedures were added after the exercise for consistency with similar procedures for inguinal, femoral and ventral hernias.
b H029, Unspecified other excision of appendix was supported by 5 of 12 panellists and was initially left out of the list of ES procedures. Later review of use of this code across hospitals identified some hospitals routinely using H029 for a majority of emergency appendicitis admissions. It was subsequently decided to include H029 as an ES procedure instead.

A cohort patient who did not meet the criteria for ES was categorised as having had a NES strategy, which could involve medical management, another operative procedure, or an ES procedure after the requisite time window.

Definition of day (time) zero
One challenge the study faced was to define day zero (i.e. the analogue in a non-randomised study to the time of randomisation). Day zero should correspond to the time when the eligibility criteria are met, and the treatment strategies commence. By establishing day zero, this can help the study reduce confounding by indication due to prognostic difference between the comparison groups. In the ESORT study, emulating the target trial's day zero raised several challenges.

For patients who had ES, the date of surgery was available from the HES inpatient data; however, for patients who had NES, the date of initiation of non-operative strategies, such as antibiotic therapy, was not recorded. Likewise, information on the date of diagnostic imaging procedures, which could be used to inform treatment assignment, was not available for all patients. Faced with these challenges, there was no perfect choice for day zero. If day zero was defined as the date of admission, then this could have led to a high risk of confounding as, for some patients, it would 'pre-date' the assessment of eligibility, and so their prognosis could differ prior to the selection of the treatment strategy.[31] Instead, the study used the date within the first eligible hospital episode from which the patient was first under the care of a consultant surgeon as day zero.

The rationale for this choice of day zero was threefold. First, this date marks the end of the eligibility assessment period, which reduces the risk of confounding due to a change in prognosis prior to treatment selection. Second, given the acute nature of these conditions, the delay between assessment by the surgeon to the end of the time window for defining the two strategies is short and, therefore, the risk of immortal time bias is minimal.[31] Third, information on this variable is available for all patients because our eligibility criteria require that all patients were at some point under the care of a consultant surgeon (see Study eligibility).

Hospital eligibility

The IV the tendency to operate (TTO) was calculated at the level of the hospital. A total of 175 acute general hospitals, with at least 200 emergency general surgery admissions per year, were included. These hospitals were identified using the site of treatment or five-character provider codes in HES. Identification of hospitals allowed for changes to codes owing to organisational changes, for example mergers of NHS trusts or the replacement of a hospital with a 'new build' hospital on a different site nearby. Reconfiguration of hospital services within and between NHS trusts has led to some hospitals closing permanently or routine emergency general surgery activity stopping before 31 December 2019. Monthly emergency general surgery activity in these hospitals was examined and judgement used to identify a date when routine activity ceased. Seventeen such hospitals were identified and contributed cohort patients only up to the date routine activity was deemed to have ceased. One new hospital, a specialist emergency centre, started ES activity after 1 April 2010. Therefore, 174 hospitals were contributing cohort patients in April 2010 and 158 hospitals were contributing patients in December 2019.

Results

Characteristics of the cohorts

Cohort sizes were as follows: 268,144 admissions with appendicitis, 240,977 admissions with cholelithiasis, 138,869 admissions with diverticular disease, 106,432 admissions with a hernia and 133,073 admissions with an intestinal obstruction (*Table 3*). The most common reasons for excluding emergency admissions from the cohort were that the patient was not under the care of a general surgeon (15.6% of cholelithiasis admissions, 13.3% of diverticular disease admissions and 14.6% of

TABLE 3 Application of ESORT study inclusion and exclusion criteria for emergency admissions to 175 acute NHS hospitals in England from April 2010 to December 2019

Inclusion/exclusion criteria	Condition, *n* (%)				
	Appendicitis	**Cholelithiasis**	**Diverticular disease**	**Hernia**	**Intestinal obstruction**
Meet panel inclusion criteria	307,890	365,791	200,021	146,601	236,791
Exclusions					
No episode with a consultant surgeon	8582 (2.8)	56,913 (15.6)	26,606 (13.3)	9457 (6.5)	34,475 (14.6)
No eligible diagnosis in the first two episodes	1492 (0.5)	5593 (1.5)	5501 (2.8)	1761 (1.2)	7208 (3.0)
Not admitted through A&E or a GP	21,053 (6.8)	21,637 (5.9)	12,182 (6.1)	11,477 (7.8)	13,595 (5.7)
Clinical panel exclusion criteria	1443 (0.5)	0 (0.0)	0 (0.0)	7574 (5.2)	10,512 (4.4)
Missing discharge data	351 (0.1)	274 (0.1)	146 (0.1)	115 (0.1)	340 (0.1)
Transfer between hospitals before index episode	634 (0.2)	588 (0.2)	406 (0.2)	173 (0.1)	503 (0.2)
Other admission meeting inclusion criteria in previous 12 months	3580 (1.2)	38,812 (10.6)	15,544 (7.8)	7908 (5.4)	31,686 (13.4)
ES prior to index episode	2518 (0.8)	406 (0.1)	573 (0.3)	842 (0.6)	1123 (0.5)
ES procedure in prior admission within 90 days	93 (0.0)	591 (0.2)	194 (0.1)	862 (0.6)	4276 (1.8)
Included in cohort	268,144	240,977	138,869	106,432	133,073

A&E, accident and emergency; GP, general practitioner.

intestinal obstruction admissions) or that the patient had been admitted as an emergency for the same condition in the previous 12 months (10.6% of cholelithiasis admissions and 13.4% of intestinal obstruction admissions).

Patient characteristics in the five cohorts are shown in *Table 4*. Patients with appendicitis were typically younger, with around half aged < 35 years and less than 1% aged ≥ 85 years. The other four cohorts included a higher percentage of patients aged ≥ 85 years (5.6% of cholelithiasis admissions, 9.9% of diverticular disease admissions, 12.3% of hernia admissions and 15.1% of intestinal obstruction admissions). There were substantially more females in the cholelithiasis and diverticular disease cohorts and substantially more males in the hernia cohort.

TABLE 4 Patient characteristics of the five cohorts

| Characteristic | Condition | | | | |
	Appendicitis (N = 268,144), n (%)	Cholelithiasis (N = 240,977), n (%)	Diverticular disease (N = 138,869), n (%)	Hernia (N = 106,432), n (%)	Intestinal obstruction (N = 133,073), n (%)
Age category (years)					
< 25	63,373 (23.6)	12,108 (5.0)	308 (0.2)	2263 (2.1)	2070 (1.6)
25–29	37,570 (14.0)	15,313 (6.4)	1076 (0.8)	3136 (3.0)	2181 (1.6)
30–34	31,388 (11.7)	16,446 (6.8)	2466 (1.8)	3998 (3.8)	2643 (2.0)
35–39	25,482 (9.5)	16,087 (6.7)	4648 (3.4)	4715 (4.4)	3341 (2.5)
40–44	21,657 (8.1)	17,748 (7.4)	7581 (5.5)	6075 (5.7)	4554 (3.4)
45–49	19,800 (7.4)	20,572 (8.5)	11,472 (8.3)	7757 (7.3)	6529 (4.9)
50–54	17,422 (6.5)	21,084 (8.8)	13,999 (10.1)	8203 (7.7)	8223 (6.2)
55–59	13,841 (5.2)	19,727 (8.2)	14,061 (10.1)	7934 (7.5)	9271 (7.0)
60–64	11,149 (4.2)	18,853 (7.8)	13,657 (9.8)	8305 (7.8)	11,076 (8.3)
65–69	9456 (3.5)	19,736 (8.2)	14,310 (10.3)	9163 (8.6)	13,899 (10.4)
70–74	6988 (2.6)	18,906 (7.9)	14,645 (10.6)	10,318 (9.7)	15,860 (11.9)
75–79	4727 (1.8)	16,796 (7.0)	14,084 (10.1)	10,755 (10.1)	16,830 (12.7)
80–84	3019 (1.1)	14,132 (5.9)	12,874 (9.3)	10,825 (10.2)	16,390 (12.3)
85–89	1604 (0.6)	9033 (3.8)	9137 (6.6)	8245 (7.8)	12,566 (9.4)
≥ 90	668 (0.3)	4436 (1.8)	4551 (3.3)	4740 (4.5)	7640 (5.7)
Sex					
Female	123,452 (46.0)	162,791 (67.6)	81,870 (59.0)	37,565 (35.3)	69,977 (52.6)
Male	144,679 (54.0)	78,177 (32.4)	56,996 (41.0)	68,863 (64.7)	63,093 (47.4)
Missing	13	9	3	4	3
Ethnicity					
Black/black mixed	6398 (2.7)	4750 (2.1)	2129 (1.6)	2602 (2.6)	3308 (2.5)
Asian/Asian mixed	12,714 (5.3)	11,328 (5.0)	2417 (1.8)	3590 (3.6)	4305 (3.4)
White	211,339 (88.0)	207,123 (90.7)	126,040 (95.2)	90,909 (91.8)	117,951 (92.3)
Chinese and other	9762 (4.1)	5089 (2.2)	1875 (1.4)	1961 (2.0)	2284 (1.8)
Missing	27,931	12,687	6408	7370	5225

TABLE 4 Patient characteristics of the five cohorts (*continued*)

Characteristic	**Condition**				
	Appendicitis (N = 268,144), n (%)	**Cholelithiasis (N = 240,977), n (%)**	**Diverticular disease (N = 138,869), n (%)**	**Hernia (N = 106,432), n (%)**	**Intestinal obstruction (N = 133,073), n (%)**
Deprivation quintile					
1 (most deprived)	53,814 (20.4)	56,424 (23.7)	24,980 (18.0)	22,845 (21.7)	23,334 (17.7)
2	54,361 (20.6)	50,640 (21.2)	27,280 (19.8)	21,893 (20.8)	25,348 (19.3)
3	53,331 (20.2)	48,202 (20.2)	29,076 (21.1)	21,702 (20.6)	27,924 (21.2)
4	51,716 (19.6)	44,378 (18.6)	29,224 (21.2)	20,495 (19.4)	27,858 (21.2)
5 (least deprived)	50,543 (19.2)	38,971 (16.3)	27,138 (19.7)	18,472 (17.5)	27,105 (20.6)
Missing	4379	2362	1171	1025	1504
Comorbidity					
None	222,846 (83.1)	157,485 (65.4)	83,246 (60.0)	65,513 (61.6)	69,648 (52.3)
One	39,710 (14.8)	62,150 (25.8)	39,588 (28.5)	29,657 (27.9)	42,145 (31.7)
Two	4750 (1.8)	17,047 (7.1)	12,676 (9.1)	8961 (8.4)	16,646 (12.5)
Three or more	838 (0.3)	4295 (1.8)	3359 (2.4)	2301 (2.2)	4634 (3.5)
Frailty index					
Fit	221,811 (82.7)	149,029 (61.8)	72,108 (51.9)	56,885 (53.5)	60,352 (45.4)
Mild frailty	38,596 (14.4)	66,237 (27.5)	44,482 (32.0)	32,712 (30.7)	44,008 (33.1)
Moderate frailty	6196 (2.3)	19,560 (8.1)	16,139 (11.6)	12,347 (11.6)	20,004 (15.0)
Severe frailty	1541 (0.6)	6151 (2.6)	6140 (4.4)	4488 (4.2)	8709 (6.5)

The majority of patients presenting as emergency admissions with appendicitis had no comorbidity and only 2.1% had at least two comorbidities. The percentage of patients with at least two comorbidities was higher for the other four conditions (cholelithiasis, 8.9%; diverticular disease, 11.5%; hernia, 10.6%; intestinal obstruction, 16.0%).

Patients with appendicitis were also least likely to be categorised as having any frailty and < 1% of patients with appendicitis were categorised as having severe frailty. The percentages of patients with any frailty were as follows: 17.3% of patients with appendicitis, 38.2% of patients with cholelithiasis, 48.1% of patients with diverticular disease, 46.5% of patients with a hernia and 54.6% of patients with an intestinal obstruction. Severe frailty was less common and reported in 2.6% of patients with cholelithiasis, 4.4% of patients with diverticular disease, 4.2% of patients with a hernia and 6.5% of patients with an intestinal obstruction.

The proportion of patients in the cohort admitted as emergency admissions who met the criteria for ES was highest for acute appendicitis (92.3%), lower for hernia (58.8%), intestinal obstruction (30.5%) and cholelithiasis (21.6%), and lowest for diverticular disease (11.4%) (*Table 5*). For two conditions (i.e. appendicitis and cholelithiasis), the patients who had ES were, on average, younger than the patients who had NES strategies.

Patients who received NES strategies were also more likely to have comorbidities than patients who received ES. In the appendicitis cohort, the percentage of patients with two or more comorbidities who received ES was 1.7% and the percentage of patients with two or more comorbidities who received

TABLE 5 Patient characteristics of the five cohorts by ES and NES groups

Characteristic	Appendicitis (n = 268,144)		Cholelithiasis (n = 240,977)		Diverticular disease (n = 138,869)		Hernia (n = 106,432)		Intestinal obstruction (n = 133,073)	
	ES (n = 247,506, 92.3%)	NES strategies (n = 20,638, 7.7%)	ES (n = 52,004, 21.6%)	NES strategies (n = 188,973, 78.4%)	ES (n = 15,772, 11.4%)	NES strategies (n = 123,097, 88.6%)	ES (n = 62,559, 58.8%)	NES strategies (n = 43,873, 41.2%)	ES (n = 40,550, 30.5%)	NES strategies (n = 92,523, 69.5%)
Female, n (%)	113,224 (45.8)	10,228 (49.6)	36,864 (70.9)	125,927 (66.6)	8698 (55.2)	73,172 (59.4)	25,035 (40.0)	12,530 (28.6)	23,269 (57.4)	46,708 (50.5)
Mean (SD) age (years)	38.3 (16.3)	47.3 (20.2)	50.7 (17.7)	56.1 (19.2)	63.9 (14.8)	64.0 (15.7)	63.1 (18.2)	62.2 (19.3)	66.6 (16.5)	67.8 (17.2)
Charlson Comorbidity Index score, n (%)										
None	207,525 (83.9)	15,321 (74.2)	36,737 (70.6)	120,748 (63.9)	9789 (62.1)	73,457 (59.7)	39,216 (62.7)	26,297 (59.9)	22,487 (55.5)	47,161 (51.0)
One	35,721 (14.4)	3989 (19.3)	12,287 (23.6)	49,863 (26.4)	4482 (28.4)	35,106 (28.5)	17,494 (28.0)	12,163 (27.7)	12,849 (31.7)	29,296 (31.7)
Two	3715 (1.5)	1035 (5.0)	2544 (4.9)	14,503 (7.7)	1222 (7.8)	11,454 (9.3)	4792 (7.7)	4169 (9.5)	4221 (10.4)	12,425 (13.4)
Three or more	545 (0.2)	293 (1.4)	436 (0.8)	3859 (2.0)	279 (1.8)	3080 (2.5)	1057 (1.7)	1244 (2.8)	993 (2.5)	3641 (3.9)
SCARF index, n (%)										
Fit	206,796 (83.6)	15,015 (72.8)	34,056 (65.5)	114,973 (60.8)	6197 (39.3)	65,911 (53.5)	33,014 (52.8)	23,871 (54.4)	17,473 (43.1)	42,879 (46.3)
Mild frailty	34,544 (14.0)	4052 (19.6)	13,608 (26.2)	52,629 (27.9)	5631 (35.7)	38,851 (31.6)	19,608 (31.3)	13,104 (29.9)	13,722 (33.8)	30,286 (32.7)
Moderate frailty	5041 (2.0)	1155 (5.6)	3385 (6.5)	16,175 (8.6)	2706 (17.2)	13,433 (10.9)	7360 (11.8)	4987 (11.4)	6511 (16.1)	13,493 (14.6)
Severe frailty	1125 (0.5)	416 (2.0)	955 (1.8)	5196 (2.8)	1238 (7.9)	4902 (4.0)	2577 (4.1)	1911 (4.4)	2844 (7.0)	5865 (6.3)

SD, standard deviation.

NES strategies was 6.4% (the equivalent figures are 5.7% vs. 9.7% for the cholelithiasis cohort, 9.6% vs. 11.8% for the diverticular disease cohort, 9.4% vs. 12.3% for the hernia cohort and 12.9% vs. 17.3% for the intestinal obstruction cohort).

The association between patients categorised as frail using the SCARF index and having ES did not follow the same pattern as comorbidity. In the appendicitis and cholelithiasis cohorts, the patients who received ES were more likely to be categorised as fit than patients who received NES strategies. However, in the diverticular disease cohort, and to a lesser extent in the intestinal obstruction cohort, frailty was more common in patients who received ES than in patients who received NES strategies. Finally, the distribution of frailty in the hernia cohort was similar between patients receiving ES and NES strategies.

Further analysis of the association between patient characteristics and receipt of ES has been reported elsewhere.[36] The most common ES procedures for cohort patients in receipt of ES are listed in *Table 6*.

Variation in emergency surgery rates between hospitals

Variation in rates of ES between the 175 hospitals contributing patients to the cohorts is shown in *Figure 1*. Greatest variation between hospitals was found in cholelithiasis, with a median rate of 18.4% [interquartile range (IQR) 11.4–28.1%; minimum, 2.3%; maximum, 66.4%]. A more moderate variation was found for hernia (median, 59.8%; IQR 54.2–65.8%; minimum, 30.8%; maximum, 79.2%) and appendicitis (median, 93.0%; IQR 91.0–94.5%; minimum, 67.5%; maximum, 98.6%). The least variation was observed with intestinal obstruction (median, 30.0%; IQR 27.0–33.2%; minimum, 20.4%; maximum, 51.4%) and diverticular disease (median, 11.2%; IQR 9.1–13.3%; minimum, 3.5%; maximum, 21.0%).

TABLE 6 Most common surgical procedures in the ES group

	Condition				
Procedure	**Appendicitis** (*n* = 247,506)	**Cholelithiasis** (*n* = 52,004)	**Diverticular disease** (*n* = 15,772)	**Hernia** (*n* = 62,559)	**Intestinal obstruction** (*n* = 40,550)
Median (IQR) days to surgery	1 (0–1)	2 (1–4)	1 (0–2)	0 (0–1)	1 (0–3)
Common main procedures (%)	Emergency excision of appendix (93.5) Appendicectomy and endoscopic resection of lesion of peritoneum (3.0) Emergency excision of appendix and drainage (1.9) Other (1.6)	Total cholecystectomy (85.3) Drainage of gall bladder (7.6) Partial cholecystectomy (3.4) Total cholecystectomy exploration bile duct (2.4) Other (1.5)	Resection with end colostomy – Hartmann's (44.1) Resection with other colostomy (e.g. loop colostomy) (25.7) Irrigation/drainage (colon, abdominal or pelvic area) (12.7) Resection and anastomosis (6.9) Colostomy with no resection on the same date (2.9) Other (7.7)	Repair of inguinal hernia (39.6) Repair of umbilical hernia (37.2) Repair of femoral hernia (19.8) Repair of ventral hernia (2.6) Other (0.1)	Freeing of adhesions of peritoneum and related procedures (48.9) Hemicolectomy (12.3) Colostomy or ileostomy (5.3) Ileectomy (10.1) Rectosigmoidectomy and related procedures (4.4) Hernia repair (3.9) Other (17.9)

IQR, interquartile range.
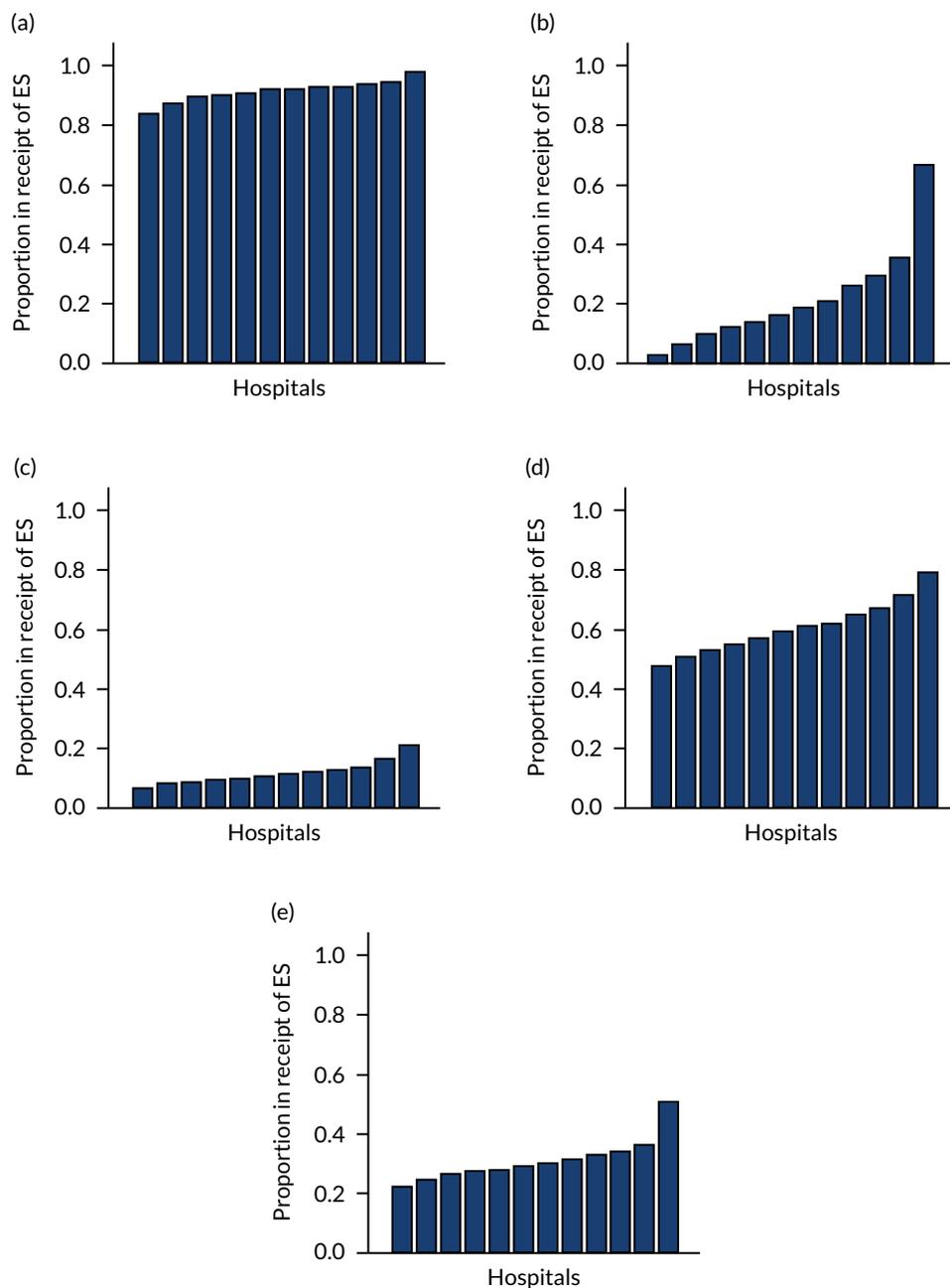
FIGURE 1 Variation in rates of ES in emergency admissions to 175 NHS acute general hospitals in England between April 2010 and December 2019. (a) Appendicitis; (b) cholelithiasis; (c) diverticular disease; (d) hernia; and (e) intestinal obstruction.

Further analysis examining case mix-adjusted rates in ES at the level of the NHS trust has been reported elsewhere.[36]

Discussion

This chapter describes the criteria used to define cohorts for the five acute conditions and the methods used to determine which patients received ES and which patients received NES strategies. The characteristics of each cohort are described and the variation in rates of ES between hospitals is reported. The process drew on the expertise of a clinical panel to identify emergency general surgical populations and to define ES following 'target trial' principles, using routine observational data.[31,32]

This research extends previous studies[28,29,37] that have found lower rates of ES for specific subgroups of older patients presenting with acute conditions within the UK. Reports from the RCS of England[28,29] have generally found lower levels of ES for patients aged over ≥ 75 years than for patients aged 65–74 years, notably for patients with acute cholelithiasis. The reports[28,29] discourage ES rationing by biological age and have called for further research on ES and age to also consider comorbidities and frailty.[28,29] The results from the ESORT study show that ES generally decreases with age, a finding that remains after adjusting for a wider range of case mix measures, including frailty.[35] For patients with acute cholelithiasis and appendicitis, the age gradient is especially steep and goes across the age distribution.

For patients presenting with hernia, previous studies reported higher ES rates for patients aged > 75, than for patients aged 65–74 years.[28,29] Findings from the ESORT study indicated that this was not the case for patients in the oldest age groups (i.e. aged ≥ 85 years), a finding that remained after allowing for differences in other patient characteristics.[36]

Large variation in the rates of ES was found for patients presenting with acute cholelithiasis, which suggests that, in some hospitals, National Institute for Health and Care Excellence (NICE) guidelines,[38] which recommend laparoscopic cholecystectomy within 7 days of diagnosis, are not being followed. These guidelines are informed by evidence from a meta-analysis that reported improved outcomes for ES compared with delayed cholecystectomy for patients with biliary colic, acute cholecystitis or gallstone pancreatitis.[39] Despite these recommendations, related research[37] has also reported large levels of unexplained variation across NHS trusts in ES over a 2-month time period.

Previous research[37] on trust-level variation for patients with benign gallbladder diseases reported higher rates of ES in centres with a specialist hepatobiliary centre available, which may reflect better availability of operating theatre space, clearer understanding of the evidence comparing emergency and delayed cholecystectomy, or the enthusiasm to deliver an emergency cholecystectomy service. This previous study[37] also found that other trust-level factors, such as the availability of ES operating lists specific to the condition and the number of consultants with expertise in the specific forms of ES, were not associated with ES rates for patients with benign gall bladder diseases.[37] Surgeon-led quality improvement initiatives, such as the Cholecystectomy Quality Improvement Collaborative, may also explain higher uptake of ES in some hospitals.[40]

More moderate variation between hospitals for acute appendicitis and abdominal wall hernia may reflect (1) the lack of evidence about which patients benefit from ES versus NES for these conditions, (2) that there are less well-defined care pathways and (3) a lack of clinical guidelines in the UK to inform the choice of whether or not the patient has ES.[41,42] It is also important to recognise that, over the ESORT study time period, patients with an abdominal wall hernia presenting as emergency admissions were not managed by a distinct surgical subspecialty, which may have hindered attempts to standardise practice.[42] In addition, different local policies on restricting elective hernia surgery may have affected emergency provision.[43,44] For patients with uncomplicated acute appendicitis, the emerging evidence for antibiotics as an alternative to ES may explain variability.[7,13]

Lower variation in rates of ES for patients with intestinal obstruction and diverticular disease may reflect increased standardisation in the clinical management of the conditions over the ESORT study's time period, the emergence of evidence from RCTs[14,45,46] and the development of clinical pathways. For diverticular disease, there is consensus in the UK about the surgical specialty (i.e. colorectal surgery) that manages patients. For patients with acute diverticular disease, ES has declined over time,[47] and the low ES rate reflects current NICE recommendations that encourage NES strategies and the lack of high-quality evidence on the effectiveness of ES for patients with acute diverticular disease.[48]

The variation that was identified in rates of ES between hospitals provides the basis for the IV (i.e. a hospital's TTO) that will be used in evaluating the clinical effectiveness (see *Chapter 3*) and cost-effectiveness (see *Chapter 4*) of ES in comparison with alternative NES treatment strategies.

Chapter 3 Clinical effectiveness

Introduction

For common acute conditions that present as emergency admissions, an important clinical decision is which patients should receive ES and which patients should receive NES strategies. Here, NES strategies include medical management, non-surgical procedures (e.g. radiological-guided drainage of abscess) and surgery deferred to the elective (planned) setting. ES rates for patients with acute conditions have declined over the last 20 years.[6] Protocols for NES strategies have been implemented as part of RCTs, and for some conditions, such as acute appendicitis, there is some evidence of improved outcomes when compared with ES strategies.[7,13] For other conditions, such as acute cholelithiasis, RCTs have found that NES strategies may have unintended consequences, with patients having recurrent symptoms and delayed surgery, therefore, leading to further pressure on surgical waiting lists.[9] However, to the best of our knowledge, none of these RCTs have included routine emergency admissions to hospitals. Similarly, for some acute conditions, no RCTs, to the best of our knowledge, have been conducted to compare ES with NES strategies.[7,9,13–15,49]

This lack of evidence has resulted in clinical uncertainty about the benefits and harms of ES for patients with acute conditions. Consequently, there is wide variation in clinical practice[2] and in rates of ES across NHS trusts in England for emergency admissions with common acute conditions, with patients of similar prognosis more likely to receive ES in some hospitals, and NES strategies in others.[36]

The aim of this chapter is to assess the relative clinical effectiveness of ES compared with NES strategies for the following five common acute conditions: (1) acute appendicitis, (2) cholelithiasis, (3) diverticular disease, (4) abdominal wall hernia and (5) intestinal obstruction. The analysis exploits the variation in ES rates across NHS hospitals in England (see *Chapter 2*), and reports relative clinical effectiveness overall and according to prespecified subgroups, in particular age, sex, number of comorbidities and level of frailty. The analyses use data on cohorts of emergency admissions to 175 NHS acute hospitals for the five common acute conditions (see *Chapter 2*). The study protocol and statistical analysis plan were developed following the principles of emulating a target trial (see *Report Supplementary Material 4*).[30]

The design and proposed analysis of the ESORT study were informed by a Patient and Public Advisory Group during two online workshops held in July 2020.[30] The workshop participants were asked to consider outcome measures for patients following emergency admission to hospital for acute conditions. The group agreed that an appropriate measure would capture mortality and the number of days in hospital.

Methods

Outcomes
The primary outcome measure was the number of days alive and out of hospital (DAOH) at 90 days. The number of DAOH is a composite measure, which encompasses mortality and total length of hospital stay, including re-admissions, for example for reinterventions. The number of DAOH has been recommended both as a standardised patient-centred outcome measure and as a core outcome measure for clinical effectiveness trials in perioperative medicine. This outcome measure has been formally validated in multiple studies following ES,[50–52] and was supported by a panel of ex-patients and public contributors.[30] The number of DAOH was measured from the date the index episode started for up to 90 days. The calculation of the number of DAOH used HES data on the total duration of

hospitalisation over the 90-day period, and the date of death from linkage to the Office for National Statistics (ONS) death record. Patients who died within the 90-day period were assigned zero DAOH. The study's sample size for each condition was projected to be sufficient to assess overall differences between the comparison groups in the mean number of DAOH of at least 1 day, with 80% power and 95% levels of statistical significance[51] (see the statistical analysis plan on the project webpage[30]). The secondary outcomes were 90-day mortality and aggregate length of stay (LOS) (i.e. the two components of DAOH), as well as any emergency re-admission within 30 days.

Instrumental variable: the tendency to operate

An IV analysis aims to approximate the random assignment of treatment in a RCT, by using an instrument to balance observed and unobserved baseline prognostic measures between the comparison groups.[18] The analysis adopted an IV that had been previously used to evaluate ES from claims data from the USA,[20] which, in turn, followed from pharmacoepidemiological research that used clinician preference as an instrument for treatment receipt.[19] In the ESORT study, the IV was the hospital's TTO, which reflects practice variation across hospitals in ES rates for these five conditions. For each eligible emergency admission, the TTO was defined as the proportion of eligible emergency admissions in the previous year at that specific hospital who received ES. Therefore, a hospital's past preference for ES is regarded as strongly predictive of the treatment choice for the current patient.

The rationale for the IV design is that, after adjustment for observed characteristics, such as age and comorbidity, patients' baseline prognosis is similar across hospitals with different levels of TTO. Hence, the patients can be 'randomised' between the ES and NES strategies according to the hospital's TTO. A valid instrument must meet two conditions.[18] First, the instrument must be associated with the treatment received, with guidance on IV methods requiring that the accompanying F-statistic exceeds a value of 10.[53] Second, the instrument should have no relation with the outcome, except through the treatment. There is no empirical approach to assess whether or not an instrument is directly associated with an outcome, but examining the extent to which observed characteristics are balanced across different levels of the instrument increases confidence that unobserved confounders are also balanced.[54] Further information on IV methods and the TTO are included in *Appendix 2*.

Patient-level covariates

The following baseline patient characteristics were extracted from HES data and were considered to potentially influence the decision as to whether the patient had ES or NES: age category, sex, ethnicity, Index of Multiple Deprivation decile, diagnostic subcategories from ICD-10 codes, Charlson Comorbidity Index score[34] and frailty measured using the SCARF index.[35] The SCARF index is based on the accumulation of deficits across a number of domains. The SCARF index uses ICD-10 codes to define 32 deficits that cover functional impairment, geriatric syndromes, problems with nutrition, cognition and mood, and medical comorbidities, with severe frailty defined as the presence of six or more deficits. Procedures during the ES window for each condition and up to 30 days were identified from OPCS-4 procedure codes in HES and categorised as panel-defined ES procedures, other operative procedures, abdominal interventional radiology procedures and imaging procedures.

Statistical analysis

The study reported absolute risk differences (for binary measures) and difference in means (for continuous measures). The IV analysis estimated the relative effectiveness of ES compared with NES for each individual, and fully accounted for heterogeneity of effects as well, as confounding (see *Appendix 2*).[55–58] The person-level treatment effects were aggregated to report the effects of ES overall and for each prespecified subgroup of interest (i.e. age category, sex, diagnostic subcategories, Charlson Comorbidity Index score, SCARF index and year of admission).

Probit regression models were used to estimate the initial propensity score (first stage), and the outcomes models used generalised linear models (GLMs) and accommodated whether each end point was continuous or binary (second stage). Models at both stages adjusted for the measures described

above, together with the time period and proxies for hospital quality. These proxies for hospital quality were defined by the rates of emergency admission and mortality for each hospital and acute condition in 2009–10 (i.e. time constant) and in the year prior to the specific admission concerned (i.e. time varying). The estimates were reported with confidence intervals (CIs) using bootstrapped (300 replications) standard errors that allowed for the clustering of individuals within hospitals. Further details of the statistical analysis methods are included in *Appendix 2*.

Sensitivity analyses

Six different sensitivity analyses were undertaken to assess whether or not the results from the main analysis were robust to alternative definitions and assumptions. First, stricter criteria for ES were applied by increasing the level of clinical panel support required for defining (1) a procedure as ES (rather than NES) and (2) the maximum time window within which ES can occur, from at least 50% to at least 75% support from the panel. Second, the ES time windows were reduced by taking the threshold as the upper quartile value of the distribution of the 'time to ES' from the main (base-case) analysis. Third, the study adjusted for 'quality of care', using external hospital performance measures from the National Emergency Laparotomy Audit (NELA).[59–61] Fourth, the study excluded observations from hospitals that provided a relatively low volume of ES procedures (i.e. at least one IQR below the median) for each condition. Fifth, the study considered a different measure of the number of DAOH, which 'counts' days out of hospital for those patients who die before 90 days.[62] Sixth, the estimates from the IV approach were compared with those from a regression approach that used logistic regression (binary outcomes) and ordinary least squares regression (continuous outcomes). The final regression approach was adjusted for observed baseline patient characteristics, but assumes no unobserved confounding after adjustment. Further details of the analysis methods and sensitivity analyses are in *Appendix 2*.

Results

Patient characteristics and clinical management

The number of patients presenting as emergency admissions who met the inclusion criteria were as follows: 268,144 admissions for appendicitis, 240,977 admissions for cholelithiasis, 138,869 admissions for diverticular disease, 106,432 admissions for a hernia and 133,073 admissions for an intestinal obstruction (see *Table 5*). The percentage of patients who had ES were 92.3% for acute appendicitis, 21.6% for cholelithiasis, 11.4% for diverticular disease, 58.8% for an abdominal wall hernia and 30.5% for an intestinal obstruction. The case mix of patients differed between patients who received ES and patients who received NES strategies, according, for example, to mean age, the proportion of patients who had comorbidities and levels of moderate or severe frailty (see *Table 5*).

All patients in the ES groups had a panel-defined ES procedure within the relevant ES time window (*Table 7*). The percentage of patients in the NES groups who had an ES procedure after the time window, but before 30 days, ranged from 0.6% (for diverticular disease) to 16.5% (for hernia). The percentages of patients who had abdominal interventional radiology procedures were low and were similar across ES and NES groups. For patients with acute appendicitis, the percentage of patients who had imaging or diagnostic procedures was higher for the NES group then for ES group.

Validity of the instrumental variable

The IV was judged to meet each of the requisite criteria for validity. First, the hospital's TTO was strongly correlated with ES receipt, with *F*-statistics that ranged from 450 (diverticular disease) to 24,517 (cholelithiasis), compared with the requirement that they exceeded 10. Second, the hospital's TTO was successful in balancing each of the observed baseline covariates (see *Appendix 2, Figure 29*).

TABLE 7 Clinical management for the ES and NES groups in and after (up to 30 days) the time window for ES

Clinical management	Appendicitis				Cholelithiasis				Diverticular disease				Hernia				Intestinal obstruction			
	ES (n = 247,506)		NES strategies (n = 20,638)		ES (n = 52,004)		NES strategies (n = 188,973)		ES (n = 15,772)		NES strategies (n = 123,097)		ES (n = 62,559)		NES strategies (n = 43,873)		ES (n = 40,550)		NES strategies (n = 92,523)	
	In time window	After time window	In time window	After time window	In time window	After time window	In time window	After time window	In time window	After time window	In time window	After time window	In time window	After time window	In time window	After time window	In time window	After time window	In time window	After time window
Operative procedures (%)																				
ES procedure	100.0	0.8	0.0	3.4	100.0	0.7	0.0	9.5	100.0	0.5	0.0	0.6	100.0	0.7	0.0	16.5	100.0	2.2	0.0	5.6
Other procedure[a]	0.0	0.1	15.5	0.4	0.0	2.3	5.1	4.1	0.0	0.1	0.1	0.2	0.0	0.2	12.9	0.1	0.0	0.1	7.4	0.1
No surgery	0.0	99.9	84.5	96.2	0.0	97.0	94.9	86.3	0.0	99.4	99.9	99.2	0.0	99.1	87.1	83.4	0.0	97.7	92.6	94.3
Interventional radiology (%)																				
Yes	0.2	0.4	0.9	0.5	0.4	0.8	0.2	0.5	6.1	0.5	1.5	0.2	0.4	0.7	0.2	0.5	0.9	1.7	0.7	0.6
No	99.8	99.6	99.1	99.5	99.6	99.2	99.8	99.5	93.9	99.5	98.5	99.8	99.6	99.3	99.8	99.5	99.1	98.3	99.3	99.4
Imaging (%)																				
Yes	27.8	4.0	56.6	8.3	32.2	8.5	37.0	11.0	78.0	4.7	80.6	8.0	16.1	2.3	16.4	1.6	70.1	7.0	54.5	2.4
No	72.2	96.0	43.4	91.7	67.8	91.5	63.0	89.0	22.0	95.3	19.4	92.0	83.9	97.7	83.6	98.4	29.9	93.0	45.5	97.6

a If there is no ES procedure.

Unadjusted outcomes

The crude outcomes for the ES and NES groups are presented in *Table 8*, together with the absolute mean differences between the ES and NES treatment strategy groups, without adjusting for case mix differences. For patients with diverticular disease, the mean number of DAOH at 90 days was 19.0 days lower in the ES group when compared with the NES strategy group. For the other four conditions, the unadjusted differences in the mean number of DAOH were relatively small. For patients with diverticular disease, a higher proportion of patients in the ES group died before 90 days, with the majority of these additional deaths occurring within the first 30 days.

Overall effectiveness of emergency surgery compared with non-emergency surgery

Table 9 presents the results from the IV analysis, which adjusts for confounding. There were small overall mean differences in number of DAOH at 90 days between the ES and NES strategy groups {appendicitis: −0.73 [95% confidence interval (CI) −2.10 to 0.64] days; cholelithiasis: 0.60 [95% CI −0.10 to 1.30] days;

TABLE 8 Unadjusted outcomes following ES and NES strategies for the primary (i.e. DAOH) and secondary outcomes

Outcome	Group		Absolute mean difference (95% CI)	
	ES	NES		
Appendicitis				
DAOH within 90 days (mean)	84.78	82.50	2.28	2.01 to 2.54
Mortality within 90 days (%)	0.19	1.09	−0.90	−1.07 to −0.73
LOS within 90 days (mean)	5.09	6.68	−1.60	−1.77 to −1.42
Emergency re-admissions within 30 days (%)	9.06	11.09	−2.03	−2.55 to −1.51
Cholelithiasis				
DAOH within 90 days (mean)	81.28	80.74	0.54	0.22 to 0.87
Mortality within 90 days (%)	0.73	1.50	−0.77	−0.9 to −0.64
LOS within 90 days (mean)	8.23	8.22	0.01	−0.26 to 0.28
Emergency re-admissions within 30 days (%)	10.63	14.43	−3.80	−4.23 to −3.37
Diverticular disease				
DAOH within 90 days (mean)	60.92	79.94	−19.0	−19.6 to −18.4
Mortality within 90 days (%)	9.30	3.04	6.27	5.74 to 6.79
LOS within 90 days (mean)	22.38	7.84	14.60	14.2 to 14.9
Emergency re-admissions within 30 days (%)	8.84	9.72	−0.88	−1.42 to 0.35
Hernia				
DAOH within 90 days (mean)	80.98	81.63	−0.65	−0.95 to −0.35
Mortality within 90 days (%)	2.68	3.69	−1.01	−1.25 to −0.78
LOS within 90 days (mean)	7.08	5.49	1.59	1.42 to 1.76
Emergency re-admissions within 30 days (%)	9.43	12.40	−2.97	−3.38 to −2.56
Intestinal obstruction				
DAOH within 90 days (mean)	66.56	68.01	−1.46	−2.00 to −0.91
Mortality within 90 days (%)	7.59	13.37	−5.78	−6.33 to −5.22
LOS within 90 days (mean)	18.13	11.80	6.33	6.07 to 6.58
Emergency re-admissions within 30 days (%)	9.50	14.72	−5.22	−5.61 to −4.82

TABLE 9 Effects of ES vs. NES strategies on DAOH, all-cause mortality, LOS and emergency re-admissions from IV analysis

End point	Mean difference	95% CI	p-value
Appendicitis			
DAOH within 90 days (mean)	–0.73	–2.10 to 0.64	0.30
Mortality within 90 days (%)	0.24	–0.04 to 0.51	0.09
LOS within 90 days (mean)	0.03	–1.04 to 1.10	0.96
Emergency re-admissions within 30 days (%)	–2.50	–10.3 to 5.26	0.53
Cholelithiasis			
DAOH within 90 days (mean)	0.60	–0.10 to 1.30	0.09
Mortality within 90 days (%)	0.18	–0.97 to 1.32	0.76
LOS within 90 days (mean)	–0.43	–0.88 to 0.03	0.07
Emergency re-admissions within 30 days (%)	–3.85	–5.54 to –2.16	< 0.001
Diverticular disease			
DAOH within 90 days (mean)	–2.66	–15.7 to 10.4	0.69
Mortality within 90 days (%)	3.34	–5.22 to 11.9	0.45
LOS within 90 days (mean)	2.28	–5.23 to 9.80	0.55
Emergency re-admissions within 30 days (%)	–12.6	–21.4 to –3.77	0.005
Hernia			
DAOH within 90 days (mean)	–0.07	–2.40 to 2.25	0.95
Mortality within 90 days (%)	–4.99	–9.92 to –0.07	0.047
LOS within 90 days (mean)	2.35	1.54 to 3.15	< 0.001
Emergency re-admissions within 30 days (%)	–4.05	–7.77 to –0.33	0.033
Intestinal obstruction			
DAOH within 90 days (mean)	3.32	–3.13 to 9.76	0.31
Mortality within 90 days (%)	1.73	–4.93 to 8.38	0.61
LOS within 90 days (mean)	–4.25	–7.50 to –1.00	0.01
Emergency re-admissions within 30 days (%)	–8.62	–18.1 to 0.88	0.075

diverticular disease: –2.66 [95% CI –15.7 to 10.4] days; hernia: –0.07 [95% CI –2.40 to 2.25] days; intestinal obstruction: 3.32 [95% CI –3.13 to 9.76] days}.

The IV analysis also found that, compared with NES strategies, for four of the conditions, the effect of ES on 90-day mortality was small and not statistically significant. For patients with an abdominal wall hernia, the absolute risk of death with ES strategies was lower by 4.99 (95% CI –9.92 to –0.07) percentage points compared with NES strategies. Statistically significant differences in mean LOS were found for the abdominal wall hernia cohort (i.e. LOS was longer by 2.35 days with ES) and the intestinal obstruction cohort (i.e. LOS was shorter by 4.25 days with ES). Compared with the NES strategies, the ES strategies led to reductions in the proportion of emergency re-admissions before 30 days for all conditions apart from acute appendicitis.

Subgroup results

Subgroup results for the relative effectiveness of ES compared with NES treatment strategies in the five conditions are reported for the number of DAOH and its two components (i.e. 90-day mortality and LOS) in *Figures 2–16*. Outcomes are reported according to age, sex, level of frailty, number of comorbidities, subdiagnosis and year of admission.

The relative effectiveness of ES compared with NES strategies was modified by age group, with ES less effective for some subgroups of older patients. For patients aged 80–84 years and ≥ 85 years, the difference in the mean number of DAOH following ES and NES strategies were, respectively, as follows: –11.81 (95% CI –16.50 to –7.12) days and –0.58 (95% CI –9.12 to 7.95) days for appendicitis (*Figure 2*), 0.76 (95% CI –2.04 to 3.57) days and –4.31 (95% CI –9.76 to 1.14) days for cholelithiasis (see *Figure 5*),–12.49 (95% CI –35.21 to 10.23) days and –23.98 (95% CI –53.65 to 5.68) days diverticular disease (see *Figure 8*), –3.34 (95% CI –7.89 to 1.21) days and –4.78 (95% CI –9.88 to 0.31) days for a hernia (see *Figure 11*) and–24.7 (95% CI –39.82 to –9.58) days and –8.64 (95% CI –19.35 to 2.07) days
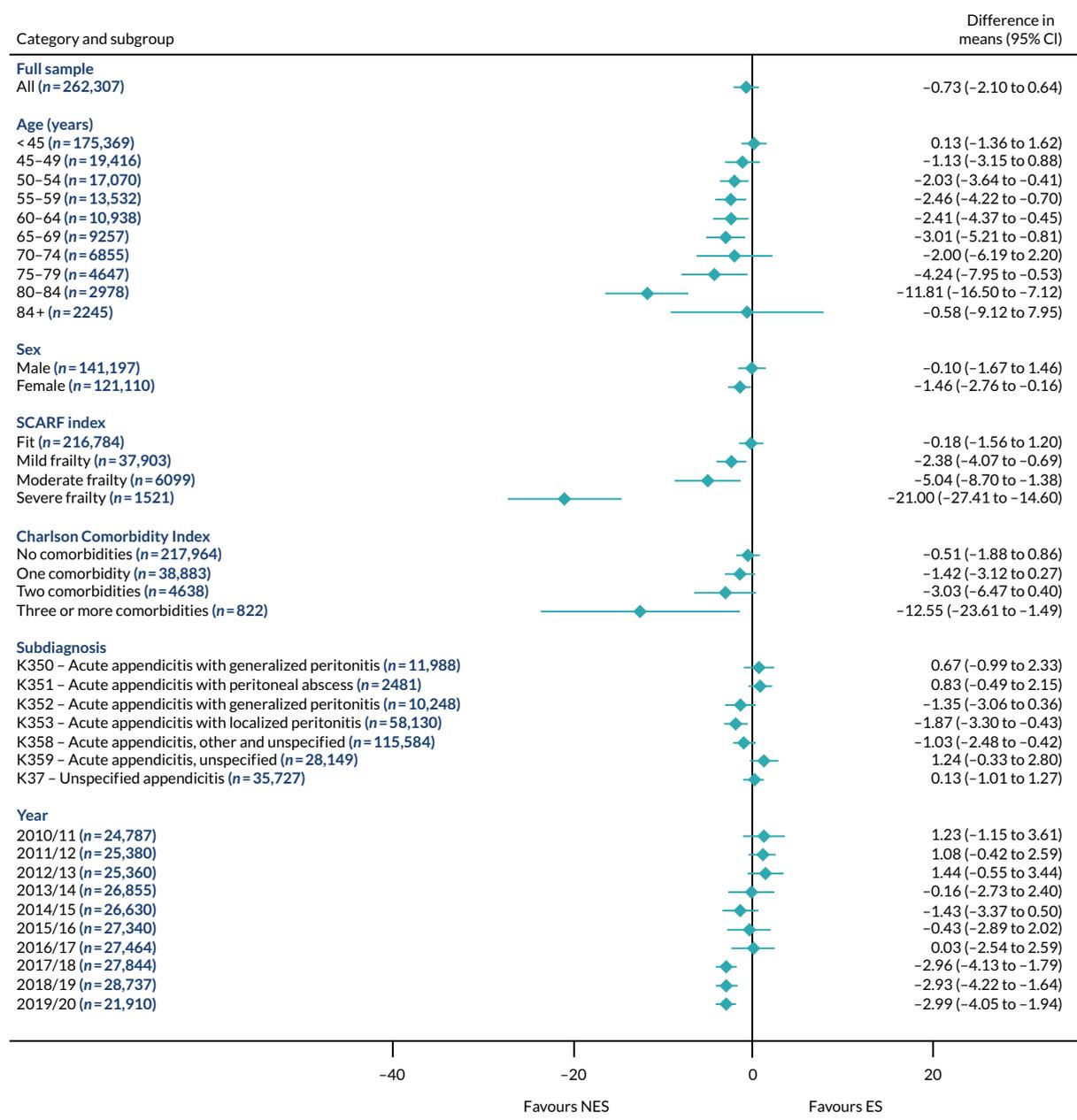
FIGURE 2 Mean differences in number of DAOH between ES and NES treatment strategies for appendicitis subgroups.

Category and subgroup	Mean absolute difference, % (95% CI)
Full sample	
All (*n* = 262,301)	0.00 (−0.00 to 0.01)
Age (years)	
< 45 (*n* = 175,364)	−0.00 (−0.00 to 0.00)
45–49 (*n* = 19,416)	−0.00 (−0.02 to 0.01)
50–54 (*n* = 17,070)	0.00 (−0.00 to 0.01)
55–59 (*n* = 13,532)	0.00 (−0.00 to 0.01)
60–64 (*n* = 10,938)	0.00 (−0.00 to 0.01)
65–69 (*n* = 9257)	0.01 (−0.00 to 0.01)
70–74 (*n* = 6855)	0.02 (0.00 to 0.03)
75–79 (*n* = 4646)	0.03 (0.00 to 0.05)
80–84 (*n* = 2978)	0.10 (0.05 to 0.16)
84+ (*n* = 2245)	0.04 (−0.04 to 0.12)
Sex	
Male (*n* = 141,197)	0.00 (−0.00 to 0.01)
Female (*n* = 121,104)	0.00 (−0.00 to 0.01)
SCARF index	
Fit (*n* = 216,779)	−0.00 (−0.00 to 0.00)
Mild frailty (*n* = 37,902)	0.01 (0.00 to 0.01)
Moderate frailty (*n* = 6099)	0.04 (0.01 to 0.07)
Severe frailty (*n* = 1521)	0.17 (0.08 to 0.27)
Charlson Comorbidity Index	
No comorbidities (*n* = 217,958)	0.00 (−0.00 to 0.00)
One comorbidity (*n* = 38,883)	0.01 (0.00 to 0.01)
Two comorbidities (*n* = 4638)	0.04 (0.01 to 0.07)
Three or more comorbidities (*n* = 822)	0.21 (0.09 to 0.34)
Subdiagnosis	
K350 – Acute appendicitis with generalized peritonitis (*n* = 11,986)	0.00 (−0.00 to 0.01)
K351 – Acute appendicitis with peritoneal abscess (*n* = 2480)	0.00 (−0.00 to 0.01)
K352 – Acute appendicitis with generalized peritonitis (*n* = 10,249)	0.01 (0.00 to 0.01)
K353 – Acute appendicitis with localized peritonitis (*n* = 58,125)	0.00 (0.00 to 0.01)
K358 – Acute appendicitis, other and unspecified (*n* = 115,584)	0.00 (−0.00 to 0.00)
K359 – Acute appendicitis, unspecified (*n* = 28,149)	0.00 (−0.00 to 0.00)
K37 – Unspecified appendicitis (*n* = 35,728)	0.00 (−0.00 to 0.00)
Year	
2010/11 (*n* = 24,786)	0.00 (−0.00 to 0.01)
2011/12 (*n* = 25,380)	0.00 (−0.01 to 0.01)
2012/13 (*n* = 25,360)	0.00 (−0.00 to 0.01)
2013/14 (*n* = 26,850)	0.00 (−0.00 to 0.01)
2014/15 (*n* = 26,634)	0.00 (0.00 to 0.01)
2015/16 (*n* = 27,336)	0.00 (−0.00 to 0.01)
2016/17 (*n* = 27,464)	0.00 (−0.01 to 0.01)
2017/18 (*n* = 27,844)	0.00 (−0.00 to 0.01)
2018/19 (*n* = 28,737)	0.00 (−0.00 to 0.01)
2019/20 (*n* = 21,910)	0.00 (−0.00 to 0.01)

Favours ES — Favours NES

FIGURE 3 Mean differences in 90-day mortality between ES and NES treatment strategies for appendicitis subgroups.

for an intestinal obstruction (see *Figure 14*). Conversely, ES was more effective than NES strategies in some younger age groups for cholelithiasis, hernia and intestinal obstruction. In four conditions, the comparatively lower number of DAOH with ES strategies in those aged ≥ 85 years was, in part, due to higher 90-day mortality. The exception was for the hernia cohort, for which the lower 90-day mortality with ES (see *Figure 12*) was offset by a longer LOS (see *Figure 13*).

There was some evidence that the mean number of DAOH was smaller by −1.46 (95% CI −2.76 to −0.16) days with ES, compared with NES strategies, for females with appendicitis (see *Figure 2*), whereas the mean number of DAOH was larger by 0.77 (95% CI 0.11 to 1.43) days with ES, compared with NES strategies, in females with cholelithiasis (see *Figure 5*) and larger by 9.69 (95% CI 4.19 to 15.20) days with ES, compared with NES strategies, in females with intestinal obstruction (see *Figure 14*). Similar differences were not observed in men with these conditions.
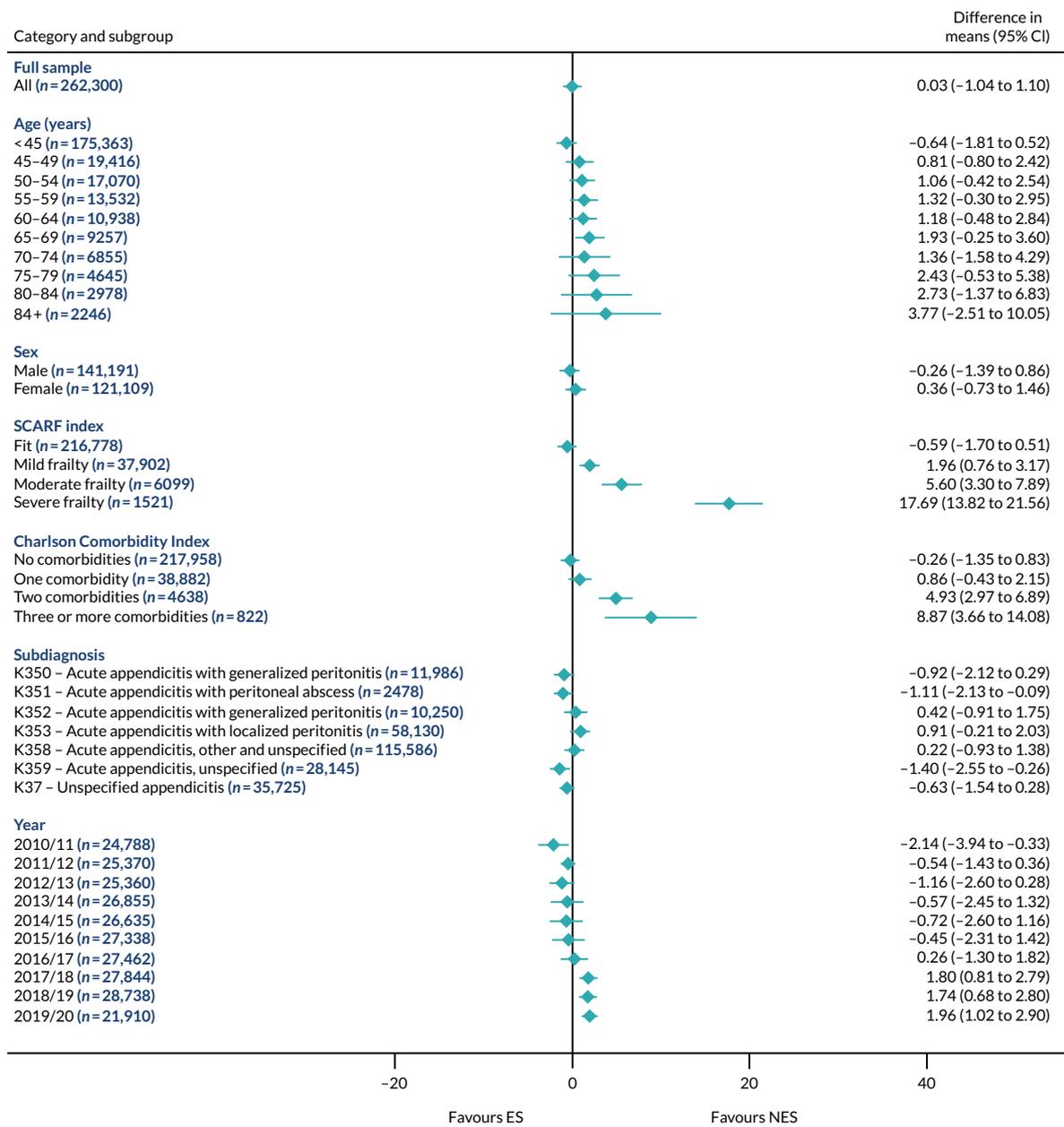
FIGURE 4 Mean differences in days in hospital in the first 90 days between ES and NES treatment strategies for appendicitis subgroups.

The largest differences in number of DAOH were observed in patients categorised as being severely frail according to the SCARF index. In all five conditions, the number of DAOH at 90 days was smaller following ES, compared with NES strategies, with mean differences of −21.0 (95% CI −27.4 to −14.6) days for appendicitis,−5.72 (95% CI −11.3 to −0.2) days for cholelithiasis,−38.9 (95% CI −63.3 to −14.6) days for diverticular disease, −19.5 (95% CI −26.6 to −12.3) days for a hernia and−34.5 (95% CI −46.7 to −22.4) days for an intestinal obstruction. For four conditions, the smaller number of DAOH was attributable to both an increase in 90-day mortality following ES and longer LOS, compared with NES strategies. The exception was with hernia cohort, where there was increased LOS but no difference in mortality.

A smaller number of DAOH with ES treatment strategies was also found for patients categorised as having 'moderate frailty' in appendicitis, hernia and intestinal obstruction cohorts, and also in patients with 'mild frailty' in the appendicitis cohort. For patients who were categorised as 'fit', the mean
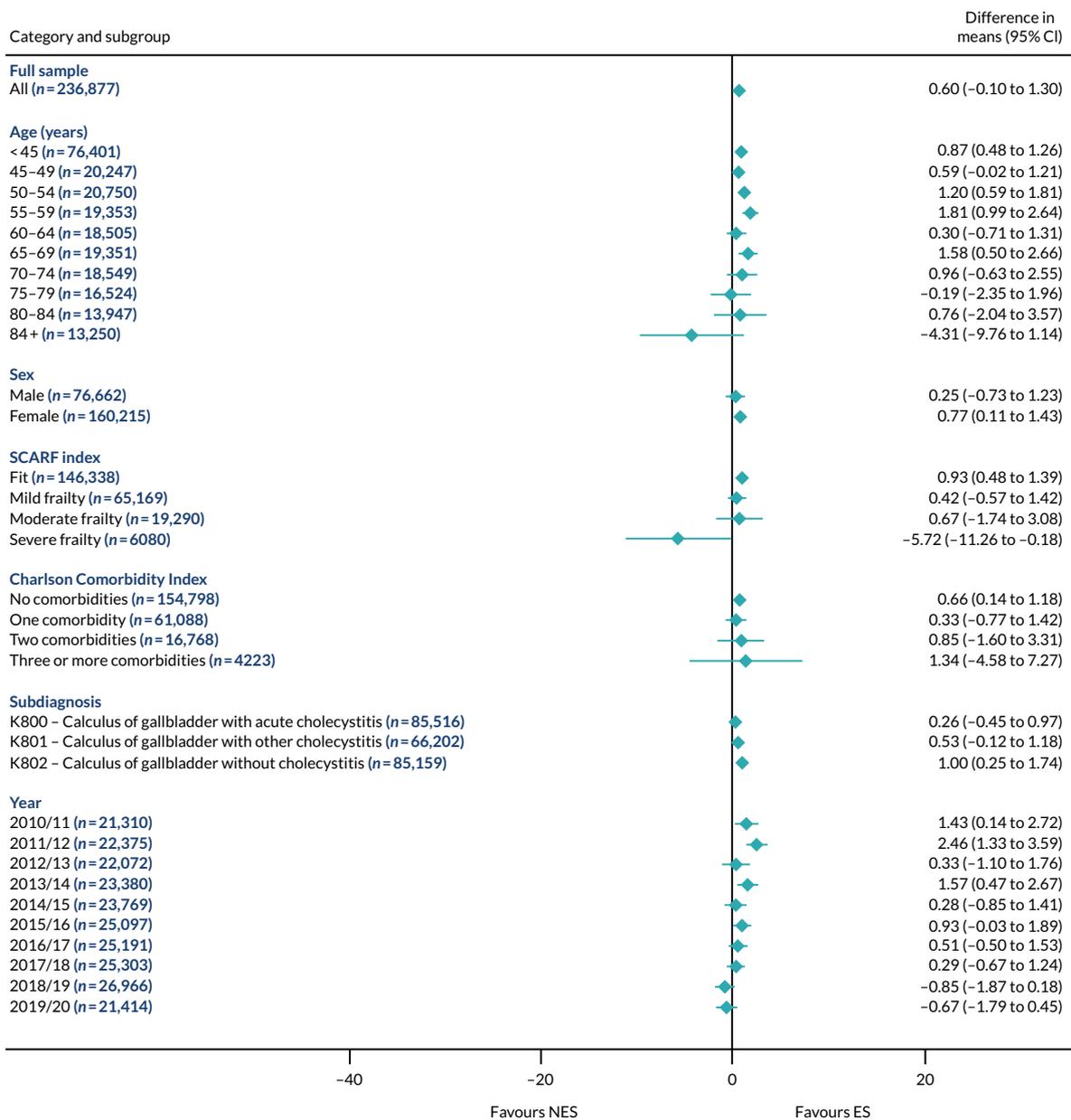
FIGURE 5 Mean differences in number of DAOH between ES and NES treatment strategies for cholelithiasis subgroups.

difference in number of DAOH tended to favour ES for diverticular disease (5.35 days, 95% CI –2.56 to 13.28 days), hernia (2.26 days, 95% CI 0.37 to 4.15 days) and intestinal obstruction (18.2 days, 95% CI 14.8 to 22.47 days), and was similar or smaller for appendicitis (–0.18 days, 95% CI –1.56 to 1.20 days) and cholelithiasis (0.93 days, 95% CI 0.48 to 1.39 days).

The relationship between comorbidity and differences in number of DAOH following ES or NES strategies did not follow the same pattern as observed for frailty. For patients with three or more comorbidities, ES strategies led to a smaller number of DAOH at 90 days for patients with acute appendicitis (mean difference –12.55 days, 95% CI –23.61 to –1.49 days), a larger number of DAOH at 90 days for patients with an intestinal obstruction (mean difference 26.37 days, 95% CI 8.71 to 44.02 days) and similar average number of DAOH at 90 days between the comparison groups for the other three conditions.
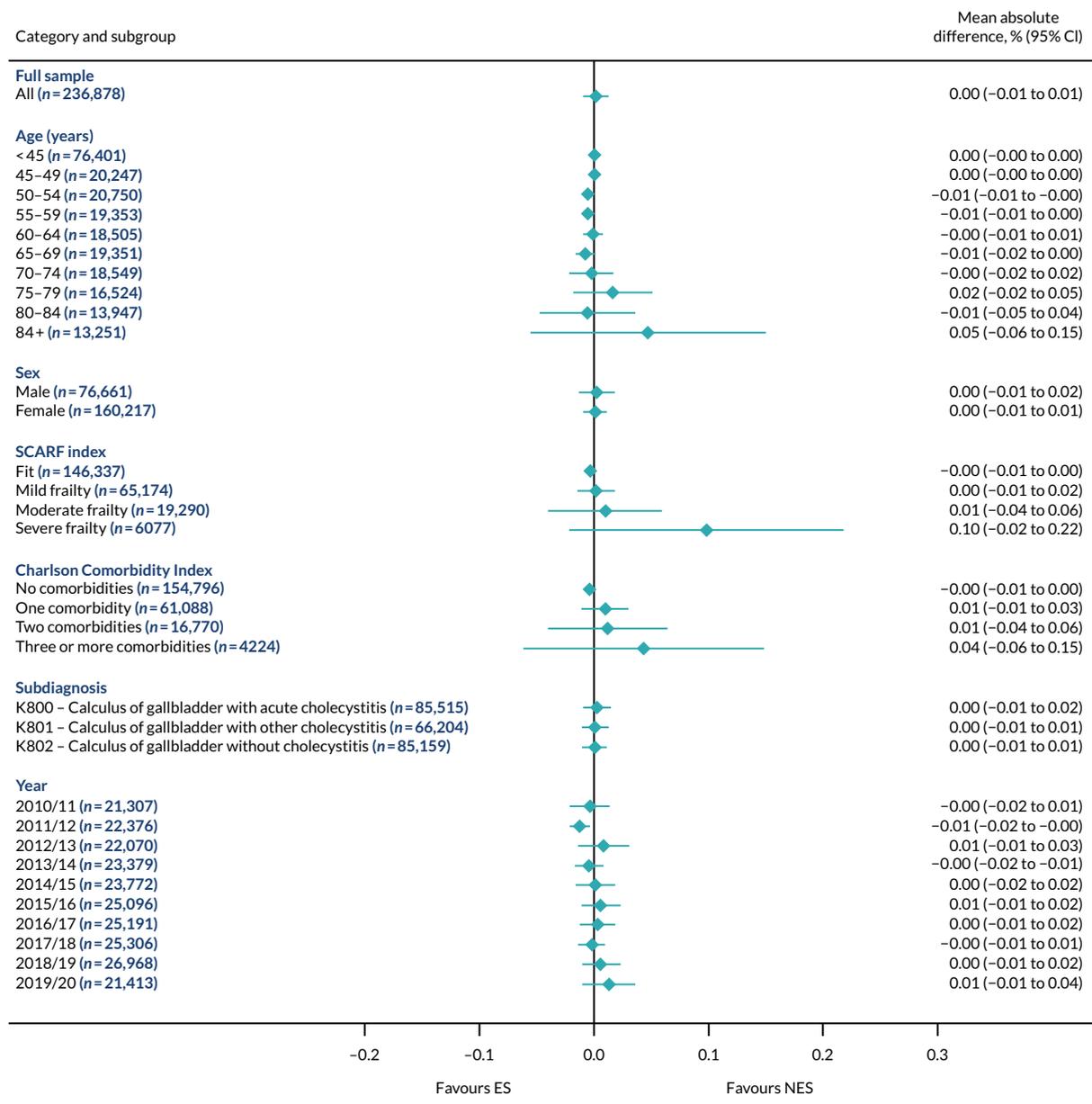
FIGURE 6 Mean differences in 90-day mortality between ES and NES treatment strategies for cholelithiasis subgroups.

Diagnostic subgroups of patients for whom ES strategies indicated smaller numbers of DAOH were patients with diverticular disease with perforation and abscess (–11.55 days, 95% CI –19.03 to –4.06 days), and patients with an intestinal obstruction with a diagnosis of gallstone ileus (–14.15 days, 95% CI –19.08 to –9.22 days). Conversely, patients with a colorectal cancer diagnosis were found to have a larger number of DAOH following ES (see *Figure 14*). For four out of the five conditions, for each year of admission, the estimates of the effectiveness of ES compared with NES were consistent with the overall estimates. The exception was for appendicitis (see *Figure 2*), for which the mean number of DAOH was larger for NES strategies for the last 3 years of admission (see *Discussion*).

Sensitivity analyses

The overall effects of the ES compared with NES strategies were similar when alternative standpoints were taken to the analysis (*Figure 17*). Specifically, when alternative, more stringent criteria for the definitions of ES were used, the mean differences in the number of DAOH remained below 4 days for each of the conditions. The overall results were also robust to the use of alternative definitions of DAOH and of the 'quality' of acute care, and the exclusion of observations from hospitals with lower

Category and subgroup		Difference in means (95% CI)
Full sample		
All (*n* = 236,880)		−0.43 (−0.88 to 0.03)
Age (years)		
< 45 (*n* = 76,401)		−0.80 (−1.11 to −0.48)
45–49 (*n* = 20,247)		−0.55 (−1.10 to −0.01)
50–54 (*n* = 20,750)		−0.75 (−1.37 to −0.13)
55–59 (*n* = 19,353)		−1.31 (−1.95 to −0.67)
60–64 (*n* = 18,505)		−0.20 (−0.93 to −0.52)
65–69 (*n* = 19,351)		−1.11 (−1.92 to −0.30)
70–74 (*n* = 18,549)		−0.54 (−1.57 to 0.50)
75–79 (*n* = 16,524)		−0.17 (−1.63 to 1.30)
80–84 (*n* = 13,947)		0.33 (−1.67 to 2.33)
84+ (*n* = 13,253)		3.42 (0.29 to 6.56)
Sex		
Male (*n* = 76,662)		−0.09 (−0.73 to 0.54)
Female (*n* = 160,218)		−0.59 (−1.03 to −0.14)
SCARF index		
Fit (*n* = 146,339)		−0.86 (−1.18 to −0.55)
Mild frailty (*n* = 65,171)		−0.50 (−1.13 to 0.12)
Moderate frailty (*n* = 19,291)		0.65 (−1.05 to 2.36)
Severe frailty (*n* = 6079)		7.42 (3.39 to 11.46)
Charlson Comorbidity Index		
No comorbidities (*n* = 154,797)		−0.64 (−1.01 to −0.27)
One comorbidity (*n* = 61,089)		−0.25 (−0.96 to 0.46)
Two comorbidities (*n* = 16,769)		0.65 (−0.94 to 2.24)
Three or more comorbidities (*n* = 4225)		0.51 (−3.14 to 4.15)
Subdiagnosis		
K800 – Calculus of gallbladder with acute cholecystitis (*n* = 85,518)		−0.13 (−0.59 to 0.33)
K801 – Calculus of gallbladder with other cholecystitis (*n* = 66,203)		−0.37 (−0.79 to 0.05)
K802 – Calculus of gallbladder without cholecystitis (*n* = 85,159)		−0.77 (−1.26 to −0.27)
Year		
2010/11 (*n* = 21,308)		−0.61 (−1.54 to 0.32)
2011/12 (*n* = 22,378)		−1.78 (−2.61 to −0.95)
2012/13 (*n* = 22,071)		−0.23 (−1.17 to 0.71)
2013/14 (*n* = 23,380)		−1.38 (−2.13 to −0.63)
2014/15 (*n* = 23,771)		−0.28 (−1.09 to 0.54)
2015/16 (*n* = 25,097)		−0.93 (−1.63 to −0.23)
2016/17 (*n* = 25,191)		−0.40 (−1.01 to 0.21)
2017/18 (*n* = 25,303)		0.13 (−0.51 to 0.77)
2018/19 (*n* = 26,968)		0.65 (−0.00 to 1.30)
2019/20 (*n* = 21,413)		0.39 (−0.43 to 1.21)

−20   0   20   40

Favours ES   Favours NES

FIGURE 7 Mean differences in days in hospital in the first 90 days between ES and NES treatment strategies for cholelithiasis subgroups.

volumes of emergency admissions. The regression analysis found that ES was associated with an average reduction in number of DAOH for patients with diverticular disease, but did not take account of unobserved confounding.

Discussion

For patients presenting as emergency admissions to hospital with acute appendicitis, cholelithiasis, diverticular disease, an abdominal wall hernia or an intestinal obstruction, the analysis of clinical effectiveness found that, overall, the average number of DAOH at 90 days was similar following ES and NES strategies. There were differences in the relative effectiveness of ES compared with NES strategies according to patients' levels of frailty, age and number of comorbidities. For patients with
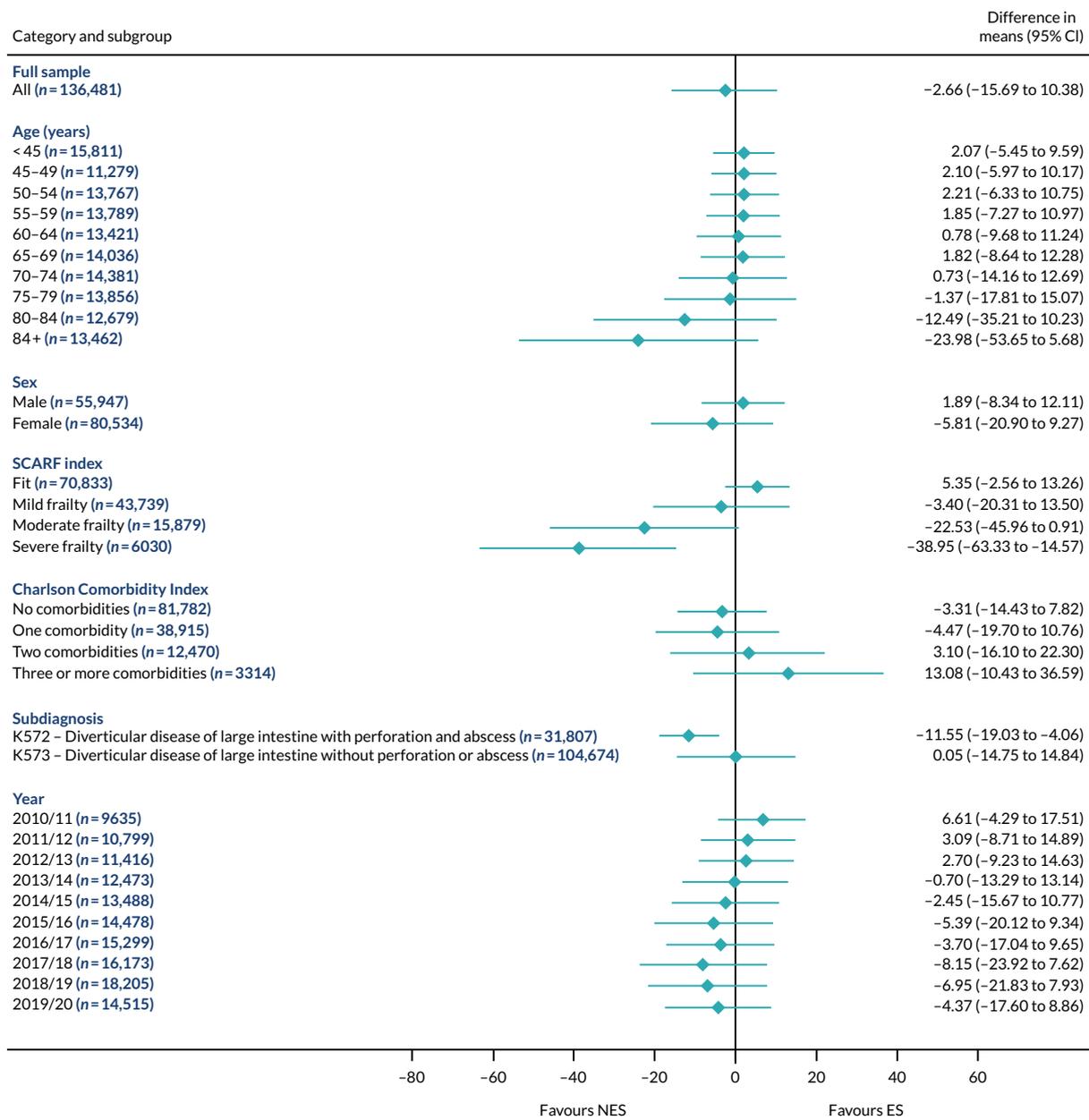
FIGURE 8 Mean differences in number of DAOH between ES and NES treatment strategies for diverticular disease subgroups.

severe frailty, according to the SCARF index, the average numbers of DAOH were smaller following ES compared with NES strategies for all five conditions. These findings have implications for clinical practice, given the importance of providing ES and NES strategies that will benefit the individual patient.

This analysis adds to the previously limited evidence[7,9,13–15,49] on the clinical effectiveness of ES compared with NES strategies, such as delayed surgery, interventional radiology or medical management. Recent trial evidence has reported equivocal results for selected patients with uncomplicated acute appendicitis, with a high proportion of patients in the NES group requiring subsequent surgery.[7,8,13] The COVID-Harem study[21] exploited the increased rates of NES strategies following the first wave of COVID-19 infections, and reported that NES strategies led to short-term cost savings. Each of these previous studies has helped developed clinical protocols for NES strategies for patients presenting as emergency admissions with acute appendicitis. The general roll-out of these protocols may be one reason for why
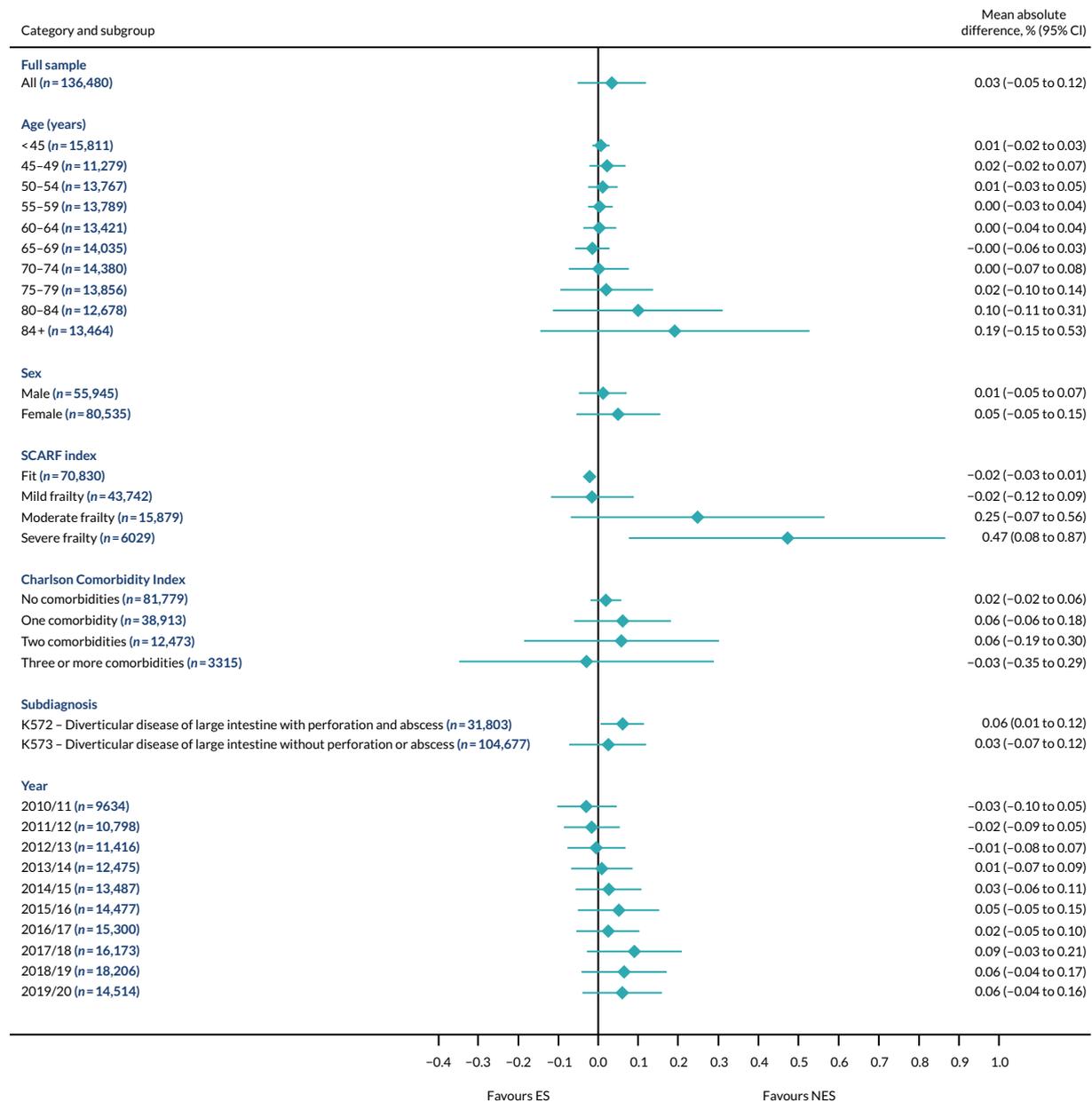
FIGURE 9 Mean differences in 90-day mortality between ES and NES treatment strategies for diverticular disease subgroups.

the ESORT study observed a relative improvement in the mean number of DAOH for NES compared with ES for the last 3 years of emergency admissions (i.e. 2017/18, 2018/19 and 2019/20). For patients presenting with acute cholelithiasis,[38] NICE guidelines, informed by meta-analyses,[39] recommend laparoscopic cholecystectomy within 7 days of diagnosis, but ES rates vary across NHS hospitals.[36,37] For patients with acute diverticular disease, published RCTs[14,15] have not recruited sufficient numbers of patients to evaluate ES and NES strategies. For patients with an abdominal wall hernia or an intestinal obstruction presenting as emergency admissions, there are, to the best of our knowledge, no published RCTs comparing ES and NES strategies. Hence, the finding that ES and NES strategies led to similar average numbers DAOH at 90 days across unselected patients routinely presenting as emergency admissions adds to limited knowledge for common acute gastrointestinal conditions.
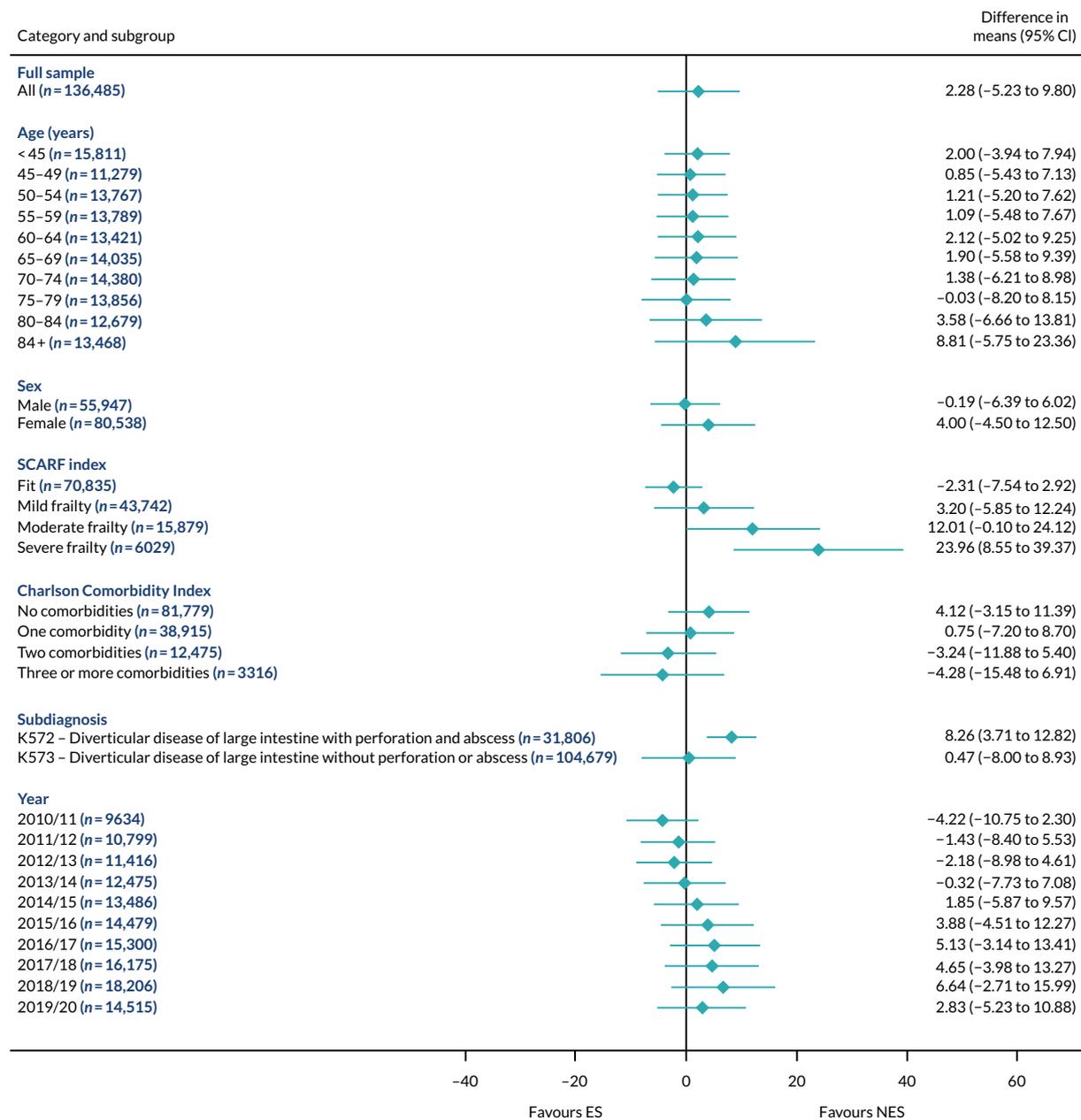
FIGURE 10 Mean differences in days in hospital in the first 90 days between ES and NES treatment strategies for diverticular disease subgroups.

National guidelines have discouraged rationing ES by age alone, and recommends further research on the relative effectiveness of ES for subgroups according to frailty and comorbidities.[28,29] In the ESORT study cohorts, for whom up to 6.5% of patients receiving ES were classified as severely frail, we found that ES was less effective than NES strategies for severely frail patients. The reductions in mean numbers of DAOH following ES were partly driven by increased all-cause mortality at 90 days (except for the hernia cohort), and partly driven by the increased LOS. RCTs of ES strategies have either excluded frail patients or not considered if frailty modifies the relative effectiveness of ES and NES strategies.[7,9,13–15,49] The current study emphasises the importance of considering frailty alongside other factors, such as age, number of comorbidities and diagnosis for patients, when deciding whether or not an individual patient should have ES.

Category and subgroup	Difference in means (95% CI)
Full sample	
All (*n* = 104,070)	−0.07 (−2.40 to 2.25)
Age (years)	
< 45 (*n* = 19,649)	2.39 (0.84 to 3.93)
45–49 (*n* = 7535)	3.48 (0.77 to 6.18)
50–54 (*n* = 7964)	4.01 (1.17 to 6.84)
55–59 (*n* = 7714)	2.29 (−0.36 to 4.95)
60–64 (*n* = 8107)	−0.72 (−3.48 to 2.03)
65–69 (*n* = 8952)	−0.19 (−3.22 to 2.83)
70–74 (*n* = 10,107)	0.24 (−3.15 to 3.62)
75–79 (*n* = 10,571)	−2.65 (−6.29 to 0.98)
80–84 (*n* = 10,673)	−3.34 (−7.89 to 1.21)
84 + (*n* = 12,798)	−4.78 (−9.88 to 0.31)
Sex	
Male (*n* = 67,378)	0.76 (−1.14 to 2.65)
Female (*n* = 36,692)	−1.59 (−5.23 to 2.05)
SCARF index	
Fit (*n* = 55,595)	2.26 (0.37 to 4.15)
Mild frailty (*n* = 32,000)	1.88 (−1.20 to 4.96)
Moderate frailty (*n* = 12,073)	−8.90 (−13.14 to −4.67)
Severe frailty (*n* = 4402)	−19.47 (−26.64 to −12.30)
Charlson Comorbidity Index	
No comorbidities (*n* = 64,096)	−0.11 (−2.17 to 1.96)
One comorbidity (*n* = 28,982)	0.56 (−2.87 to 4.00)
Two comorbidities (*n* = 8752)	−1.07 (−5.90 to 3.76)
Three or more comorbidities (*n* = 2240)	−3.45 (−12.62 to 5.72)
Subdiagnosis	
Inguinal (*n* = 50,236)	0.17 (−1.67 to 2.00)
Femoral (*n* = 13,263)	−3.30 (−9.54 to 2.93)
Umbilical (*n* = 38,430)	0.86 (−1.23 to 2.94)
Ventral (*n* = 2141)	−2.28 (−4.83 to 0.27)
Bilateral (*n* = 3288)	−0.39 (−2.19 to 1.42)
Obstruction (*n* = 46,048)	−1.45 (−5.76 to 2.85)
Gangrene (*n* = 3221)	−3.28 (−8.86 to 2.30)
Year	
2010/11 (*n* = 8952)	3.12 (−1.94 to 8.19)
2011/12 (*n* = 9356)	1.94 (−2.70 to 6.58)
2012/13 (*n* = 9587)	1.56 (−2.90 to 6.02)
2013/14 (*n* = 9919)	0.99 (−2.95 to 4.94)
2014/15 (*n* = 10,037)	−0.07 (−3.55 to 3.40)
2015/16 (*n* = 10,235)	0.07 (−3.41 to 3.56)
2016/17 (*n* = 11,283)	−1.59 (−4.54 to 1.35)
2017/18 (*n* = 11,648)	−0.49 (−3.15 to 2.16)
2018/19 (*n* = 12,851)	−1.34 (−3.75 to 1.06)
2019/20 (*n* = 10,202)	−3.67 (−6.37 to −0.96)

Favours NES   Favours ES

FIGURE 11 Mean differences in number of DAOH between ES and NES treatment strategies for hernia subgroups.

FIGURE 12 Mean differences in 90-day mortality between ES and NES treatment strategies for hernia subgroups.

Category and subgroup		Difference in means (95% CI)
Full sample		
All (*n* = 104,069)		2.35 (1.54 to 3.15)
Age (years)		
< 45 (*n* = 19,649)		−0.10 (−0.72 to 0.52)
45–49 (*n* = 7535)		−0.01 (−1.06 to 1.04)
50–54 (*n* = 7964)		0.65 (−0.43 to 1.73)
55–59 (*n* = 7714)		0.82 (−0.51 to 2.14)
60–64 (*n* = 8107)		2.29 (0.93 to 3.64)
65–69 (*n* = 8951)		1.65 (0.23 to 3.08)
70–74 (*n* = 10,107)		2.55 (1.17 to 3.93)
75–79 (*n* = 10,571)		4.27 (2.72 to 5.81)
80–84 (*n* = 10,673)		4.74 (2.98 to 6.49)
84+ (*n* = 12,798)		6.25 (4.26 to 8.25)
Sex		
Male (*n* = 67,378)		1.38 (0.63 to 2.13)
Female (*n* = 36,691)		4.12 (2.90 to 5.34)
SCARF index		
Fit (*n* = 55,594)		0.40 (−0.26 to 1.07)
Mild frailty (*n* = 32,000)		2.03 (0.92 to 3.15)
Moderate frailty (*n* = 12,073)		7.35 (5.50 to 9.20)
Severe frailty (*n* = 4402)		15.47 (12.09 to 18.85)
Charlson Comorbidity Index		
No comorbidities (*n* = 64,095)		2.02 (1.29 to 2.76)
One comorbidity (*n* = 28,982)		2.35 (1.05 to 3.65)
Two comorbidities (*n* = 8752)		4.30 (2.29 to 6.31)
Three or more comorbidities (*n* = 2240)		3.98 (−0.23 to 8.19)
Subdiagnosis		
Inguinal (*n* = 50,235)		1.57 (0.81 to 2.33)
Femoral (*n* = 13,263)		6.87 (4.81 to 8.93)
Umbilical (*n* = 38,430)		1.71 (0.95 to 2.48)
Ventral (*n* = 2141)		3.87 (2.81 to 4.93)
Bilateral (*n* = 3288)		1.47 (0.52 to 2.42)
Obstruction (*n* = 46,048)		4.70 (3.21 to 6.19)
Gangrene (*n* = 3221)		6.58 (4.61 to 8.55)
Year		
2010/11 (*n* = 8952)		1.58 (−0.48 to 3.63)
2011/12 (*n* = 9356)		2.71 (1.00 to 4.42)
2012/13 (*n* = 9587)		3.07 (1.53 to 4.61)
2013/14 (*n* = 9919)		1.41 (−0.26 to 3.09)
2014/15 (*n* = 10,037)		2.01 (0.40 to 3.61)
2015/16 (*n* = 10,235)		2.17 (0.92 to 3.42)
2016/17 (*n* = 11,282)		3.12 (1.86 to 4.38)
2017/18 (*n* = 11,648)		2.10 (0.99 to 3.22)
2018/19 (*n* = 12,851)		2.24 (1.09 to 3.40)
2019/20 (*n* = 10,202)		2.98 (1.77 to 4.19)

−20   0   20   40

Favours ES   Favours NES

FIGURE 13 Mean differences in days in hospital in the first 90 days between ES and NES treatment strategies for hernia subgroups.

Category and subgroup	Difference in means (95% CI)
Full sample	
All (*n* = 130,618)	3.32 (−3.13 to 9.76)
Age (years)	
< 45 (*n* = 14,446)	7.78 (3.63 to 11.93)
45–49 (*n* = 6382)	14.77 (9.41 to 20.12)
50–54 (*n* = 8032)	15.85 (10.41 to 21.28)
55–59 (*n* = 9073)	16.22 (10.62 to 21.83)
60–64 (*n* = 10,849)	19.18 (12.84 to 25.53)
65–69 (*n* = 13,634)	16.03 (10.14 to 21.91)
70–74 (*n* = 115,543)	7.84 (0.14 to 15.54)
75–79 (*n* = 16,552)	2.20 (−6.81 to 11.21)
80–84 (*n* = 16,137)	−8.64 (−19.35 to 2.07)
84+ (*n* = 19,970)	−24.70 (−39.82 to −9.58)
Sex	
Male (*n* = 61,895)	−3.76 (−11.95 to 4.42)
Female (*n* = 68,723)	9.69 (4.19 to 15.20)
SCARF index	
Fit (*n* = 59,029)	18.72 (14.98 to 22.47)
Mild frailty (*n* = 43,291)	3.21 (−5.53 to 11.95)
Moderate frailty (*n* = 19,696)	−26.08 (−37.58 to −14.58)
Severe frailty (*n* = 8602)	−34.54 (−46.70 to −22.39)
Charlson Comorbidity Index	
No comorbidities (*n* = 68,274)	−4.45 (−9.69 to 0.79)
One comorbidity (*n* = 41,417)	6.08 (−1.50 to 13.66)
Two comorbidities (*n* = 16,377)	22.30 (11.54 to 33.07)
Three or more comorbidities (*n* = 4550)	26.37 (8.71 to 44.02)
Subdiagnosis	
K561 – Intussusception (*n* = 1431)	4.42 (−1.33 to 10.16)
K562 – Volvulus (*n* = 15,823)	−2.35 (−11.44 to 6.74)
K563 – Gallstone ileus (*n* = 1990)	−14.15 (−19.08 to −9.22)
K565 – Intestinal adhesions (bands) with obstruction (*n* = 45,198)	−0.14 (−3.78 to 3.50)
K566 – Other and unspecified intestinal obstruction (*n* = 66,176)	7.53 (−1.80 to 16.86)
C18 – Malignant neoplasm of colon (*n* = 3462)	15.61 (7.66 to 23.55)
C19 – Malignant neoplasm of rectosigmoid junction (*n* = 318)	15.51 (8.35 to 22.67)
C20 – Malignant neoplasm of rectum (*n* = 555)	24.35 (18.30 to 30.39)
Non-cancer (*n* = 126,283)	2.86 (−3.81 to 9.53)
Year	
2010/11 (*n* = 10,825)	8.73 (−0.16 to 17.61)
2011/12 (*n* = 11,692)	3.99 (−4.89 to 12.86)
2012/13 (*n* = 11,906)	11.64 (4.20 to 19.08)
2013/14 (*n* = 12,696)	1.15 (−7.06 to 9.36)
2014/15 (*n* = 13,231)	2.00 (−5.71 to 9.70)
2015/16 (*n* = 13,701)	0.34 (−7.69 to 8.37)
2016/17 (*n* = 14,026)	3.72 (−3.60 to 11.05)
2017/18 (*n* = 14,621)	1.97 (−5.48 to 9.42)
2018/19 (*n* = 15,665)	0.80 (−6.35 to 7.95)
2019/20 (*n* = 12,255)	1.17 (−6.79 to 9.12)

−60   −40   −20   0   20   40   60

Favours NES   Favours ES

FIGURE 14 Mean differences in number of DAOH between ES and NES treatment strategies for intestinal obstruction subgroups.

This analysis has several strengths and limitations. The study included a sufficient number and range of patients with these common acute conditions to provide precise estimates of the clinical effectiveness of ES compared with NES strategies as provided in routine practice. The study used a previously developed IV method to address confounding, and provides comparative effectiveness estimates that apply to both the overall populations and subpopulations of interest. The limitations of this paper are that detailed information from imaging or diagnostic procedures was unavailable, and so the definitions of subgroups were broad. In addition, the categorisation of ES and NES strategies assumes accurate coding of OPCS-4 procedures and episode dates. Although it is conceivable that there were coding differences across NHS trusts or over time, it is unlikely that this led to substantive differences in the

Category and subgroup	Mean absolute difference, % (95% CI)
Full sample	
All (*n* = 130,617)	0.02 (−0.05 to 0.08)
Age (years)	
< 45 (*n* = 14,446)	−0.07 (−0.12 to −0.02)
45–49 (*n* = 6382)	−0.19 (−0.25 to −0.13)
50–54 (*n* = 8032)	−0.18 (−0.24 to −0.12)
55–59 (*n* = 9073)	−0.19 (−0.26 to −0.13)
60–64 (*n* = 10,848)	−0.21 (−0.27 to −0.15)
65–69 (*n* = 13,634)	−0.17 (−0.23 to −0.11)
70–74 (*n* = 15,543)	−0.05 (−0.13 to 0.04)
75–79 (*n* = 16,552)	0.01 (−0.10 to 0.12)
80–84 (*n* = 16,137)	0.20 (0.04 to 0.35)
84 + (*n* = 19,970)	0.48 (0.28 to 0.67)
Sex	
Male (*n* = 61,895)	0.09 (0.01 to 0.16)
Female (*n* = 68,722)	−0.04 (−0.11 to 0.02)
SCARF index	
Fit (*n* = 59,029)	−0.17 (−0.21 to −0.14)
Mild frailty (*n* = 43,290)	−0.00 (−0.10 to 0.10)
Moderate frailty (*n* = 19,696)	0.41 (0.26 to 0.56)
Severe frailty (*n* = 8602)	0.52 (0.36 to 0.67)
Charlson Comorbidity Index	
No comorbidities (*n* = 68,273)	0.02 (−0.02 to 0.05)
One comorbidity (*n* = 41,417)	0.03 (−0.05 to 0.12)
Two comorbidities (*n* = 16,377)	−0.02 (−0.18 to 0.14)
Three or more comorbidities (*n* = 4550)	0.02 (−0.25 to 0.28)
Subdiagnosis	
K561 – Intussusception (*n* = 1431)	−0.13 (−0.20 to −0.06)
K562 – Volvulus (*n* = 15,823)	0.11 (0.01 to 0.22)
K563 – Gallstone (*n* = 1990)	0.12 (0.05 to 0.18)
K565 – Intestinal adhesions (bands) with obstruction (*n* = 45,198)	−0.04 (−0.08 to 0.01)
K566 – Other and unspecified intestinal obstruction (*n* = 66,175)	0.03 (−0.06 to 0.13)
C18 – Malignant neoplasm of colon (*n* = 3462)	−0.17 (−0.25 to −0.09)
C19 – Malignant neoplasm of rectosigmoid junction (*n* = 318)	−0.12 (−0.21 to −0.02)
C20 – Malignant neoplasm of rectum (*n* = 555)	−0.21 (−0.29 to −0.14)
Non-cancer (*n* = 126,282)	0.02 (−0.04 to 0.09)
Year	
2010/11 (*n* = 10,825)	−0.03 (−0.12 to 0.06)
2011/12 (*n* = 11,692)	0.02 (−0.06 to 0.11)
2012/13 (*n* = 11,906)	−0.07 (−0.15 to 0.01)
2013/14 (*n* = 12,696)	0.06 (−0.02 to 0.15)
2014/15 (*n* = 13,231)	0.04 (−0.04 to 0.13)
2015/16 (*n* = 13,700)	0.05 (−0.04 to 0.13)
2016/17 (*n* = 14,026)	0.03 (−0.06 to 0.11)
2017/18 (*n* = 14,621)	0.02 (−0.06 to 0.09)
2018/19 (*n* = 15,665)	0.04 (−0.03 to 0.12)
2019/20 (*n* = 12,255)	−0.00 (−0.09 to 0.09)

Favours ES · Favours NES

FIGURE 15 Mean differences in 90-day mortality between ES and NES treatment strategies for intestinal obstruction subgroups.

estimates of relative effectiveness. The primary outcome measure (i.e. DAOH) did not consider other dimensions of health, such as health-related quality of life (HRQoL), as this information was not available within the HES data. The definition of NES strategies was limited to OPCS-4 procedure codes and, therefore, could not capture other aspects of clinical management, for example the type and duration of antibiotic therapy. The ability of the IV approach to deal with unmeasured confounding may be undermined if the requisite assumptions do not hold. The finding that the IV balanced important
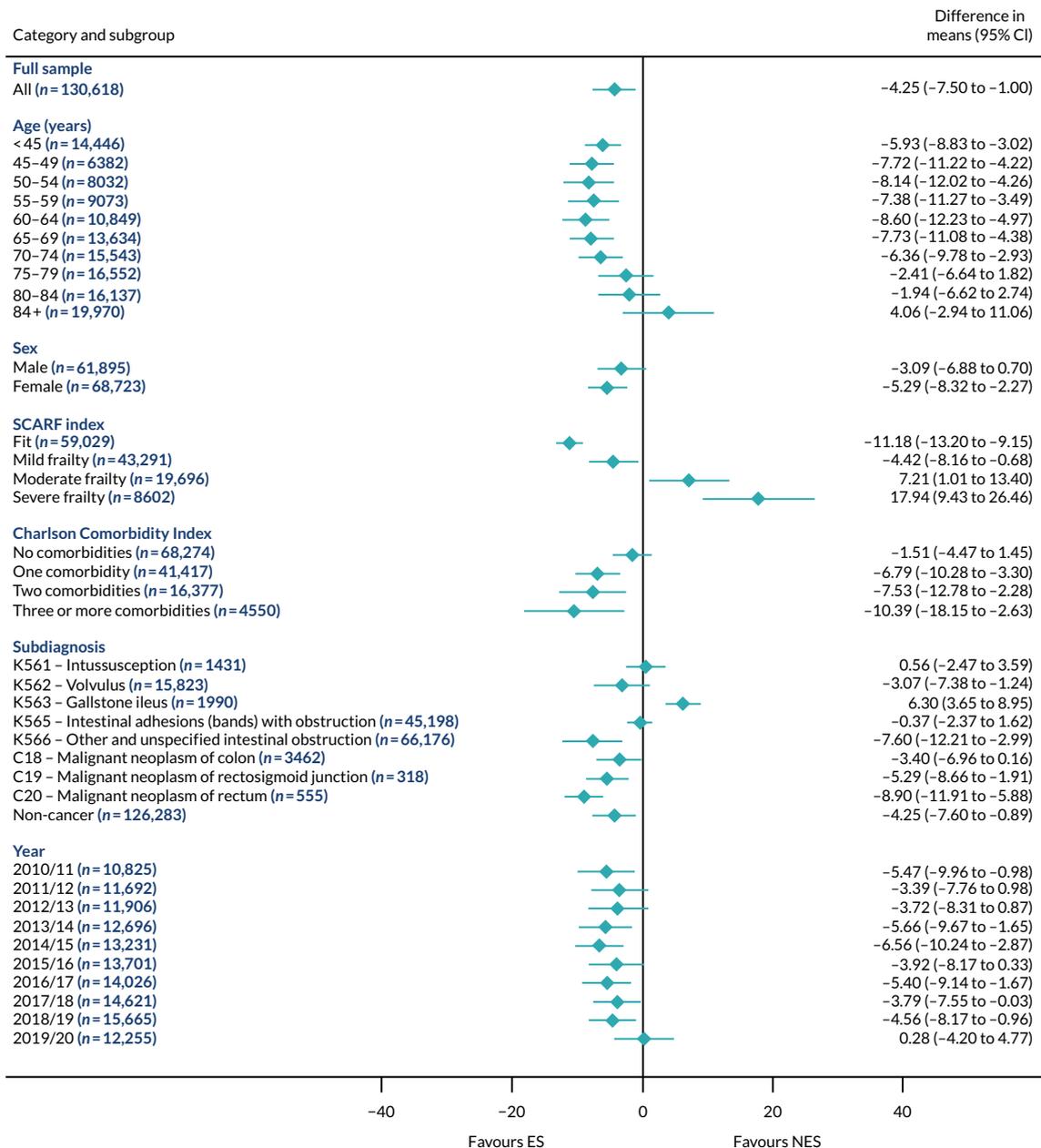
FIGURE 16 Mean differences in days in hospital in the first 90 days between ES and NES treatment strategies for intestinal obstruction subgroups.

case mix measures, such as frailty and comorbidity, provides assurance that it would also ensure that indicators of disease severity that are not available in HES data were also similarly balanced between the ES and NES groups.

In summary, this analysis found that ES and NES strategies for patients presenting as emergency hospital admissions with common acute conditions led to similar average numbers of DAOH at 90 days. For patients with severe levels of frailty, ES strategies led to worse outcomes than NES strategies for each of the five acute conditions.
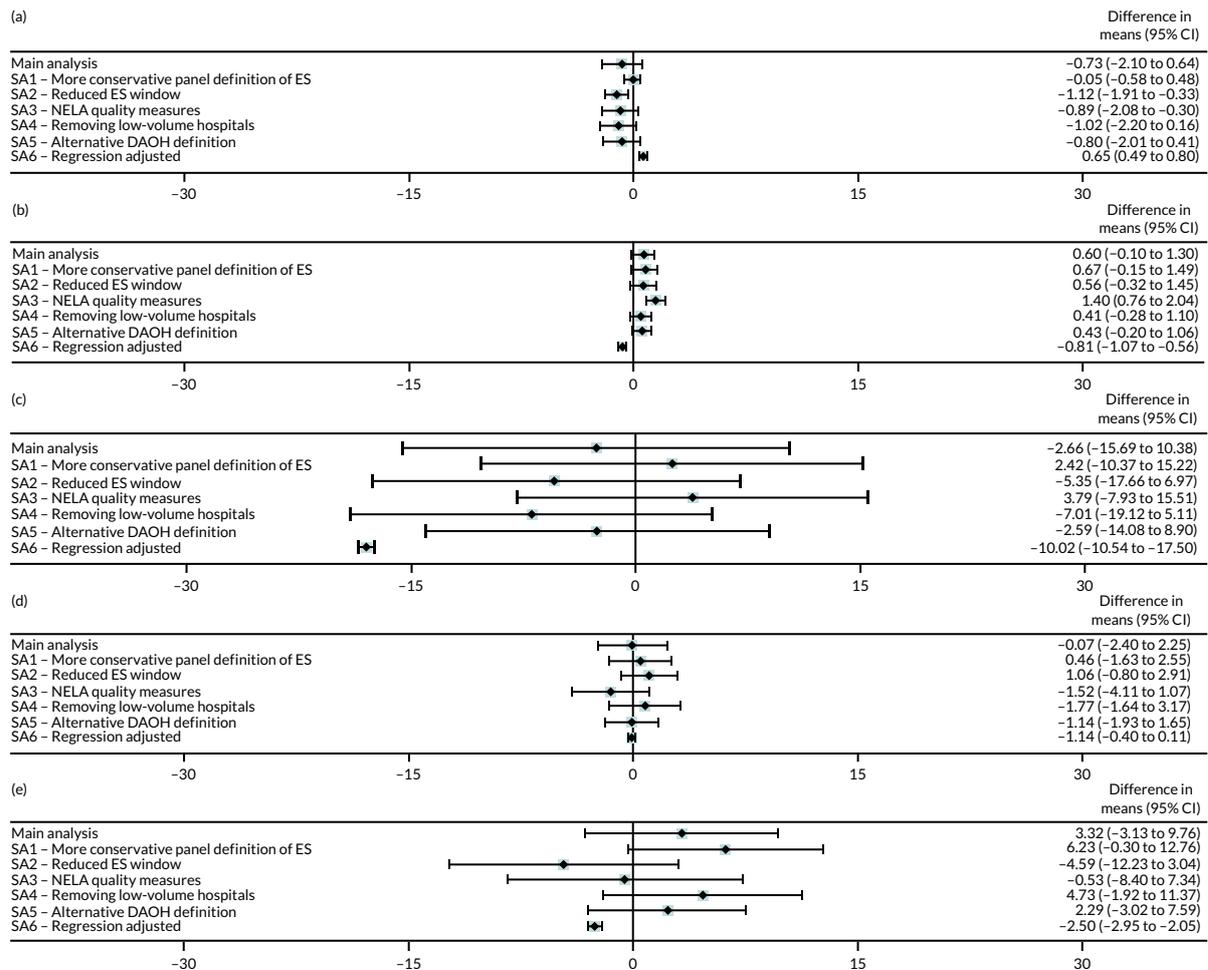
FIGURE 17 Sensitivity analysis for DAOH. (a) Appendicitis; (b) cholelithiasis; (c) diverticular disease; (d) hernia; and (e) intestinal obstruction.

Chapter 4 Cost-effectiveness

Introduction

The second and third objectives of the ESORT study required assessment of the cost-effectiveness of ES compared with NES strategies for patients with common acute gastrointestinal conditions presenting as emergency hospital admissions. The second and third objectives emphasised the importance of providing evidence on relative costs and cost-effectiveness of ES, recognising the opportunity costs of providing ES and that, for some patients, NES strategies may prove more cost-effective. Although CEA methods are relatively well established for evaluating health technologies, the evaluation of health service interventions in general, and ES in particular, raise particular challenges for recommended CEA approaches.[63,64]

The ESORT study, in common with previous CEAs of ES for cholelithiasis or appendicitis,[24,25] did not have suitable evidence from RCTs to inform the assessment of relative effectiveness within the CEA. For patients with acute appendicitis, it was unclear if the published RCT evidence[7,8,65–67] applied directly to provision of ES versus NES for NHS funded emergency admissions in England. In particular, these RCTs did not recruit from hospitals in England, nor did they include the full range of patients who would present as routine hospital admissions. For the other four acute conditions, there was, to the best of our knowledge, no RCTs that have compared ES with NES strategies for patients presenting as emergency hospital admissions. The CEA for the ESORT study, therefore, used the same IV design as for the evaluation of clinical effectiveness (see *Chapter 3*) to address concerns about confounding when evaluating the cost-effectiveness of ES compared with NES strategies.[55,68–70] The local instrumental variable (LIV) approach taken also allowed the study to meet the third objective in reporting cost-effectiveness results according to patient subgroup.[55,71]

A further challenge that arose in assessing the relative cost-effectiveness of ES from routine (HES) data was that requisite information about HRQoL was not available. Although previous CEAs of alternative forms of surgery in the elective (planned) setting have used HRQoL measures linked to HES data,[72] such linked data were not available in the emergency setting. Hence, literature reviews were undertaken for each of the five conditions to identify the most appropriate HRQoL estimates, acknowledging the challenges of HRQoL assessment for patients presenting as emergency hospital admissions.[73] We now describe the CEA methods used in the ESORT study.

Methods

Overview
The economic evaluation took the form of a CEA, with outcomes reported as life-years and quality-adjusted life-years (QALYs). The CEA took the perspective of the NHS acute hospital trust, and included the costs of inpatient hospital stays. The rationales for excluding the costs of outpatient and personal health services[63] were that the required data items were not available from HES, that it was infeasible to survey a sufficiently large patient sample within the study time frame and that previous studies reported that the majority of the incremental costs were those incurred by inpatient hospital stays.[21] The study populations were the same as those described in the evaluation of clinical effectiveness (see *Chapter 3*). In the base case, the study reported results over a 1-year time horizon (and so were undiscounted) and incorporated patient-level resource use data (from HES) and linked all-cause mortality data (from ONS). Previous studies have reported that the majority of costs following either ES or NES strategies are incurred within 1 year.[74] A recent study using data from NELA showed that most of the gains in HRQoL after ES occurred within the first 3–6 months.[75]

We carried out sensitivity analyses to assess whether or not the base-case results were robust to extending this time horizon to 5 years, with costs and QALYs in years 2–5 discounted at NICE-recommended rates of 3.5%.

Definition of emergency surgery and non-emergency surgery strategy groups

The definition of the comparison groups was identical to that for the assessment of relative effectiveness, with ES defined from OPCS-4 procedure codes and within the maximum specified time windows of within 3 days (for hernia), 7 days (for appendicitis, cholelithiasis and intestinal obstruction) or any time within the emergency admission (for diverticular disease). The NES strategies were the counterfactuals to the ES strategies and included non-operative care (e.g. antibiotic therapy) and operative procedures that did not constitute ES (e.g. a procedure that was not considered a qualifying ES procedure or was undertaken after the maximum time window for ES, or both). The CEA recognised that the ES group could have similar aspects of clinical management to the NES group, in addition to the ES strategies.

Resource use measurement

Resource use and cost measurement was carried out from the date the patient was judged eligible for the study (i.e. the date the patient was first seen by the surgical team following the index emergency admission). The study extracted resource use items from the HES data that were judged to be the major drivers of incremental cost and cost-effectiveness of ES and NES strategies for patients with the five acute conditions.[76,77] The resource use categories were the use of operative procedures (including procedures that constituted ES), the use of non-operative procedures (including interventional radiology), the use of imaging and other radiological investigation, the duration of hospital stay for the index admission (including transfers to critical care) and subsequent re-admissions up to 1 year.

For each admission, we collated data on the procedures received according to OPCS-4 procedure codes, including combinations of multiple OPCS-4 codes, where appropriate. For each index emergency hospital admission, we defined 'more common' operative procedures as any operative procedure with a prevalence of at least 1% within the index admission for each comparison group. This conservative definition of a 'more common' procedure was taken to reduce the risk of excluding important cost differences between the comparison groups. The overall list of the 'more common' procedures for the index emergency admissions across the comparison groups was then applied to define the procedures that would be costed within each re-admission, for both comparison groups. The 'common' operative procedures within re-admissions were costed using the same methods and assumptions used for operative procedures within the index admission.

For each of the 'more common' procedures, we assumed that the procedure costs were additional to those included within the unit costs per bed-day. If multiple operative procedures within the admission met the criteria for 'more common', only one procedure was costed to minimise the risk of double-counting. In identifying 'more common' procedures within each index admission or re-admission, we first considered ES procedures, then operative procedures that did not meet the ES criteria, including interventional radiology, and then, finally, non-operative procedures (e.g. catheterisation of bladder).

For admissions with procedures that did not meet the threshold for a common procedure because they were all 'low volume' (i.e. < 1% for each comparison group), we assumed that the cost was captured within the bed-day costs. We included additional costs of diagnostic tests that were 'more common', defined, again, as a prevalence within each group of at least 1%. The costs of other diagnostic tests were assumed to be covered by the general bed-day costs.

We identified all re-admissions to hospitals in England up to 1 year after the index admission (i.e. the base case). We extracted resource use for all hospital re-admissions irrespective of the diagnoses accompanying the re-admission. We extracted information for the OPCS-4 codes within the re-admission and used this information to define whether or not, within the re-admission, the patient received one of the 'more common' operative procedures either as delayed surgery (NES group) or as a

reintervention (ES or NES group). We extracted information on the overall hospital LOS within each index admission and re-admission, and summed this across all admissions to calculate the total days in hospital prior to 1 year. We accessed data on the number of critical care bed-days, with the level of care defined by the number of organs supported, according to the Adult Critical Care data linked to HES.[78] We extracted data on hospital transfers to recognise that patients may transfer to another hospital, for example for rehabilitation.

Unit costs and total costs

To calculate unit costs for each of the 'more common' operative procedures, we extracted information, from the precedent literature and expert opinion on the expected durations of the procedure, as well as the number and grade of staff involved in the procedure (see *Appendix 3, Table 23*). The use of disposables (e.g. reload staplers), equipment (e.g. ultrasound systems), surgical instruments (e.g. laparoscopic sets) and overheads were informed by expert opinion. Direct personnel costs were calculated as the costs per hour of employing each grade of staff. Overhead costs included drugs, direct Central Sterile Supply Department costs and allocated costs (e.g. rent, property and equipment maintenance, cleaning costs) associated with the provision of the procedure.[79] Purchase prices of disposables, instruments and equipment for each procedure were retrieved from different sources, including the finance department of an NHS trust hospital (i.e. Royal Devon and Exeter NHS Foundation Trust). The unit costs for each item were assigned according to the expected number of times that item would be used over the lifetime, recognising any additional costs of the sterilisation process required to enable reuse.[80]

The unit costs of critical care bed-days according to the number of organs supported, for bed-days on general wards and 'more common' diagnostic procedures were taken from 2017/18 NHS reference costs.[81] For diagnostic procedures for which the unit costs included the bed-day costs, the unit cost was calculated after subtracting the average cost per day spent on the general ward.

All unit costs were inflated to 2019/20 prices (GBP) using the UK Gross Domestic Product deflator published by HM Treasury (*Table 10*; and see also *Appendix 3, Tables 24* and *25*).[82] Each resource use measure was combined with the relevant unit costs to report (undiscounted) total costs per patient up to 1 year.

TABLE 10 Unit costs (£ 2019/20) for potential cost drivers

Item	Unit	Unit cost (£)	Source, definitions and assumptions
Inpatient stay			
General ward	Day	347	NHS reference costs:[81] weighted average of codes FD05A (Abdominal Pain with Interventions) and FD05B (Abdominal Pain without Interventions) in non-elective excess bed-days sheet
ICU ward			
Level 2 ICU	Day	1190	NHS reference costs:[81] XC06Z (Adult Critical Care, 1 organ supported)
Level 3 ICU	Day	1890	NHS reference costs:[81] weighted average of codes XC01Z (Adult Critical Care, 6 or more Organs Supported) to XC05Z (Adult Critical Care, 2 Organs Supported)

ICU, intensive care unit.

Note
The complete list of unit cost items can be found in *Appendix 3, Tables 24* and *25*.

Outcomes

Survival time up to 1 year was calculated for all patients from HES records linked to ONS death records, and the survival time was used to report life-years for each patient. To calculate QALYs, HRQoL measures were required at 'baseline' and '1-year follow-up'. Although HRQoL measures were not available from HES, we identified appropriate estimates by reviewing the published literature for each of the five acute conditions (see *Appendix 4, Search for appropriate health-related quality-of-life scores and adjustment* and *Box 1*). In brief, the inclusion criteria for each review required that studies assessed HRQoL for patients with each of the five acute conditions. From the published sources identified, we selected studies according to the following criteria (in order of priority): (1) the study measured HRQoL following emergency admission for patients with each of the diagnoses described previously (see *Chapter 2*), (2) the study used a recommended generic HRQoL instrument (e.g. EuroQol-5 Dimensions) for patients in the UK or a country with similar health state preferences and (3) the study was undertaken in the emergency setting. The HRQoL values selected for the 'baseline' and 'follow-up' time points for each of the five conditions are listed in *Table 11* (see *Appendix 4, Search for appropriate health-related quality-of-life scores and adjustment*).

In calculating QALYs, for survivors at 1 year, it was assumed that a patient's HRQoL was reduced for the duration of the initial emergency hospital admission and that following hospital discharge the patient's HRQoL level immediately recovered to the average HRQoL level reported in the literature at the 1-year follow-up. For those patients who had an emergency re-admission during the 1-year follow-up, it was assumed that the HRQoL following the emergency re-admission reverted to the same 'baseline' level as for the initial (index) emergency admission. It was also assumed that following hospital discharge the HRQoL levels reverted to those at 1-year follow-up (see *Appendix 4, Figure 30*). The assumption that HRQoL reverted to follow-up levels immediately after hospital discharge was challenged in a sensitivity analysis (SA) in which QALYs were calculated using linear interpolation between the index emergency admission and 1-year follow-up (see *Appendix 4, Figure 30*). For patients who died prior to 1 year, a HRQoL score of zero was applied.

Therefore, the approach to estimating QALYs assumed that events that did not lead to emergency re-admissions (e.g. planned surgery for recurrence) had no impact on a patient's HRQoL, and this assumption was informed by the precedent HRQoL literature, for example on HRQoL for hernia repairs in the elective setting.[86,87] It was also assumed that the only differential effect of the comparison group on 1-year QALYs was according to 1-year mortality or the rate or duration of emergency re-admissions, both of which were derived from individual-level HES data.

TABLE 11 Health-related quality-of-life scores from the literature and sources

Condition	Source	Mean age (years)[a]	EQ-5D-3L score Baseline Females	Males	1 year Females	Males
Appendicitis	O'Leary *et al.*[8]	33	0.751	0.768	0.967	0.989
Cholelithiasis	Sutherland *et al.*[83]	58	0.853	0.832	0.916	0.893
Diverticular disease	Thornell *et al.*[14]	68	0.649	0.666	0.866	0.889
Hernia	Rutegård *et al.*[84]	59	0.848	0.870	0.936	0.960
Intestinal obstruction	Young and Zahid[85]	66	0.706	0.687	0.173	0.169

EQ-5D-3L, EuroQol-5 Dimensions, three-level version.
a Mean age at trial start in the study.

The QALY calculations used HRQoL average levels extracted from the literature, according to the age and sex of each individual patient.[88,89] The study reported incremental net monetary benefits (INBs) of ES compared with NES strategies by multiplying the incremental QALYs by the NICE-recommended willingness-to-pay threshold of £20,000 per QALY, and subtracting this from the total cost (£) at 1 year.[63]

Statistical analysis

The CEA used the same LIV approach as for the evaluation of clinical effectiveness (see *Chapter 3*).[19,20,90] In the ESORT study, the IV was the hospital's TTO, which reflected practice variation across hospitals in ES rates for these conditions (see *Appendix 2, Tendency to operate as an instrumental variable* and *Person-level instrumental variable approach*). For each qualifying emergency admission, the TTO was defined as the proportion of eligible emergency patient admissions in that specific hospital who received ES in the previous 12 months, therefore, requiring that the hospital's past preference for ES strongly predicted the choice of ES or a NES strategy for the current patient. As described in *Chapter 3*, although the underlying assumptions behind the IV approach were plausible, we also adjusted for a rich set of potential confounders, including proxies for quality of acute care (see *Appendix 2, Proxies for the quality of acute care*).

The LIV approach was used to estimate the potential outcomes for each individual following both ES and NES strategies. The predicted outcomes were used in summarising main resource use measures, such as the percentages of patients who had common operative procedures, were transferred to critical care units or had a hospital re-admission or reintervention, and the mean total number of days in hospital up to 1 year.

We calculated person-centred treatment effects of ES and NES strategies for the total number of hospital days, costs, life-years, QALYs and INBs for each individual, allowing for treatment effect heterogeneity and confounding.[14,83–85] The estimated effects were aggregated to report the effects of ES overall and for each prespecified subgroup of interest (see *Chapter 3, Subgroup results*). Probit regression models were used to estimate the initial propensity score (first stage). GLMs were then applied to each end point, with the most appropriate GLM chosen, amongst those models that converged, for each end point and condition [according to root mean squared error (RMSE)] (see *Appendix 5*). Hosmer–Lemeshow and Pregibon tests were also used to check model fit and appropriateness.[91,92] For LOS within the index admission, the Poisson family with log-link was chosen for all conditions. For the total LOS, life-years and QALYs at 1 year, the end points were rescaled [0,1], which enabled a model with binomial family and log-link to be applied, which was also the best-fitting model for the binary resource use measures (all five conditions). For costs, the Gaussian family with log-link was selected for the appendicitis, diverticular disease and cholelithiasis cohorts, the Gaussian family with identify link was selected for the hernia cohort and the Poisson family with log-link to handle the heavily right-skewed costs was selected for the intestinal obstruction cohort.

Models at both stages adjusted for confounding factors for costs and outcomes, together with the time period and proxies for hospital quality. These proxies for hospital quality were defined by the rates of emergency admission and mortality for each hospital and acute condition in 2009–10 (i.e. time constant) and in the year prior to the specific admission concerned (i.e. time varying) (see *Appendix 2, Proxies for the quality of acute care*).

Overall estimates of incremental resource use, costs, life-years, QALYs and INBs were reported with CIs using bootstrapped (300 replications) standard errors that allowed for the clustering of individuals within hospitals, and the correlation of individual-level costs and effects. The individual-level estimates of incremental costs and QALYs were also plotted on the cost-effectiveness plane, stratified by subgroups of policy relevance.

Sensitivity analyses

Sensitivity analyses were undertaken to assess whether or not the results from the main analysis were robust to alternative definitions and assumptions. First, the CEA, like the evaluation of clinical effectiveness, considered whether or not the results were robust to the choice of adjustment measure for 'quality of care' by using external hospital performance measures from NELA, rather than those from the preceding HES admissions data (i.e. SA1).[59–61] Second, to assess the sensitivity of the results to assumptions about unit costs, the sensitivity analyses considered alternative scenarios. Specifically, the inclusion of the full unit costs of operative and diagnostic procedures for the 'more common' operative and diagnostic procedures risks double counting of items (e.g. consumables) that may be included within the overall costs per bed-day. Conversely, the exclusion of specific additional costs for the 'less common' operative procedures for both comparison groups may have led to an underestimate of costs for both groups. To investigate whether or not either standpoint led to a large inaccuracy in the INB estimates, we increased and decreased unit costs by 10% (i.e. SAs 2 and 3).

Third, we considered an alternative approach to QALY calculation that assumed a gradual increase in HRQoL following emergency (re-)admission by applying a linear interpolation to the HRQoL estimates between the baseline and 1-year follow-up (i.e. SA4) (see *Appendix 4, Figure 30*). Fourth, to consider whether or not there were important differences between the comparison groups after 1 year, we repeated the base-case analysis, but with a longer time horizon of 5 years. We conducted this 5-year analysis by restricting the sample to patients who were admitted in 2010–14, which reduced the sample size for each cohort, but avoided making the additional assumptions implied by extrapolating for the periods beyond the observed data (i.e. SA5). Fifth, to investigate the role of unmeasured confounding, we undertook conventional risk-adjustment (GLM regression) approaches, which followed the same approach to model specification and selection, to report overall estimates of INBs (i.e. SA6). For each of these sensitivity analyses, we reported estimates of the relative cost-effectiveness of ES compared with NES strategies for each acute condition (see *Table 18*).

Results

Unadjusted resource use and costs

Tables 12–15 describe differences in resource use and costs for the ES and NES groups prior to any adjustment for confounding. *Table 12* reports the unadjusted resource use for the index admissions and re-admissions up to 1 year. The re-admission rates were generally higher for the NES group than for the ES group across all the conditions. The mean numbers of days in hospital at 1 year were broadly similar across the comparison groups for appendicitis, cholelithiasis and hernia. However, the mean numbers of days in hospital at 1 year were larger in the ES group, than in NES group, for diverticular disease and intestinal obstruction.

For the appendicitis, cholelithiasis and hernia cohorts, the mean LOSs in the index admission were similar across the comparison groups. For the diverticular disease and intestinal obstruction cohorts, the mean LOS was longer for the ES group. For patients with appendicitis, 53% of patients in the NES group had a hospital re-admission within 1 year of the index admission and 16% had a common operative procedure (see *Table 12*). For patients with cholelithiasis, the majority (80%) of patients in the NES group had a re-admission, including patients who had a later common operative procedure (60%), whereas the corresponding proportions of patients in the ES group were lower (45% and 14%, respectively). The majority of patients with diverticular disease and intestinal obstruction in the NES group did not receive common operative procedures in the index admission or in re-admissions. For patients with a hernia, the proportion of patients who had a re-admission in the NES group was approximately twice that in the ES group (73.8% vs. 38.9%).

The most common operative procedures for each condition during and after the ES window are shown in *Tables 13* and *14*. For patients with appendicitis, cholelithiasis and hernia, the most common

TABLE 12 Unadjusted hospital resource use up to 1 year for ES and NES strategies

Resource use	Appendicitis (N = 268,144)		Cholelithiasis (N = 240,977)		Diverticular disease (N = 138,869)		Hernia (N = 106,432)		Intestinal obstruction (N = 133,073)	
	ES (n = 247,506)	NES strategies (n = 20,638)	ES (n = 52,004)	NES strategies (n = 188,973)	ES (n = 15,772)	NES strategies (n = 123,097)	ES (n = 62,559)	NES strategies (n = 43,873)	ES (n = 40,550)	NES strategies (n = 92,523)
Index admission										
'More common' operative procedure,[a] n (%)	244,782 (98.9)	3650 (17.7)	47,798 (91.9)	12,651 (6.7)	14,097 (89.4)	406 (0.3%)	58,662 (93.8)	2133 (4.9)	36,846 (90.9)	4209 (4.6)
Critical care, n (%)	5249 (2.1)	215 (1.0)	1827 (3.5)	1410 (0.8)	8543 (54.2)	792 (0.6)	4949 (7.9)	634 (1.5)	15,497 (38.2)	3026 (9.3)
Mean (SD) days in hospital	4.4 (3.7)	5.2 (8.1)	6.6 (6.5)	5.3 (6.3)	19.5 (17.9)	5.3 (6.4)	5.6 (18.3)	3.1 (8.4)	15.5 (22.9)	8.3 (34.1)
Re-admissions										
Re-admission (any time prior to 1 year), n (%)	66,615 (26.9)	10,970 (53.2)	23,294 (44.8)	152,697 (80.8)	10,289 (65.2)	93,234 (75.7)	26,428 (42.2)	32,583 (74.3)	22,110 (54.5)	58,851 (63.6)
'More common' operative procedure[a] (within any re-admission), n (%)	5578 (2.3)	3338 (16.2)	7220 (13.9)	112,849 (59.7)	3289 (20.9)	16,248 (13.2)	5053 (8.1)	20,712 (47.2)	2682 (6.6)	10,460 (11.3)
Total hospital days at 1 year	5.9 (7.9)	8.9 (13.0)	10.2 (13.5)	12.1 (16.05)	28.1 (27.4)	11.7 (17.3)	10.2 (17.8)	9.8 (18.7)	23.4 (25.8)	18.4 (25.1)
Total hospital days at 1 year: incremental effect (95% CI)	−3.1 (−3.23 to −2.87)		−1.91 (−2.0 to −1.8)		16.5 (16.0 to 16.9)		0.3 (0.1 to 0.5)		5.0 (4.7 to 5.3)	

SD, standard deviation.
a Percentage of patients who received a procedure with a prevalence that exceeded 1% in the index admission.

TABLE 13 Operative procedures during ES window for ES and NES strategies

Operative procedure	Appendicitis (N = 268,144)		Cholelithiasis (N = 240,977)		Diverticular disease (N = 138,869)		Hernia (N = 106,432)		Intestinal obstruction (N = 133,073)	
	ES (n = 247,506)	NES strategies (n = 20,638)	ES (n = 52,004)	NES strategies (n = 188,973)	ES (n = 15,772)	NES strategies (n = 123,097)	ES (n = 62,559)	NES strategies (n = 43,873)	ES (n = 40,550)	NES strategies (n = 92,523)
'More common' operative procedures[a] (%)	Emergency excision of abnormal appendix (63.0) Unspecified other excision of appendix (16.6) Emergency excision of abnormal appendix and drainage (9.9) Other (5.6)	Interval appendicectomy (6.7) Other specified other excision of appendix (4.2) Planned delayed appendicectomy (1.6) Other (2.5)	Total cholecystectomy (66.8) Percutaneous drainage of gall bladder (3.52) Partial cholecystectomy (2.58) Other (4.2)	Endoscopic sphincterotomy of sphincter of Oddi and removal of calculus (1.8) Unspecified urethral catheterisation of bladder (0.3) Other (0.0)	Rectosigmoidectomy and closure of rectal stump and exteriorisation of bowel (55.6) Irrigation of peritoneal cavity (7.7) Sigmoid colectomy and exteriorisation of bowel (6.6) Other (19.5)	Fibreoptic endoscopic snare resection of lesion of colon (0.1) Endoscopic division of adhesions of peritoneum (0.1) Endoscopic snare resection of lesion of lower bowel using fibreoptic sigmoidoscope (0.1) Other (0.0)	Primary repair of inguinal hernia using insert of prosthetic material (27.6) Repair of umbilical hernia using sutures (17.9) Repair of umbilical hernia using insert of prosthetic material (12.3) Other (31.7)	Unspecified urethral catheterisation of bladder (0.5) Ileectomy and anastomosis of ileum to ileum (0.3) Unspecified excision of ileum (0.1) Other (0.0)	Freeing of adhesions of peritoneum (28.0) Ileectomy and anastomosis of ileum to ileum (7.6) Endoscopic division of adhesions of peritoneum (6.4) Other (37.0)	Other specified other therapeutic endoscopic operations on lower bowel using fibreoptic sigmoidoscope (1.3) Unspecified opening of abdomen (0.5) Other (0.0)
% with no 'more common' operative procedures[b]	4.9	84.9	22.9	97.8	10.6	99.7	10.5	98.9	21.1	98.2

a 'Other' includes procedures with > 1% volume in index admission appearing in ES window.
b Includes patients for whom no procedures were recorded and patients who had 'low-volume' (i.e. < 1%) procedures.

Note
The denominator is the total number of patients in the group.

TABLE 14 Operative procedures after the ES window and up to one year for ES and NES strategies

| Operative procedure | Appendicitis (N = 268,144) | | Cholelithiasis (N = 240,977) | | Diverticular disease (N = 138,869) | | Hernia (N = 106,432) | | Intestinal obstruction (N = 133,073) | |
	ES (n = 247,506)	NES strategies (n = 20,638)	ES (n = 52,004)	NES strategies (n = 188,973)	ES (n = 15,772)	NES strategies (n = 123,097)	ES (n = 62,559)	NES strategies (n = 43,873)	ES (n = 40,550)	NES strategies (n = 92,523)
'More common' operative procedures[a] (%)	Emergency excision of abnormal appendix (2.8) Emergency excision of abnormal appendix (0.8) Emergency excision of abnormal appendix (0.45) Other (1.9)	Unspecified urethral catheterisation of bladder (4.3) Emergency excision of abnormal appendix (4.3) Unspecified other excision of appendix (2.8) Other (6.8)	Total cholecystectomy (20.8) Endoscopic sphincterotomy of sphincter of Oddi and removal of calculus (2.0) Partial cholecystectomy (1.1) Other (3.1)	Total cholecystectomy (48.3) Endoscopic sphincterotomy of sphincter of Oddi and removal of calculus (5.8) Partial cholecystectomy (1.4) Other (2.4)	Closure of colostomy (9.4) Rectosigmoidectomy and closure of rectal stump and exteriorisation of bowel (8.0) Freeing of adhesions of peritoneum (1.3) Other (9.4)	Fibreoptic endoscopic snare resection of lesion of colon (4.0) Fibreoptic endoscopic resection of lesion of colon (1.7) Rectosigmoidectomy and closure of rectal stump and exteriorisation of bowel (1.2) Other (5.4)	Primary repair of inguinal hernia using insert of prosthetic material (3.8) Repair of umbilical hernia using sutures (1.7) Unspecified urethral catheterisation of bladder (1.5) Other (4.8)	Primary repair of inguinal hernia using insert of prosthetic material (26.9) Repair of umbilical hernia using insert of prosthetic material (6.2) Repair of umbilical hernia using sutures (5.6) Other (10.6)	Freeing of adhesions of peritoneum (5.7) Ileectomy and anastomosis of ileum to ileum (1.5) Endoscopic division of adhesions of peritoneum (1.2) Other (9.2)	Freeing of adhesions of peritoneum (2.9) Other specified other therapeutic endoscopic operations on lower bowel using fibreoptic sigmoidoscope (1.8) Ileectomy and anastomosis of ileum to ileum (1.1) Other (7.4)
% with no 'more common' opeative procedures[b]	94.1	81.8	73.0	42.0	72.0	87.7	88.1	50.7	82.4	86.8

a 'Other' includes procedures with > 1% volume in index admission appearing in ES window.
b Includes patients for whom no procedures were recorded and patients who had 'low-volume' (i.e. < 1%) procedures.

Note
The denominator is the total number of patients in the group.

procedure within the ES window was received by more than 50% of patients in the ES group (see *Table 13*). For the NES group, the rates of later surgery were high for cholelithiasis (58.0%) and hernia (49.3%), and lower for appendicitis (18.2%), intestinal obstruction (13.2%) and diverticular disease (12.3%) (see *Table 13*). For the appendicitis cohort, most patients who had delayed surgery had appendicectomies. For the cholelithiasis cohort, more than 50% of patients in the NES group had a total or partial cholecystectomy. For the hernia cohort, the most common delayed surgery was hernia repair.

Table 15 presents the unadjusted costs of ES and NES at 1 year. For patients with diverticular disease, the mean total costs for the ES group at 1 year were higher than for the NES strategy (£16,498 vs. £4673), reflecting the higher initial admission costs, including operative costs, and also the higher re-admission costs. For intestinal obstruction, the costs were substantially higher in the ES group than in the NES group (£12,324 vs. £7791), despite the higher re-admission costs observed in the NES group. For the other three conditions, the average 1-year costs of ES and NES were similar, with the higher operative costs of ES offset by the higher re-admission costs for NES.

Estimates of resource use, costs, life-years, quality-adjusted life years and incremental net monetary benefits, after adjustment for confounding

The results of the analysis following adjustment with the LIV approach are shown in *Tables 16* and *17*. Following adjustment for confounding between the comparison groups, the mean differences in LOS within the index admission (all conditions) and in the proportion of re-admissions (appendicitis and cholelithiasis) were reduced. For the appendicitis, cholelithiasis, diverticular disease and hernia cohorts, the difference between ES and NES groups in average number of hospital days at 1 year was small, following adjustment. For patients with an intestinal obstruction, the ES group had a 7-day reduction in the number of hospital days at 1 year, compared with the NES group, after adjustment.

Table 17 presents the results of the LIV models for incremental costs, life-years, QALYs and INBs of ES compared with NES strategies. For each of the five conditions, the 95% CIs around the INB estimates included zero, which suggests that, overall, neither ES or NES strategies could be judged as relatively cost-effective, given the uncertainty. For patients with appendicitis, cholelithiasis or a hernia, the mean INBs were all close to zero. For patients with diverticular disease or an intestinal obstruction, the INBs were £2664 and £728, respectively, but with wide 95% CIs surrounding these point estimates. For appendicitis and cholelithiasis, the choice of ES or NES strategies resulted in small overall differences in mean costs (−£109 vs. −£76.8), life-years (−0.00334 vs. −0.00862) and QALYs (−0.00973 vs. 0.00720). For hernia, the ES strategy had higher mean costs (£891), life-years (0.0461) and QALYs (0.0386) than the NES strategies. The positive INBs following ES for diverticular disease and intestinal obstruction were driven by mean cost reductions (of −£1724 and −£2328, respectively) and, in the case of diverticular disease, also average QALY gains (0.0471) for the ES group; however, the uncertainty around each of these adjusted estimates was high.

Subgroup results for the cost-effectiveness analysis

Figures 18–22 reported underlying heterogeneity in the INB estimates according to subgroup. Three results are worth emphasising (for a summary, see *Table 20*). First, the relationship between age group and the mean INB estimates differed by condition. Although ES was more cost-effective for younger patients with cholelithiasis or a hernia, for patients with diverticular disease the estimated INBs were higher for patients aged ≥ 55 years. For patients with acute appendicitis or an intestinal obstruction, there is not a clear pattern in the INB estimates across age groups. Second, the relative cost-effectiveness of ES compared with NES strategies differed according to the number of comorbidities. For example, on average, for patients with two or more comorbidities, ES was the more cost-effective strategy for patients with diverticular disease or an intestinal obstruction, whereas for patients with appendicitis, cholelithiasis or a hernia, NES was the more cost-effective strategy. However, the CIs for the INBs for both sets of estimates included zero. Third, the cost-effectiveness of ES differed by frailty subgroups. For the appendicitis, cholelithiasis, hernia and intestinal obstruction cohorts, there was strong evidence that NES strategies were cost-effective among patients with severe frailty, with the

TABLE 15 Unadjusted costs (£ 2019/20) of ES and NES strategies

Cost	Appendicitis (N = 268,144)		Cholelithiasis (N = 240,977)		Diverticular disease (N = 138,869)		Hernia (N = 106,432)		Intestinal obstruction (N = 133,073)	
	ES (n = 247,506)	NES strategies (n = 20,638)	ES (n = 52,004)	NES strategies (n = 188,973)	ES (n = 15,772)	NES strategies (n = 123,097)	ES (n = 62,559)	NES strategies (n = 43,873)	ES (n = 40,550)	NES strategies (n = 92,523)
Index admission costs, mean (SD)										
Bed-day costs	1610 (2080)	1850 (3150)	2480 (3450)	1880 (2610)	10,600 (12,900)	1880 (2510)	2250 (7040)	1180 (3850)	7280 (10,700)	3050 (12,100)
Cost of diagnostic procedures	28.0 (54.2)	57.8 (69.1)	49.2 (94.0)	60.6 (107)	108 (104)	86.5 (81.4)	20.3 (52.3)	18.2 (45.1)	95.9 (94.5)	72.4 (98.7)
Cost of operative procedures	1130 (127)	192 (429)	1100 (387)	63.3 (248)	1950 (938)	1.68 (32.8)	809 (244)	42.3 (209)	1500 (684)	68.6 (339)
Total costs in index admission	2770 (1970)	2100 (3210)	3630 (3500)	2000 (2720)	12,700 (13,100)	1970 (2540)	3080 (7070)	1240 (3940)	8880 (10,800)	3200 (12,100)
Re-admissions up to 1 year costs, mean (SD)										
Bed-day costs	541 (2600)	1410 (4210)	1360 (4930)	2570 (5960)	3440 (8030)	2420 (6170)	1790 (6000)	2580 (7410)	3250 (8370)	4280 (9570)
Cost of diagnostic procedures	22.5 (80.2)	70.2 (142)	50.0 (131)	83.3 (168)	94.4 (149)	146 (174)	33.5 (100)	45.7 (120)	66.3 (140)	107 (200)
Cost of operative procedures	18.5 (139)	178 (419)	164 (430)	779 (689)	270 (628)	137 (496)	62.7 (242)	406 (457)	121 (519)	213 (681)
Total costs in re-admissions	582 (2650)	1660 (4340)	1570 (5070)	3440 (6080)	3810 (6370)	2710 (6740)	1880 (6100)	3030 (7470)	3440 (8580)	4600 (9860)
Total costs at 1 year	3360 (3520)	3760 (5660)	5200 (6450)	5440 (6900)	16,500 (16,000)	4670 (7150)	4960 (9670)	4280 (8680)	12,300 (14,500)	7790 (15,800)

SD, standard deviation.

TABLE 16 Estimated resource use for ES and NES strategies from the LIV approach, which adjusts for confounding

Resource use	Appendicitis (n = 268,144)		Cholelithiasis (n = 240,977)		Diverticular disease (n = 138,869)		Hernia (n = 106,432)		Intestinal obstruction (n = 133,073)	
	ES	NES strategies	ES	NES strategies	ES	NES strategies	ES	NES strategies	ES	NES strategies
Index admission										
'More common' operative procedure[a] (%)	96.7	100	55.3	7.4	29.9	3.3	89.5	7.3	45.9	4.3
Critical care (%)	2.5	4.7	1.8	1.0	9.5	2.2	10.0	9.1	53.5	11.0
Mean (SD) days in hospital	4.4 (4.3)	4.9 (4.4)	5.7 (6.5)	5.2 (6.3)	5.1 (12.2)	4.9 (8.7)	4.5 (15.4)	2.5 (15.0)	7.0 (33.0)	9.3 (31.7)
Re-admissions										
Re-admission (any time prior to 1 year) (%)	27.9	37.0	53.4	83.3	56.6	77.4	38.9	73.8	38.6	61.0
'More common' operative procedure (within any re-admission)[a] (%)	2.6	0.5	19.2	63.2	18.4	11.9	6.5	47.9	20.7	11.7
Total hospital days at 1 year	6.0 (8.9)	6.4 (8.5)	11.0 (16.9)	12.4 (15.5)	15.1 (23.9)	10.7 (19.2)	9.4 (19.0)	7.9 (18.0)	10.4 (31.6)	17.4 (27.2)
Total hospital days at 1 year: incremental effect	−0.4		−1.4		4.4		1.5		−7	

SD, standard deviation.
a Percentage of patients who received a procedure with a prevalence that exceeded 1% in the index admission (see *Resource use measurement*).

TABLE 17 Estimated INB (£ 2019/20), costs (£ 2019/20), life-years and QALYs of ES vs. NES strategies, from the LIV approach, adjusting for confounding

End point	Appendicitis (n = 268,144)	Cholelithiasis (n = 240,977)	Diverticular disease (n = 138,869)	Hernia (n = 106,432)	Intestinal obstruction (n = 133,073)
Mean NMB (£)					
ES	15,475	12,059	15,045	12,257	–5836
NES strategies	15,561	11,838	12,382	12,377	–6563
INBª (£), mean (95% CI)	–86.2 (–1163 to 991)	221 (–450 to 892)	2664 (–4298 to 9626)	–119 (–1282 to 1043)	728 (–2161 to 3617)
Mean costs (£)					
ES	3366	5477	2643	5052	9281
NES strategies	3475	5554	4365	4161	9817
Incremental costs (£), mean (95% CI)	–109 (–1130 to 913)	–76.8 (–702 to 548)	–1724 (–7878 to 4430)	891 (20.7 to 1762)	–535 (–3448 to 2376)
Mean QALYs					
ES	0.942	0.877	0.884	0.865	0.172
NES strategies	0.952	0.870	0.837	0.827	0.163
Incremental QALYs, mean (95% CI)	–0.00973 (–0.0226 to 0.00316)	0.00720 (–0.000871 to 0.0153)	0.0471 (–0.0829 to 0.177)	0.0386 (0.00430 to 0.0729)	0.00962 (–0.0155 to 0.0347)
Mean life-years					
ES	0.996	0.970	0.967	0.956	0.832
NES strategies	0.999	0.978	0.963	0.910	0.825
Incremental life-years, mean (95% CI)	–0.00334 (–0.00572 to –0.000965)	–0.00862 (–0.0219 to 0.00466)	0.00442 (–0.0730 to 0.0818)	0.0461 (0.00428 to 0.0880)	0.00693 (–0.0583 to 0.0722)

NMB, net monetary benefit.
a Mean difference in the NMB between the ES and NES strategies.
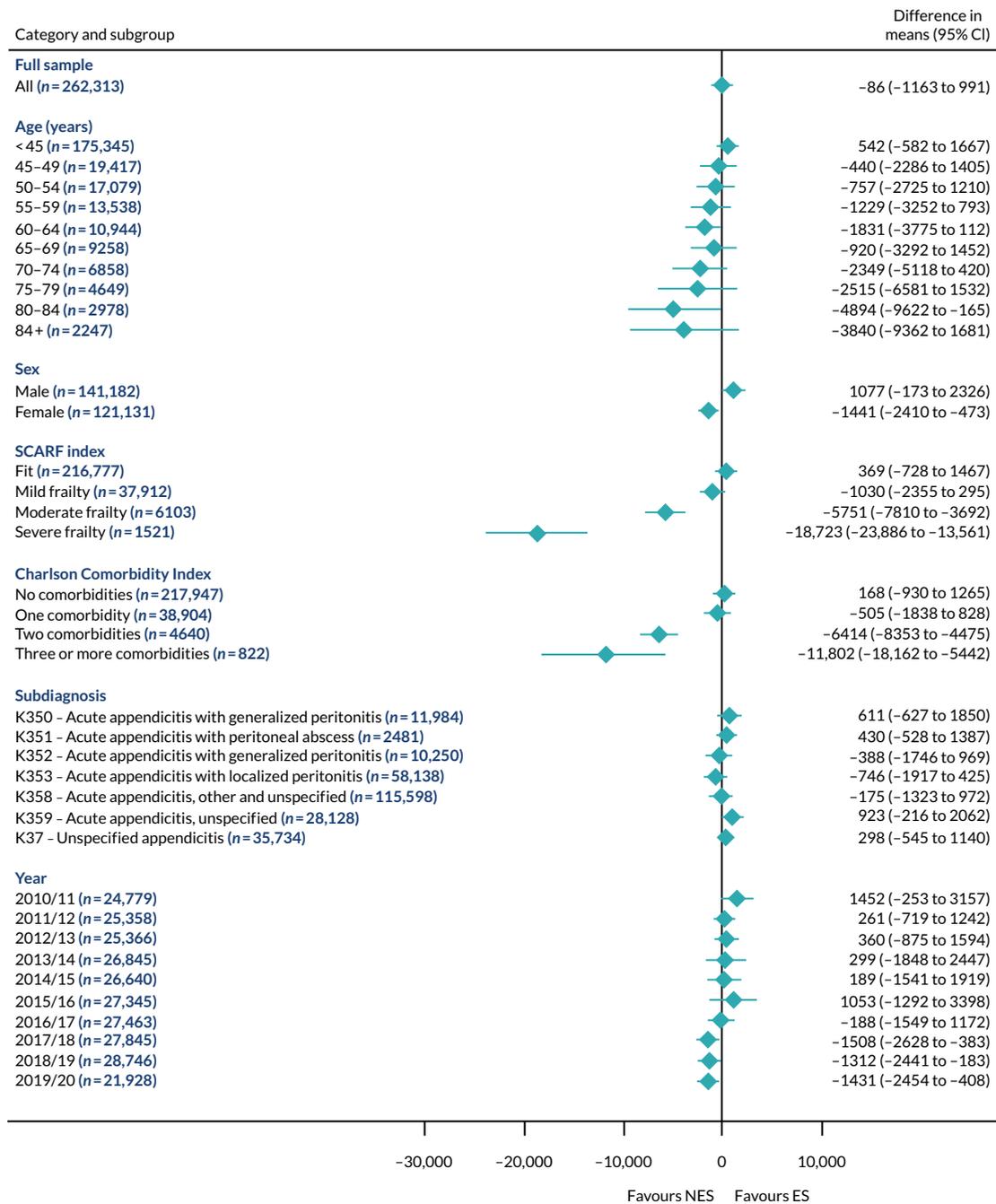
FIGURE 18 Estimated INBs (£) of ES vs. NES strategies for appendicitis subgroups.

respective mean INBs for the subgroups being −£18,700 (95% CI −£23,900 to −£13,600), −£7700 (95% CI −£13,000 to −£2370), −£16,600 (95% CI −£21,100 to −£12,000) and −£19,300 (95% CI −£25,600 to −£13,000). The results for diverticular disease suggested that NES strategies were also cost-effective for this subgroup, albeit with greater uncertainty around the point estimate (−£9230, −£24,300 to £5860). NES strategies were more cost-effective than ES among patients with moderate frailty, with a mean INB estimates of −£5750 (95% CI −£7810 to −£3690) for appendicitis, −£5630 (95% CI −£8150 to −£3110) for a hernia and−£9320 (95% CI −£14,900 to −£3740) for an intestinal obstruction. For patients who were otherwise 'fit' at baseline, ES was more cost-effective, with mean INB estimates of £718 (95% CI £294 to £1140) for cholelithiasis, £5180 (95% CI £684 to £9680) for diverticular disease, £2041 (95% CI £996 to £3090) for a hernia and £7850 (95% CI £5020 to £10,700) for an intestinal obstruction.
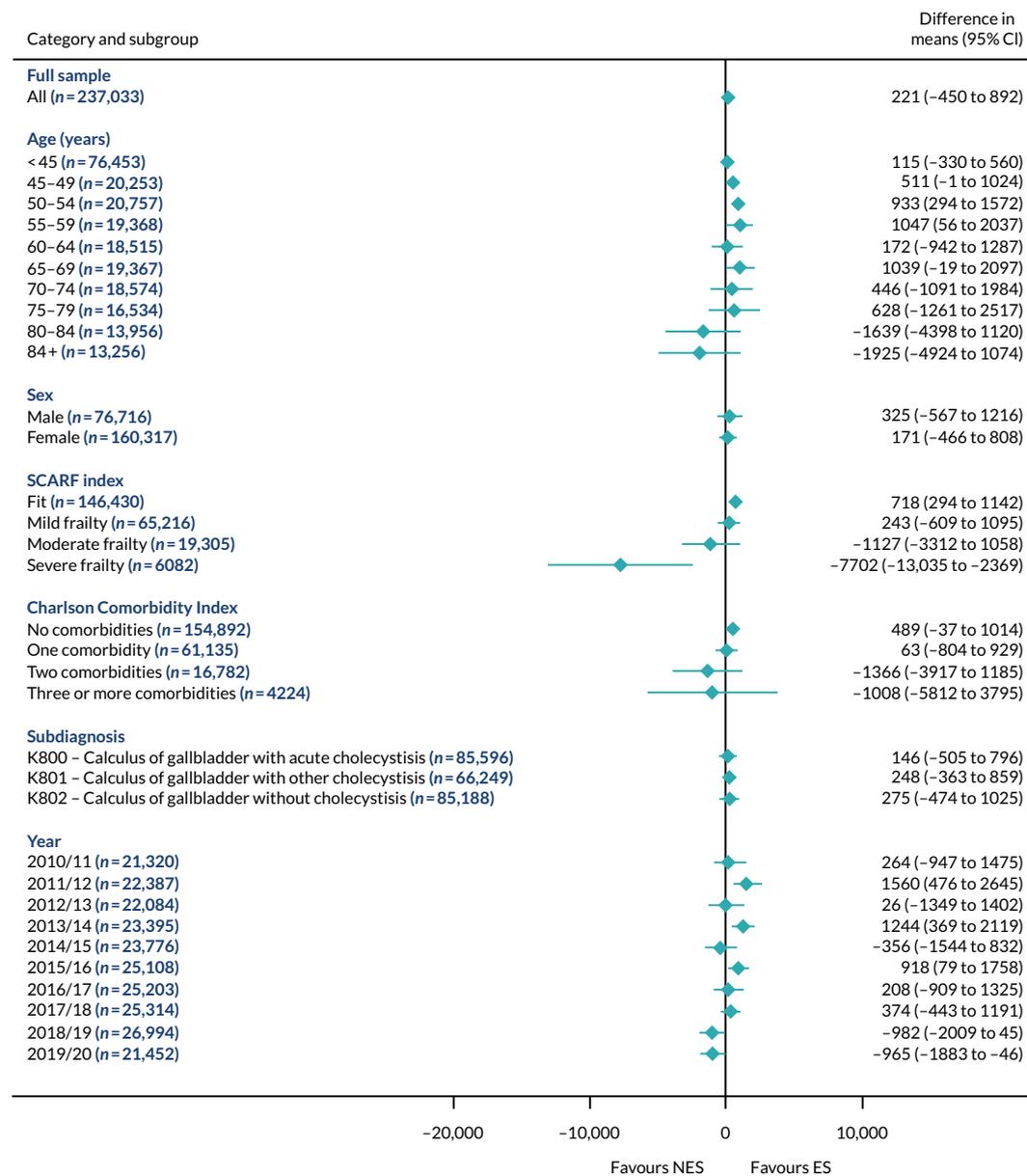
FIGURE 19 Estimated INBs (£) of ES vs. NES strategies for cholelithiasis subgroups.

Figures 23–27 report the individual-level estimates of incremental costs and QALYs for the five conditions. Here, for illustration, the results are stratified by frailty level. For patients with severe frailty, the proportions for whom ES is estimated to be cost-effective are 0.0657% (appendicitis), 0.443% (cholelithiasis), 46.9% (diverticular disease), 0.00% (hernia) and 7.52% (intestinal obstruction). For patients who were fit, the corresponding proportions are 59.0% (appendicitis), 92.5% (cholelithiasis), 87.1% (diverticular disease), 82.0% (hernia) and 61.6% (intestinal obstruction).

Sensitivity analyses

The overall results were robust to alternative assumptions (*Table 18*), including alternative proxies for the hospital's quality of acute care (i.e. SA1), higher or lower unit costs (i.e. SAs 2 and 3, respectively) and the use of linear interpolation for calculating QALYs (i.e. SA4).

For the scenario in which the time horizon was extended to 5 years (i.e. SA5), the CIs surrounding the INB estimates were particularly wide and, like the base case, included zero. The extension to a 5-year time horizon compared with the base case resulted in estimates of the mean INB that were positive,

Category and subgroup		Difference in means (95% CI)
Full sample		
All (*n* = 137,028)		2664 (–4298 to 9626)
Age (years)		
< 45 (*n* = 15,859)		–887 (–7640 to 5867)
45–49 (*n* = 11,324)		–74 (–6949 to 6801)
50–54 (*n* = 13,812)		–749 (–9764 to 8265)
55–59 (*n* = 13,844)		2568 (–4361 to 9497)
60–64 (*n* = 13,483)		2974 (–4005 to 9953)
65–69 (*n* = 14,108)		4653 (–1810 to 11,116)
70–74 (*n* = 14,446)		5710 (–895 to 12,315)
75–79 (*n* = 13,915)		6269 (–1978 to 14,516)
80–84 (*n* = 12,733)		3997 (–5422 to 13,416)
84+ (*n* = 13,504)		2101 (–7939 to 12,141)
Sex		
Male (*n* = 56,196)		3001 (–2946 to 8947)
Female (*n* = 80,832)		2430 (–5328 to 10,188)
SCARF index		
Fit (*n* = 71,036)		5180 (684 to 9676)
Mild frailty (*n* = 43,942)		2649 (–5692 to 10,991)
Moderate frailty (*n* = 15,970)		–3959 (–16,427 to 8509)
Severe frailty (*n* = 6080)		–9230 (–24,323 to 5863)
Charlson Comorbidity Index		
No comorbidities (*n* = 82,115)		1143 (–5457 to 7742)
One comorbidity (*n* = 39,067)		3904 (–3704 to 11,511)
Two comorbidities (*n* = 12,526)		6379 (–2163 to 14,921)
Three or more comorbidities (*n* = 3320)		11,684 (1864 to 21,503)
Subdiagnosis		
K572 – Diverticular disease of large intestine with perforation and abscess (*n* = 32,207)		–4754 (–9047 to –460)
K573 – Diverticular disease of large intestine without perforation or abscess (*n* = 104,821)		4943 (–2888 to 12,774)
Year		
2010/11 (*n* = 9678)		8325 (3011 to 13,639)
2011/12 (*n* = 10,856)		5541 (–944 to 12,027)
2012/13 (*n* = 11,473)		5818 (63 to 11,573)
2013/14 (*n* = 12,531)		4287 (–2684 to 11,257)
2014/15 (*n* = 13,551)		3233 (–3482 to 9949)
2015/16 (*n* = 14,546)		906 (–7264 to 9076)
2016/17 (*n* = 15,354)		1106 (–7099 to 9311)
2017/18 (*n* = 16,223)		1216 (–6615 to 9047)
2018/19 (*n* = 18,262)		519 (–7389 to 8426)
2019/20 (*n* = 14,554)		46 (–8538 to 8630)

–30,000 –20,000 –10,000 0 10,000 20,000 30,000

Favours NES   Favours ES

FIGURE 20 Estimated INBs (£) of ES vs. NES strategies for diverticular disease subgroups.

and of a larger magnitude for patients with cholelithiasis (£792 vs. £221), a hernia (£700 vs. –£119) or an intestinal obstruction (–£1502 vs. £2664). In the scenario with a 5-year time horizon, the INB estimates were negative for patients with appendicitis [–£3786 vs. –£86 (base case)] and diverticular disease [–£1502 vs. £2664 (base case)]. For patients with cholelithiasis, a hernia or an intestinal obstruction, the positive INB after 5 years was driven by the lower proportion of patients who had any re-admission following ES and NES strategies in years 2–5, after case mix adjustment (*Table 19*).

For patients with diverticular disease, the negative INB was driven by the higher adjusted re-admission rates, the proportions of patients having operative procedures and the days in hospital for the ES and NES strategy groups, during years 2–5, following the index admission (see *Table 19*). For patients with acute appendicitis, the differences in adjusted resource use between the comparison groups was small for years 2–5. Hence, the negative estimates of the average INB for patients with acute appendicitis may relate to the reduced sample included (i.e. around 50%).

Finally, in interpreting these findings, it should be recognised that the IV assumption that the TTO does not affect outcome except through the treatment received may be less tenable over a 5-year time period, as there may be other longer-term differences between the comparison groups in the care received that are not directly related to ES.
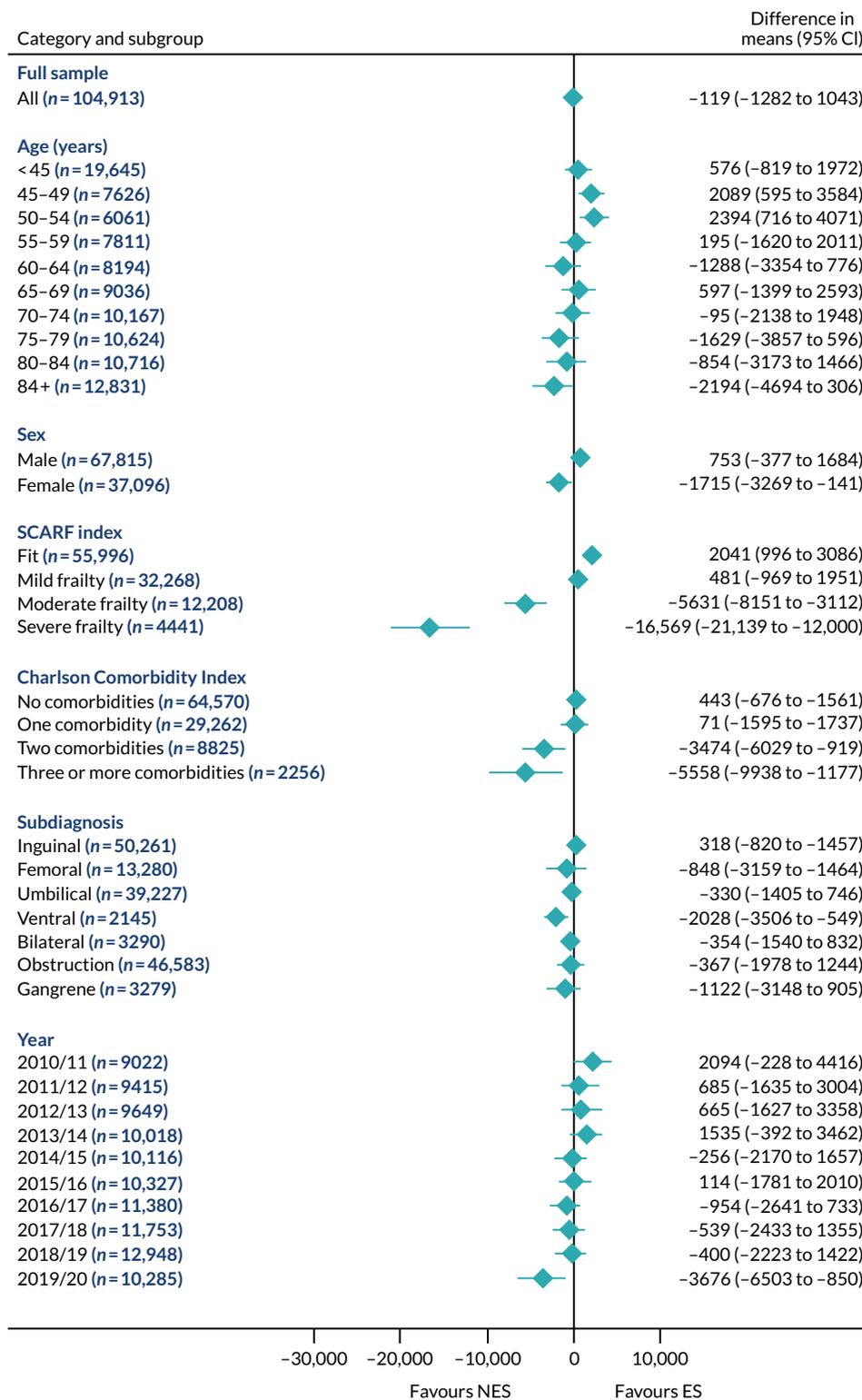
FIGURE 21 Estimated INBs (£) of ES vs. NES strategies for hernia subgroups.

The final SA that used risk adjustment (GLM regression) approaches, assuming that all confounding factors have been observed (i.e. SA6), reported similar results to the base case for the appendicitis, cholelithiasis and hernia cohorts. For patients with diverticular disease or an intestinal obstruction, the negative INBs for ES following risk adjustment may reflect that patients who had worse prognosis according to unobserved characteristics were more likely to receive ES than NES strategies.

Category and subgroup		Difference in means (95% CI)
Full sample		
All (*n* = 130,921)		728 (–2161 to 3617)
Age (years)		
< 45 (*n* = 14,476)		–486 (–4331 to 3394)
45–49 (*n* = 6401)		4078 (–652 to 8607)
50–54 (*n* = 8052)		8493 (450 to 16,538)
55–59 (*n* = 9067)		5437 (–2502 to 13,376)
60–64 (*n* = 10,868)		3935 (–575 to 8446)
65–69 (*n* = 13,678)		3738 (–485 to 7958)
70–74 (*n* = 16,591)		1633 (–1994 to 5859)
75–79 (*n* = 16,568)		534 (–4761 to 5629)
80–84 (*n* = 16,183)		–2378 (–6258 to 1503)
84+ (*n* = 19,997)		–6733 (–11,348 to –2118)
Sex		
Male (*n* = 62,035)		–1069 (–4529 to 2391)
Female (*n* = 68,696)		2348 (–585 to 5278)
SCARF index		
Fit (*n* = 59,168)		7650 (5018 to 10,683)
Mild frailty (*n* = 43,401)		–439 (–3537 to 2860)
Moderate frailty (*n* = 19,738)		–9314 (–14,889 to –3740)
Severe frailty (*n* = 8614)		–19,306 (–25,637 to –12,974)
Charlson Comorbidity Index		
No comorbidities (*n* = 68,426)		–605 (–3711 to 2501)
One comorbidity (*n* = 41,505)		1501 (–1791 to 4793)
Two comorbidities (*n* = 16,424)		2261 (–2227 to 6749)
Three or more comorbidities (*n* = 4566)		8159 (–1662 to 17,981)
Subdiagnosis		
K561 – Intussusception (*n* = 1435)		2030 (–1087 to 5147)
K562 – Volvulus (*n* = 15,843)		–1456 (–4601 to 1688)
K563 – Gallstone ileus (*n* = 1996)		–3004 (–6257 to 250)
K565 – Intestinal adhesions (bands) with obstruction (*n* = 45,287)		1100 (–1712 to 3912)
K566 – Other and unspecified intestinal obstruction (*n* = 66,353)		1080 (–1865 to 4024)
C18 – Malignant neoplasm of colon (*n* = 3473)		6423 (2914 to 9931)
C19 – Malignant neoplasm of rectosigmoid junction (*n* = 319)		5489 (1716 to 9261)
C20 – Malignant neoplasm of rectum (*n* = 558)		8257 (4803 to 11,911)
Year		
2010/11 (*n* = 10,844)		377 (–3809 to 4562)
2011/12 (*n* = 11,713)		–120 (–4167 to 3947)
2012/13 (*n* = 11,933)		113 (–3860 to 4105)
2013/14 (*n* = 12,722)		1781 (–2356 to 5919)
2014/15 (*n* = 13,263)		3069 (–912 to 7050)
2015/16 (*n* = 13,726)		126 (–3943 to 4195)
2016/17 (*n* = 14,058)		1739 (–1627 to 5104)
2017/18 (*n* = 14,680)		–786 (–4288 to 2717)
2018/19 (*n* = 15,706)		4221 (–1912 to 10.355)
2019/20 (*n* = 12,296)		–4313 (–8425 to –201)

–40,000 –30,000 –20,000 –10,000 0 10,000 20,000 30,000

Favours NES   Favours ES

FIGURE 22 Estimated INBs (£) of ES vs. NES strategies for intestinal obstruction subgroups.

Discussion

This CEA finds that for patients presenting as emergency admissions with the five acute conditions, neither ES nor NES strategies were more cost-effective after applying an IV approach to address confounding. For all five conditions, the mean INBs were surrounded by 95% CIs that included zero. For patients with an abdominal wall hernia, ES led to a small increase in mean QALYs and in mean costs at 1 year. For the other four conditions, the mean costs and mean QALYs were similar for ES and NES strategies. The finding that any overall difference between the comparison groups in average cost
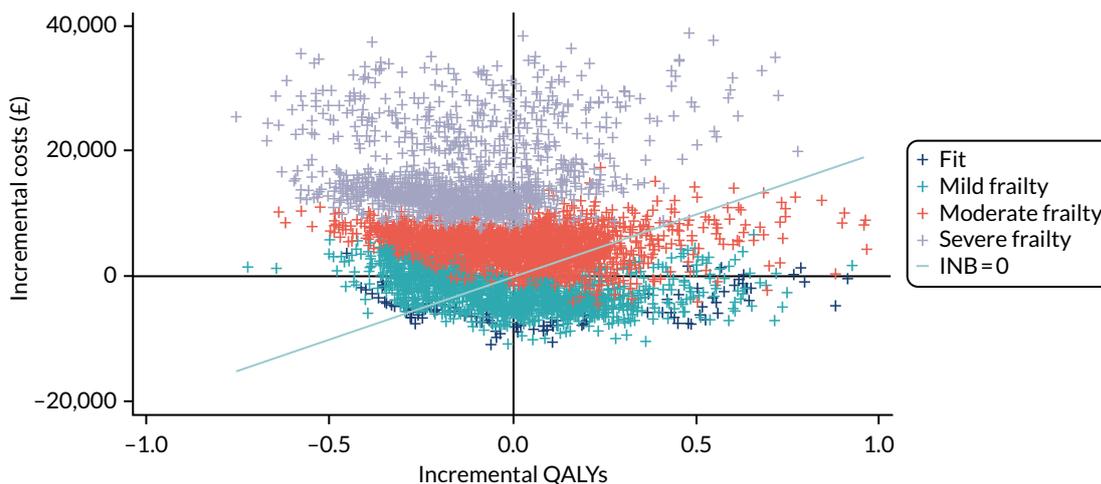
FIGURE 23 Cost-effectiveness plane of person-centred treatment effects for costs (£) and QALYs for appendicitis.
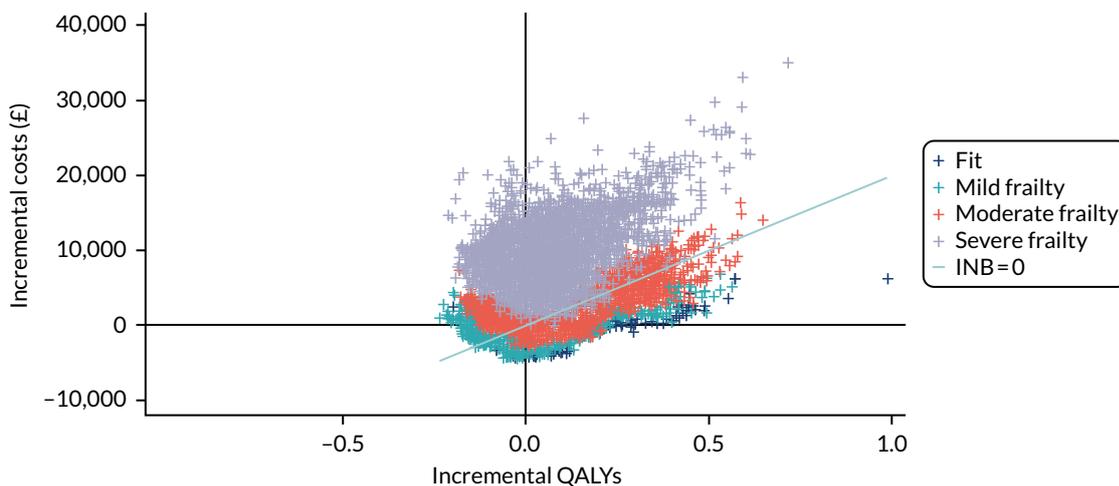
FIGURE 24 Cost-effectiveness plane of person-centred treatment effects for costs (£) and QALYs for cholelithiasis.
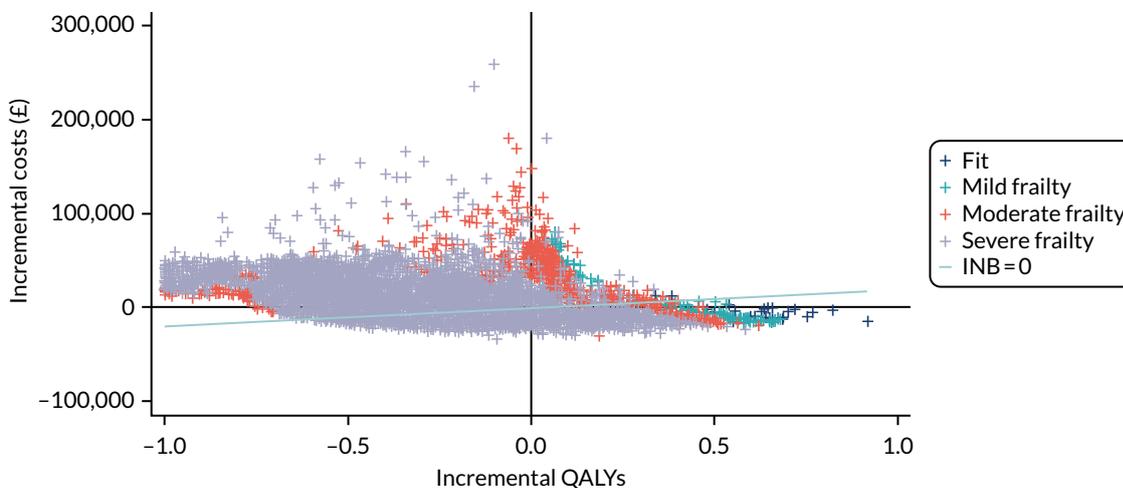
FIGURE 25 Cost-effectiveness plane of person-centred treatment effects for costs (£) and QALYs for diverticular disease.
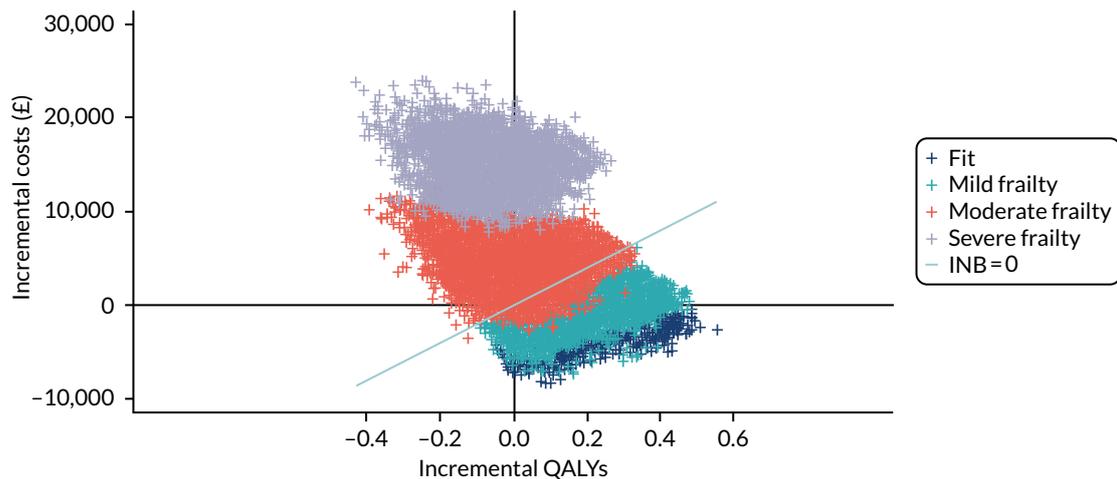
FIGURE 26 Cost-effectiveness plane of person-centred treatment effects for costs (£) and QALYs for hernia.
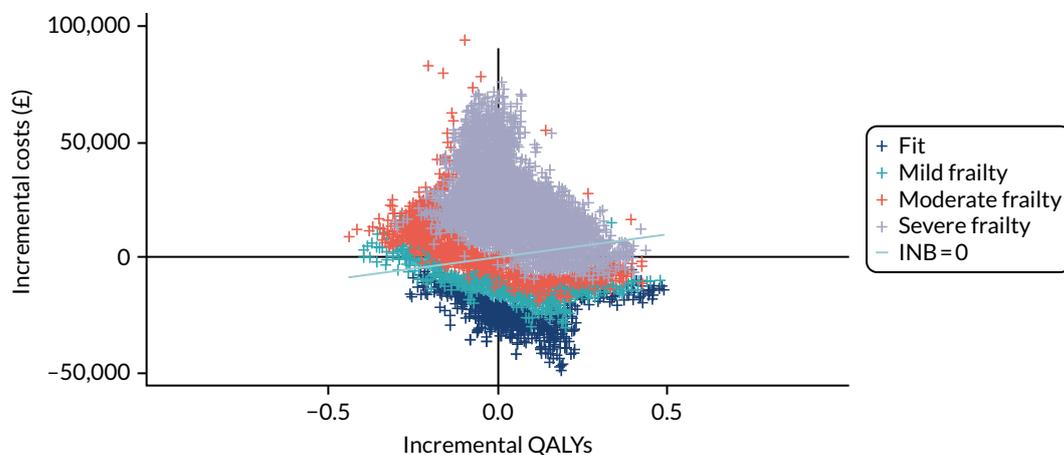
FIGURE 27 Cost-effectiveness plane of person-centred treatment effects for costs (£) and QALYs for intestinal obstruction.

was moderate and uncertain reflected that, although the mean index admission costs were higher for the ES groups than for the NES groups, the proportions of patients who were re-admitted to hospital before 1 year were higher in the NES groups than in the ES groups (see *Table 16*).

Beneath these overall findings, the CEA identified differences in relative cost-effectiveness according to patient subgroup. For patients with severe frailty, the mean costs were higher (see *Appendix 6*) and the average QALYs lower (see *Appendix 6*) and, hence, the estimated INBs were negative following ES, compared with NES strategies (*Table 20*). Conversely, for subgroups of patients who were 'fit' or without comorbidities, the ES strategies tended to be more cost-effective (see *Table 20*). The subgroup findings reflected that the higher risks of all-cause mortality were in the ES group for patients with severe frailty and more comorbidities (apart from for patients with diverticular disease). Conversely, ES strategies were generally more cost-effective for patients with lower baseline risks of death, such as patients who were 'fit' and without comorbidities.

This paper contributes to the limited previous literature[7,8,24,25,65–67] evaluating the cost-effectiveness of ES for these common acute gastrointestinal conditions. For patients with acute appendicitis, some previous studies[8,21] have suggested that NES strategies can reduce average costs. The studies[8,21] reported that around 20% of patients in the NES arm had appendicectomies by 90 days[21] and 1 year,[8,23] but that the additional costs of the NES strategy were insufficient to offer the additional costs of ES within the

TABLE 18 Summary of SA results: overall INB (£ 2019/20) of ES vs. NES strategies

Analysis	Description	INB, mean (95% CI)				
		Appendicitis (n = 268,144)	Cholelithiasis (n = 240,977)	Diverticular disease (n = 138,869)	Hernia (n = 106,432)	Intestinal obstruction (n = 133,073)
Base case	See *Chapter 2, Methods*	−86.2 (−1163 to 991)	221 (−450 to 892)	2664 (−4298 to 9626)	−119 (−1282 to 1043)	728 (−2161 to 3616)
SA1	Considered alternative measures of hospital quality derived from the 2016–2018 NELA reports (see *Sensitivity analyses*)	408 (−787 to 1605)	664 (47.1 to 1281)	5823 (1029 to 10,617)	125 (−1027 to 1276)	316 (−2901 to 3532)
SA2	Considered a 10% decrease in all unit costs in total cost calculation (see *Sensitivity analyses*)	−96.9 (−1078 to 884)	213 (−343 to 769)	2491 (−3674 to 8655)	−30.4 (−964 to 903)	674 (−2032 to 3381)
SA3	Considered a 10% increase in all unit costs in total cost calculation (see *Sensitivity analyses*)	−75.2 (−1257 to 1107)	228 (−436 to 893)	2836 (−4357 to 10,029)	−208 (−1254 to 837)	782 (−2518 to 4082)
SA4	Used linear interpolation between baseline and 1-year HRQoL end points for calculating QALYs (see *Sensitivity analyses*)	−202 (−1514 to 1110)	221 (−364 to 807)	2796 (−2796 to 8389)	−125 (−1216 to 967)	1130 (−2349 to 4609)
SA5	Evaluated costs and effects of ES and NES over a 5-year time horizon (see *Sensitivity analyses*)	−3786 (−9113 to 1541)	792 (−2761 to 4345)	−1502 (−27,066 to 24,062)	700 (−6812 to 8212)	5399 (−633 to 11,430)
SA6	Risk adjustment using GLM regression (see *Sensitivity analyses*)	−165 (−287 to −42)	−190 (−290 to −90.7)	−12,381 (−12,848 to −12,058)	−50.1 (−241 to 141)	−3991 (−4242 to −3740)

index admission. The ESORT study results differed in that almost all of the patients in the ES group would have had appendicectomy by laparoscopic, rather than by open repair, and that the broad inclusion criteria would have included patients with complicated as well as uncomplicated appendicitis and patients with suspected appendicitis who did not have clinically or radiologically confirmed diagnoses. In the ESORT study, the NES group had rates of re-admission and reintervention that led to sufficiently higher costs to offset the additional index admission costs of the ES strategy (see *Table 16*). The differences between the ESORT study and these previous cost results[8,21] may reflect the differences in the study designs and that, unlike the 90-day results for the COVID-Harem study,[21] the ESORT study follow-up extended to 1 year (base case) and 5 years (sensitivity analyses) and, therefore, included higher (adjusted) rates of re-admission and reintervention in the NES group, than in the ES group, beyond 90 days.

TABLE 19 Estimated incremental effects of ES vs. NES strategies on main resource measures for a 5-year time horizon (i.e. SA5)

Resource use item	Year				
	1	2	3	4	5
Acute appendicitis (n = 125,670)					
Per cent with re-admission	−7.5	−1.8	−0.1	−0.4	−3.6
Per cent with critical care admission (within any re-admission)	0.5	0.2	0.1	0.1	0.1
Per cent with costed operative procedure[a] (within any re-admission)	1.9	0.2	0.4	0.1	0.2
Mean days in hospital during each year (all admissions)	0.7	0.0	0.2	0.2	0.1
Cholelithiasis (n = 110,732)					
Per cent with re-admission	−36.8	−1.4	−1.0	−1.6	1.2
Per cent with critical care admission (within any re-admission)	−1.7	−0.2	0.0	−0.2	0.5
Per cent with costed operative procedure[a] (within any re-admission)	−47.7	0.5	−0.1	−0.4	1.0
Mean days in hospital during each year (all admissions)	−0.5	0.2	0.6	0.0	0.6
Diverticular disease (n = 59,945)					
Per cent with re-admission	−33.6	9.5	33.9	−10.1	27.1
Per cent with critical care admission (within any re-admission)	9.9	1.5	14.2	2.0	7.8
Per cent with costed operative procedure[a] (within any re-admission)	1.0	8.1	17.0	4.5	−1.4
Mean days in hospital during each year (all admissions)	7.9	4.8	0.8	3.7	11.2
Abdominal wall hernia (n = 47,580)					
Per cent with re-admission	−39.6	−8.3	0.6	−3.0	−7.2
Per cent with critical care admission (within any re-admission)	−4.7	0.7	−3.5	−2.8	−0.5
Per cent with costed operative procedure[a] (within any re-admission)	−55.1	0.3	0.3	0.7	0.8
Mean days in hospital during each year (all admissions)	1.9	−0.1	0.2	−0.2	−0.7
Intestinal obstruction (n = 64,170)					
Per cent with re-admission	−23.1	−9.7	−8.0	−14.2	−17.2
Per cent with critical care admission (within any re-admission)	−0.7	−0.3	−1.6	−0.8	1.7
Per cent with costed operative procedure[a] (within any re-admission)	12.0	−1.6	−1.0	−1.2	0.1
Mean days in hospital during each year (all admissions)	−5.0	−1.8	−3.0	−3.1	−2.7

a Percentage of patients who received a procedure with a prevalence that exceeded 1% volume in the index admission (see *Resource use measurement*).

Note
All results are changes in percentage points unless stated.

For patients with acute cholelithiasis, the ESORT study differed from CEA based on decision models that found that 'early' versus 'delayed' cholecystectomy was less costly and more effective[25,93] and, therefore, more cost-effective.[26] Our results also differed from a trial-based CEA, which found that delayed cholecystectomy was more cost-effective than early cholecystectomy.[27] The distinguishing features of the ESORT study are that it had broader eligibility criteria than the RCT or cohort studies underlying the previous CEA and it considered re-admissions and reinterventions for up to 1 year (rather than 30 days). In addition, unlike the previous cohort studies, the ESORT study used an

TABLE 20 Estimated INB (£ 2019/20) of ES vs. NES, overall and by subgroups

Group	Estimated INB (£), mean (95% CI)				
	Appendicitis (n = 268,144)	Cholelithiasis (n = 240,977)	Diverticular disease (n = 138,869)	Hernia (n = 106,432)	Intestinal obstruction (n = 133,073)
Overall	−86.2 (−1160 to 991)	221 (−450 to 892)	2660 (−4300 to 9630)	−119 (−1280 to 1040)	728 (−2160 to 3620)
Age group (years)					
< 45	0542 (−582 to 1670)	114 (−330 to 560)	−887 (−7640 to 5870)	576 (−819 to 1970)	−468 (−4330 to 339)
45–49	−440 (−2280 to 1410)	511 (−1.20 to 1020)	−73.9 (−6950 to 6800)	2090 (595 to 3580)	4080 (−652 to 8810)
50–54	−757 (−2720 to 1210)	933 (294 to 1570)	−749 (−9760 to 8260)	2390 (716 to 4070)	8490 (450 to 16,535)
55–59	−1230 (−3250 to 793)	1050 (56.2 to 2040)	2568 (−4360 to 9500)	0195 (−1620 to 2010)	5440 (−2500 to 13,400)
60–64	−1830 (−3770 to 112)	172 (−942 to 1290)	2970 (−4000 to 9950)	−1290 (−3350 to 778)	3940 (−575 to 8450)
65–69	−920 (−3290 to 1450)	1040 (−19.3 to 2100)	4650 (−1810 to 11,100)	597 (−1400 to 2590)	3740 (−485 to 7960)
70–74	−2350 (−5120 to 420)	446 (−1090 to 1980)	5710 (−895 to 12,320)	−94.8 (−2140 to 1950)	1830 (−1990 to 5660)
75–79	−2510 (−6560 to 1530)	628 (−1260 to 2520)	6270 (−1980 to 14,500)	−1630 (−3860 to 598)	534 (−4760 to 5830)
80–84	−4890 (−9610 to −165)	−1640 (−4400 to 1120)	4000 (−5420 to 13,400)	−854 (−3170 to 1470)	−2380 (−6260 to 1500)
> 84	−3840 (−9360 to 1680)	−1920 (−4920 to 1070)	2100 (−7940 to 12,100)	−2190 (−4690 to 306)	−6730 (−11,300 to −2120)
Frailty level					
Fit	369 (−728 to 1470)	718 (294 to 1140)	5180 (684 to 9680)	2040 (996 to 3090)	7850 (5020 to 10,700)
Mild frailty	−1030 (−2360 to 295)	243 (−609 to 1090)	2650 (−5690 to 11,000)	481 (−989 to 1950)	−439 (−3540 to 2660)
Moderate frailty	−5750 (−7810 to −3690)	−1130 (−3310 to 1060)	−3960 (−16,400 to 8510)	−5630 (−8150 to −3110)	−9320 (−14,900 to −3740)
Severe frailty	−18,700 (−23,900 to −13,600)	−7700 (−13,000 to −2370)	−9230 (−24,300 to 5860)	−16,600 (−21,100 to −12,000)	−19,300 (−25,600 to −13,000)
Number of comorbidities					
Zero	168 (−930 to 1270)	489 (−36.6 to 1010)	1140 (−5460 to 7740)	443 (−676 to 1560)	−605 (−3710 to 2500)
One	−505 (−1840 to 828)	62.6 (−804 to 929)	3900 (−3700 to 11,500)	71.3 (−1590 to 1740)	1500 (−1790 to 4790)
Two	−6410 (−8350 to −4480)	−1370 (−3920 to 1190)	6380 (−2160 to 14,900)	−3470 (−6030 to −919)	2260 (−2230 to 6750)
Three or more	−11,800 (−18,200 to −5440)	−1000 (−5810 to 3800)	11,700 (1860 to 21,500)	−5560 (−9940 to −1180)	8160 (−1660 to 18,000)

Colours denote statistical significance at 5% level. Dark purple cells indicate that the LIV method favours ES (i.e. a INB > 0) and blue cells indicate that the LIV method favours NES (i.e. a INB < 0).

advanced quantitative approach (i.e. IV estimation) to balance confounding factors between the groups. A fundamental difference between the ESORT study and the preceding CEA was that previous studies defined the comparator group as patients having 'delayed surgery' who may have been differed in prognosis from patients who had 'early surgery'. By contrast, in the ESORT study, the ES and NES groups were defined at the same point in time (when they were seen by a surgeon), and the IV approach aimed to balance the patient characteristics at that time point. A key reason why the ESORT study found that average costs were similar between the comparison groups at 1 year is that, although the NES strategy arm had higher rates of re-admission and (re)intervention after the index admission, the differences in adjusted rates [53.4 vs. 83.3 for re-admission, 19.2 vs. 63.2 for (re)intervention] were relatively small.

For patients with acute diverticular disease or an abdominal wall hernia, a previous CEA of surgery compared with watchful waiting has been undertaken in the elective setting;[94,95] however, there is no precedent CEA of direct relevance to compare with the ESORT study. For patients with an intestinal obstruction, the ESORT study included a broad range of patients, including patients admitted with cancer diagnoses and patients with adhesions from previous surgery. Previous CEAs have found NES (i.e. stenting) to be cost-effective for patients with an incurable large bowel obstruction[85] and ES to be cost-effective for patients with an adhesive small bowel obstruction.[96]

The ESORT study has several strengths. First, the CEA extended the previously validated IV approach,[20] used in the assessment of clinical effectiveness (see *Chapter 3*), to the CEA context. The study exemplified how an IV approach can be applied to electronic health record data to address confounding and heterogeneity in providing cost-effectiveness estimates overall, and according to patient subgroups. The ESORT study, therefore, addresses a gap in the applied methods literature. (Note that a review[97] of CEA, which used observational data, highlighted that < 5% of studies used IV approaches.) In addition, the LIV approach fully recognises heterogeneity, but has not previously been applied to CEA. In extending the LIV approach to the CEA context, it was necessary to allow for individual-level correlation between costs and outcomes. It was also important to apply appropriate costs models for handling the highly right-skewed nature of the cost data, notably for conditions such as diverticular disease and intestinal obstruction, which included patients with extremely high 1-year costs (i.e. costs in excess of £100,000). The LIV approach provided cost-effectiveness estimates according to the subgroups of interest. For the appendicitis, cholelithiasis and hernia cohorts, the LIV approach (base case) provided similar cost-effectiveness estimates to estimates from regression approaches that assumed all confounding factors were observed (i.e. SA6) (see *Table 18*). For patients with acute diverticular disease or an intestinal obstruction, the regression analysis provided quite different results to the LIV method. The most plausible explanation for the difference is that for acute diverticular disease and intestinal obstruction the choice of ES or NES strategy was informed by baseline prognostic factors that are not available within the HES data, for example the patient's disease stage. Hence, it seems plausible that patients with these acute conditions who were of worse prognosis according to unobserved baseline measures were more likely to receive ES than NES. Although the IV approach was able to ensure that these factors were balanced between the comparison groups, the regression analysis was subject to residual confounding in favour of the NES strategies group.

Second, the application of IV methods to CEA raised the danger of imprecise estimates. Although the overall CEA estimates were, indeed, uncertain, the large representative samples of patients enabled the study to report sufficiently precise subgroup results. The ESORT study, therefore, demonstrates the value of applying IV approaches to CEA, with concerns addressed at the study design stage by undertaking sample size calculations using pilot data as part of prespecified analysis plans (see statistical analysis plan on project webpage[30]) and by including sufficient hospitals (*n* = 175) and time periods (10 years) to provide estimates with sufficient precision at subgroup level.

Third, the HES data, although having common features of electronic health record data (notably the potential for confounding and heterogeneity), were of generally high quality, with baseline covariates,

all-cause mortality and resource use data available for ≈95% of patients. Economic evaluations that use routine data face inevitable challenges, above and beyond those that face comparative effectiveness studies. CEAs are required to offer a comprehensive description of the comparison groups. For both groups, the information available on defining the strategies was limited to that available from the OPCS-4 procedure codes. The information available from OPCS-4 procedure codes was sufficient to categorise patients into the ES and NES strategy groups, and to define use of specific operative procedures in both comparison groups, but OPCS-4 procedure codes did not provide full information about non-operative strategies, for example use of antibiotics for patients with uncomplicated acute appendicitis. It is possible that the HES data were subject to coding errors that were then incorporated into the estimates of cost and cost-effectiveness, although previous research found that costs estimated from HES data were very similar to those derived from medical records.[98]

The base-case analysis took a 1-year time horizon. The 1-year time horizon was judged to be sufficient to capture important differences between the comparator groups, such as differences in all-cause mortality or re-admissions. The advantage of this time horizon is that it enabled all patients to be included in the analysis without censoring or making assumptions about future costs and outcomes beyond the observed data. The sensitivity analyses found that the overall results were similar if the time horizon was extended to 5 years. For patients with appendicitis or diverticular disease, the SA suggested that the point estimates for the INB were somewhat different with the 5- year time horizon compared with 1-year time horizon (see *Table 18*), but this SA should be interpreted with caution. The 5-year results were highly uncertain, with 95% CIs that overlaped the base-case estimates, and reflected the much-reduced sample sizes. The extended time horizon raised additional concerns for the IV approach, in that the underlying assumption that the hospital's TTO had no direct effect on the end points, beyond that through the comparison groups, was less tenable over the longer time horizon.

The CEA did not take the recommended perspective of agencies such as NICE, which would have required consideration of all costs from the health and personal social services perspective. The ESORT study used resource data from the HES inpatient database, and only the ensuing inpatient costs were included in the analysis. The inclusion of this narrow range of costs raises the concern that for some patient groups the total costs per patient were underestimated, for example those patients who had severe levels of frailty might have had high use of NHS-funded social care. However, the rationale for only including inpatient costs was that differences between ES and NES strategies in inpatient costs were anticipated to be the main driver of the incremental costs. Although, the ESORT study could have also used linked data on HES outpatient data, these data are of poor quality and are likely to be inaccurate. Similarly, although information could have been used on patients' discharge destination (e.g. care home), patients' discharge information was judged as being unreliable and its use would have required further untestable assumptions. An alternative approach would have been to survey samples of patients and their carers, and ask them about their use of health and personal social services outside the acute hospital. However, this approach would have added to the study's burden and duration, and is prone to problems of missing data and recall bias. Previous CEAs of ES compared with NES strategies for patients with acute appendicitis,[21] and of alternative forms of ES for patients with ruptured aortic aneurisms,[77] reported that the main drivers of relative costs of the alternative strategies are inpatient costs.

Finally, in common with any approach to address confounding, the implementation of the LIV methods made assumptions; in particular, that the relationships of the covariates and the IV with both the treatment receipt and the outcomes were correctly specified. Here, more flexible data-adaptive approaches may be helpful, although these approaches have not yet been extended to this context. A further consideration is that the subgroup analyses presented here represent the average estimated effect for individuals within the group, rather than the causal effect of group membership per se. Although the subgroups used here were prespecified within a statistical analysis plan, in other contexts spurious subgroup effects may be obtained by '*p*-hacking'.[99]

Chapter 5 Discussion

Overview

The ESORT study evaluated the clinical effectiveness and the cost-effectiveness of ES compared with NES strategies [e.g. medical management, non-surgical procedures (e.g. radiological-guided drainage of abscess), surgery deferred to the elective (planned) setting]. The ESORT study considered patients presenting as emergency admissions to hospital with suspected acute appendicitis, cholelithiasis, diverticular disease, an adonomial wall hernia or an intestinal obstruction. The study used HES data to define the study cohorts and the comparator strategies. The ESORT study also used the same HES data linked to ONS mortality to derive the primary clinical outcome (i.e. DAOH at 90 days) and to calculate QALYs and hospitalisation costs at 1 year. The ESORT study reported the clinical effectiveness and cost-effectiveness of ES compared with NES with an IV approach that addressed confounding and recognised heterogeneity in the relative effectiveness and cost-effectiveness of ES according to patient subgroups.

The main findings were that, overall, the average numbers of DAOH at 90 days were similar following ES and NES strategies for each of the five cohorts of patients with acute gastrointestinal conditions requiring emergency admission. The CEA did not provide strong evidence that either strategy (i.e. ES or NES strategies) was more cost-effective overall across each of the five cohorts of patients.

The ESORT study found that for each of the five conditions the relative clinical effectiveness and cost-effectiveness of ES compared with NES strategies differed by patient subgroup, notably according to number of comorbidities (using the Charlson Comorbidity Index) and frailty level (using the SCARF index). For patients with acute appendicitis or and abdominal wall hernia who had two or more comorbidities, NES strategies led to increased average number of DAOH at 90 days, and was cost-effective for these two conditions and for patients with cholelithiasis. For patients with acute diverticular disease or an intestinal obstruction who had two or more comorbidities, on average, ES was cost-effective compared with NES strategies.

For patients with severe frailty at baseline, the NES strategies led to an increase in mean number of DAOH (for all conditions), a reduction in 90-day mortality (for all conditions apart from hernia), a reduction in LOS (for all conditions) and were cost-effective compared with the ES strategies. For patients who were 'fit' at presentation, on average, ES strategies were relatively cost-effective compared with NES strategies.

Public and patient involvement

The views of patients who had direct experience of admission to a hospital in an emergency informed the design of the study, the presentation of the findings and dissemination proposals. A key element of this was that, during the two 'design workshops', 14 PPI panellists endorsed two of the outcome measures: (1) mortality and (2) DAOH. As part of the two 'translation workshops', 12 PPI panellists carefully considered the *Plain English summary* and the ways in which the research should be presented, particularly in relation to the key findings regarding frailty (see *Report Supplementary Material 5*).

Our approach to 'patient and public' in the context of recruitment was to interpret it widely, and we invited patients, family members and those interested in supporting research to our PPI panels. This broad approach meant that we had a breadth of perspectives among our PPI panellists and advisors and, therefore, PPI voices were not simply individuals who had had particular positive or negative experiences of either ES or NES.

The COVID-19 pandemic necessitated flexibility in our approach to PPI, with the main change being that all panels were held virtually over Zoom. Briefing materials were provided in advance, mostly accessed on the Sway platform.[30] Many panellists commented favourably on this approach and praised the briefing materials.

Panellists also commented that the virtual format improved accessibility, as many panellists would have not been able to participate in an in-person meeting in London because of physical disabilities or caring commitments precluding their ability to travel. The benefits of the virtual format, coupled with close attention to ensure that any accessibility issues (e.g. sight impediments) were addressed, meant that we were able to attract a diverse range of individuals to the PPI panels. The diversity in age, sex and experience also helped reduce the potential for 'group-think' in our PPI panels and meant that panellists freely expressed different views and perspectives in their groups.

Findings in the context of related research

The ESORT study extends previous literature[2,5] that has raised concern about variations in ES rates. The ESORT study identified unexplained variations in ES rates across NHS acute hospitals for each of the five conditions, which remained after adjusting for differences in prognostic measures, such as age, sex, number of comorbidities and frailty level.[36] This variation was observed within clinically relevant definitions of the target populations of interest. The ESORT study design was informed by the concept of the target trial,[31,32] and used the consensus of a clinical panel to define population eligibility criteria and the range and timing of surgical procedures that constituted ES.

The ESORT study exploited this variation in ES rates across hospitals, within an IV design,[18] and compared patient outcomes from hospitals that differed in their TTO[20] to report the relative effectiveness of ES compared with NES strategies. The ESORT study, therefore, adds to the limited evidence on the relative clinical effectiveness[7,8,13,14,21,23,66,67] and cost-effectiveness of ES for common acute conditions presenting as emergency admissions. There are relatively few RCTs that have compared ES with NES strategies in the emergency setting[7,8,13,23] and, to the best of our knowledge, none that have considered relative clinical effectiveness and cost-effectiveness according to patient subgroups.

Previous RCTs[7,8] of ES compared with NES strategies for patients with acute appendicitis have reported mixed results, with some RCTs reporting better outcomes following ES[8] and others reporting NES strategies as more effective.[7] None of the RCTs included the broad range of patients, including patients with complications, who present with suspected acute appendicitis to NHS hospitals in England, and the CODA (Comparison of Outcomes of Antibiotic Drugs and Appendectomy) and APPAC (Appendicitis Acuta) trials considered open rather than laparoscopic appendicectomy,[7,23] which is the common surgical technique for appendicectomy in the NHS. It is, therefore, challenging to interpret the results of these previous studies for NHS decision-making.

The COVID-Harem study was an observational study that considered antibiotics as a first-line alternative to appendicectomy, which for about two-thirds of patients was by the laparoscopic route.[21] The COVID-Harem study had similarly broad inclusion criteria to the ESORT study, and included patients following the initial COVID-19 lockdown in the UK on 23 March 2020. At the time the COVID-Harem study was recruiting patients, there were concerns about viral transmission of COVID-19 via aerosolisation during laparoscopy, which led professional societies to recommend management with antibiotics (NES strategy) rather than ES for patients with uncomplicated acute appendicitis.[100] The COVID-Harem study found that the NES strategy led to fewer days in hospital and lower average costs, compared with ES, at 90 days.[101,102]

The differences between the findings of the ESORT and COVID-Harem studies may reflect the following. First, the COVID-Harem study used information from imaging to confirm diagnosis of acute appendicitis; however, such information was not available in the ESORT study and so the 'acute appendicitis cohort' included patients with 'suspected' rather than 'confirmed' appendicitis. In the COVID-Harem study, around 3% of patients had negative appendicectomies,[21] which is lower than rates reported by the RIFT (Right Iliac Fossa Treatment) study for routine clinical practice in the UK,[103] prior to the onset of the COVID-19 pandemic. The RIFT study reported negative appendicectomy rates of 28% for women and 12% for men, which are more representative of routine practice and likely to apply to the ESORT study, which included broad populations of patients with suspected acute appendicitis. The ESORT study found that for men ES was the most cost-effective strategy, whereas for women NES strategies were more cost-effective than ES, which may reflect the higher negative appendicectomy rate in routine practice. A second important difference between the ESORT and COVID-Harem studies is the duration of follow-up. The CEA for the ESORT study reported that the proportions of patients who had re-admissions for the NES and ES strategy groups were 53% and 27%, respectively, with 16% and 2.3%, respectively, having further operative procedures over the 1-year follow-up period. By contrast, the proportion of patients in the COVID-Harem study who had re-admissions following NES and ES were 37% and 20, respectively; however, this was over the shorter follow-up period of 90 days. The COVID-Harem study will report further results at 12 months. The APPAC trial[66,67] assessed 5- and 7-year outcomes and resource use for antibiotics compared with open appendicectomy for uncomplicated acute appendicitis. The APPAC trial found that the median total days in hospital at 5 years and HRQoL at 7 years were similar between the groups.[66,67]

For patients with acute diverticular disease, RCTs have not contrasted ES versus NES strategies, but rather different forms of ES, such as lavage versus resection.[14,15] There are no previous observational studies of ES versus NES strategies for patients with acute diverticular disease, although the ongoing DAMASCUS study plans to contrast outcomes for patents following ES versus NES strategies.[104]

For patients with acute cholecystitis, which was the most common subcondition (≈35%) within the more general diagnosis of 'cholelithiasis' used in the ESORT study cohort, meta-analyses of published RCTs indicate that ES leads to improved outcomes,[9,39] and that it is relatively cost-effective.[25,93] However, these studies compared ES with later surgery[9,25,39,93] and did not include a 'no operative procedure', which is within the definition of NES used in the ESORT study. Hence, the finding from the ESORT study, that is for patients with 'acute cholelithiasis' outcomes and costs were similar following ES and NES, is not directly comparable to previous studies.

For patients with an abdominal wall hernia, the proportions of patients in the ESORT study receiving ES and NES strategies were more similar (ES, 58.8%; NES strategies, 41.2%) than for the other four conditions. This similarity may indicate greater equipoise in the choice of ES compared with NES strategies. In addition, for patients with an abdominal wall hernia, any differences between the comparison groups in the prognosis of patients at baseline was small. In the elective setting, RCTs have been carried out for patients with a hernia, but these RCTs have mainly considered the clinical effectiveness and cost-effectiveness of laparoscopic surgery compared with open repair,[86] and included patients with only minimally symptomatic disease.[105,106] Previous observational studies have suggested that watchful waiting for patients with ventral hernias may be safe[107] and other studies have suggested that repair at diagnosis is cost-effective.[108] However, these observational studies all face concerns about confounding by indication.[109] The importance of a RCT to address the evidence gaps concerning ES and NES strategies for patients with acute hernia is also highlighted by an ongoing observational study that is assessing patient outcomes, including complications following ES and NES strategies.[42] The RCT[42] is anticipated to provide complementary evidence to the ESORT study (see *Implications for further research*).

The ESORT study found particularly large heterogeneity in the estimates of the relative clinical effectiveness and cost-effectiveness for patients with intestinal obstruction according to patients'

baseline prognosis. This level of heterogeneity reflected differences in the underlying reasons for emergency admission within the two broad categories of obstruction: (1) following intestinal adhesions from previous surgery (small bowel obstruction) and (2) from cancer (large bowel obstruction). Published RCTs[85,110] have suggested that, for patients with intestinal obstruction due to left colon or rectal cancer, forms of NES, such as endoscopic stenting, could lead to similar outcomes[110] and reduced costs,[85] compared with ES, but these findings are from RCTs[85,110] that recruited small samples of patients with short periods of follow-up. For patients with a small bowel obstruction, some studies have suggested that ES may be more clinically effective and cost-effective than NES,[96] whereas other studies report that outcomes and costs were similar following ES and watchful waiting.[111,112] Although this evidence, like the ESORT study, was based on administrative hospital data, the study designs did not address concerns about confounding by indication.[96,111,112]

Subgroups

The ESORT study reported that for some conditions the relative clinical effectiveness and cost-effectiveness of ES compared with NES strategies differed according to the patients' chronological age, number of comorbidities assessed (according to the Charlson Comorbidity Index score) and frailty level (assessed using the SCARF index). The ESORT study found that the relationship between age group and cost-effectiveness was somewhat different for diverticular disease compared with the other four conditions. For patients with diverticular disease, ES was more cost-effective for patients aged 55–80 years and also, with greater uncertainty, for patients aged > 80 years. However, for younger patients (aged ≤ 55 years) with diverticular disease there was no strong evidence that either strategy was more cost-effective. For the other four conditions, NES strategies were generally more clinically effective and cost-effective for older patients, although the results did not support a particular age 'cut-off' point. For example, for patients with acute appendicitis, the age at which there was reasonably strong evidence that NES strategies were relatively cost-effective was around 70 years old; however, for patients with an abdominal wall hernia, the analogous age at which NES strategies were relatively cost-effective was ≥ 85 years.

The ESORT study found that for patients with appendicitis, cholelithiasis or a hernia, and without comorbidities, ES was the more cost-effective strategy, although the results were somewhat uncertain, with the 95% CIs around the INBs including zero. For patients with diverticular disease or an intestinal obstruction, ES was more clinically effective and cost-effective, on average, for patients with two or three comorbidities, but, again, with 95% CIs around the mean INBs including zero.

The most consistent subgroup effects across conditions were found for the level of frailty. ES strategies were generally favoured in patients classified as 'fit' and NES strategies were generally favoured in patients classified as 'severely frail'. The ESORT study found that for patients who had severe levels of frailty at emergency hospital admission, the NES strategy led to an increase in average number of DAOH at 90 days and was cost-effective compared with ES strategies. The proportion of patients with severe frailty ranged from 0.5% (in the appendicitis cohort) to 6.5% (in the intestinal obstruction cohort). The ESORT study found that for patients with moderate levels of frailty, NES strategies were also cost-effective for patients with appendicitis, an abdominal wall hernia or an intestinal obstruction. Conversely, for patients who were categorised as 'fit', for the cholelithiasis, hernia and intestinal obstruction cohorts ES strategies were more effective and were more cost-effective for all conditions (although with some uncertainty for the appendicitis cohort). The proportion of patients classified as 'fit' ranged from 45% in the intestinal obstruction cohort to 83% in the appendicitis cohort.

The concept of frailty is distinct from biological age and comorbidity, and has been described as a 'dynamic and heterogeneous manifestation of age-related decline in physiological reserve and increased vulnerability to stressors'.[35,113,114] The ESORT study assessed frailty with the SCARF index, which uses ICD-10 codes to define deficits according to functional impairment, geriatric syndromes,

problems with nutrition, cognition and mood, and medical comorbidities, with severe frailty defined as the presence of six or more deficits.[35] The SCARF index can provide a frailty assessment solely using routine data sources, such as HES. Previous research in a quite different context (breast cancer) found that frailty assessed by the SCARF index was independently associated with rates of surgery and outcomes, after adjustment for other baseline characteristics, including age and number of comorbidities.[35]

The construct behind the SCARF index is similar to the Clinical Frailty Score[115] and the Electronic Frailty Index,[116] which have previously been used to assess frailty prior to ES, and found to be independently associated with greater risks of postoperative mortality and morbidity in reports by NELA and as part of the Emergency Laparotomy Study.[59,117] The ESORT study extends these findings by including conditions that do not require emergency laparotomy, and by considering frailty level as a potential effect modifier in assessing the relative effectiveness of ES compared with NES strategies. The finding that the relative clinical effectiveness and cost-effectiveness of ES is modified by frailty level, after adjusting age and number of comorbidities, supports the inclusion of frailty assessments as part of routine preoperative assessments prior to either ES or NES strategies. A RCS report on the 'high-risk general surgical patient'[118] states that all patients over aged ≥ 65 years should have frailty assessed, but this does not happen for all relevant emergency admissions.[59,117] The ESORT study, therefore, provides evidence to support recent guidelines that have emphasised the importance of frailty assessment as part of preoperative assessment.[119]

The findings from the ESORT study also support previous recommendations that the choice of ES or NES strategies should not be according to chronological age,[28,29] but should recognise the importance of other factors, such as comorbidity and frailty, in advising patients and their carers of the relative benefits and risks of ES and NES strategies. NES strategies include later surgery, which offers the opportunity to stabilise the patient's condition, as well as potential to improve outcomes and avoid subsequent re-admissions (see *Implications for further research*).

Strengths

The ESORT study addresses an important question about the provision of ES, by considering five different common acute conditions that often present as emergency hospital admissions. This study uses a large number of nationally representing data, comprising a total of 887,495 emergency hospital admissions to 175 hospitals over a 10-year period. Previous evaluations of the clinical effectiveness or cost-effectiveness of ES compared with NES strategies for patients with acute conditions have included a narrow range of patients and small samples of patients or hospitals.[7,8,13,21,23,66] By using routine data from all NHS acute hospitals, the ESORT study was able to include large representative populations of patients, including subgroups such as patients with severe levels of frailty, older patients or patients with long-term conditions, who have been excluded from previous studies.[36] The ESORT study reported estimates of relative clinical effectiveness and cost-effectiveness for prespecified subgroups, with sufficient precision to provide useful estimates for informing clinical practice.

The study exemplifies how an advanced IV approach can be applied to routine data. Here, conventional risk adjustment approaches are problematic, as they rely on the assumption that all the potential confounders have been observed. This assumption is implausible in this context, as the choice of ES or NES strategies reflects the patient's prognosis according to measures not available within the HES data, such as the patient's disease state. Hence, estimates from conventional risk adjustment approaches using standard regression analysis, such as those considered in the sensitivity analyses, will be prone to bias. In particular, for patients with acute diverticular disease or an intestinal obstruction, we found that patients who had ES rather than NES strategies were of a more severe case mix, according to both measured

characteristics (e.g. age, frailty, number of comorbidities) and – it is likely – unmeasured factors (e.g. size of abscess). Hence, the finding from the regression analysis, that is that ES led to increases in mortality for patients with diverticular disease, is likely to reflect confounding by indication. By contrast, the IV approach that avoids assuming that all measured confounders have been observed found that, overall, clinical outcomes and costs were similar for the ES and NES strategies, and that the estimated INBs were close to zero, with substantial uncertainty surrounding the estimates. Although the IV approach also made assumptions, we found that, given the large sample size and the high levels of variation in ES rates across hospitals, the IV was sufficiently strong and balanced the observed covariates between the comparison groups. This research, therefore, extends previous methodological research that has reported that a preference-based IV can be developed from routine data and pass stringent assessments of IV validity.[20] Hence, although such IV approaches are unusual within the economic evaluation or health services literature, the ESORT study suggests that they may have wider applicability in evaluations that use routine data, which are prone to the major problem of confounding by indication.[120]

Limitations

The ESORT study's inclusion criteria and definition of baseline, subgroup and resource use variables were reliant on the coding of diagnosis and procedures available from the HES data. One concern this raises is that for those patients without diagnostic or imaging information available, the diagnostic categorisation was imprecise. For example, the 'acute appendicitis cohort' would have included patients with 'suspected appendicitis' who did not have appendicitis. The ESORT study, as with any observational study that attempts to estimate relative clinical effectiveness and cost-effectiveness, was required to make assumptions. The ESORT study design took an IV approach that avoided making the unrealistic assumption that all the potential confounding variables were observed. Instead, the IV approach assumed that the hospitals' TTO did not have a direct effect on the outcome. Although this assumption could not be verified from the data, the study provided evidence to support the underlying assumptions, and found that the study conclusions were robust when alternative standpoints were taken, for example to the different variables used as proxies for the quality of acute care within the hospital.

The ESORT study measured inpatient costs from the perspective of the NHS acute hospital, and so costs of other aspects of health and personal social services were excluded. The choice of costing perspective was driven by the availability of the resource use data that were available from the HES data requested, which were limited to inpatient admissions to NHS trust hospitals. The exclusion of broader costs, for example of NHS-funded care outside the acute hospital, could have led to the underestimation of the incremental costs of the NES strategy for some subgroups, such as patients with severe levels of frailty. These subgroups may have incurred these costs following discharge from the acute hospital.

The ESORT study was able to provide recommendations according to patient subgroup definitions available in routine data. The study did not have information pertaining to disease stage, physiology or underlying health conditions, which are generally available to the clinical team, patient and carers when deciding whether ES or NES strategies are more appropriate for the individual patient. In addition, the analysis of the ESORT study does not consider relative clinical effectiveness and cost-effectiveness according to combinations of subgroup variables.

Finally, although the ESORT study was able to define ES according to OPCS-4 procedure codes and clinically relevant time windows, information was not fully available on specific NES strategies, for example whether or not the patient had antibiotics. Further research would be useful to define protocols for NES strategies, especially for those conditions, such as abdominal wall hernia, for which previous RCTs have not been undertaken (see *Implications for further research*).

Conclusions and implications for provision of emergency surgery and non-emergency surgery strategies

- The ESORT study finds that, overall, for patients presenting as emergency hospital admissions with common acute conditions, ES and NES strategies led to similar average numbers of DAOH at 90 days, and neither strategy was more cost-effective. The initial additional costs of ES procedures were generally offset by higher re-admission and later operative costs following the NES strategies, over a 1-year period.

- The study found differences in the relative clinical effectiveness and cost-effectiveness of ES compared with NES strategies according to patients' chronological age, number of comorbidities and frailty levels (see *Table 20*).

- For patients with appendicitis, cholelithiasis, a hernia or an intestinal obstruction, NES strategies were generally more clinically effective and cost-effective for older patients, although the results did not support a general age 'cut-off' point.

- The ESORT study found that relative cost-effectiveness differed according to the number of comorbidities. NES strategies were more cost-effective, on average, for those patients with acute appendicitis, cholelithiasis or an abdominal wall hernia, who had two or more comorbidities. For patients with acute appendicitis, cholelithiasis or a hernia, ES was, on average, more cost-effective for those patients without comorbidities. For patients with acute diverticular disease or an intestinal obstruction, ES was more clinically effective and cost-effective for patients with two or more comorbidities.

- For patients with severe frailty at emergency admission, NES strategies led to better outcomes at 90 days and 1 year, and were relatively cost-effective for all five conditions. ES strategies were generally, on average, more cost-effective for patients who were 'fit' at presentation.

- The ESORT study results emphasise the importance of considering NES as well as ES strategies for patients presenting as emergency hospital admissions with common acute conditions. Overall, neither ES nor NES strategies were more cost-effective. For some patient subgroups, such as patients with severe frailty, patients with certain acute conditions with two or more comorbidities and some older age groups, NES is more cost-effective. For other patient subgroups, including patients who were fit at presentation, ES was estimated to improve outcomes and to be cost-effective.

- The evidence from the ESORT study on the relative clinical effectiveness and cost-effectiveness of ES compared with NES strategies can complement that from ongoing cohort studies, and inform future updates to NICE guidelines for emergency and acute care, including guidelines for cholelithiasis and diverticular disease.[38,48]

- The ESORT study supports ongoing RCS initiatives to encourage frailty assessment at presentation as part of comprehensive geriatric assessments, which can help in deciding the best care pathway for patients with acute conditions presenting as emergency admissions.

- The ESORT study provides additional information, alongside that from previous and ongoing research, to inform patients and carers about the potential benefits and risks of NES, as well as ES strategies, and that these may differ according to factors beyond chronological age, including frailty level and number of comorbidities.

Implications for further research

- The ESORT study highlights the importance of the appropriate choice of ES or NES strategies for patients with comorbidities, including patients with multiple long-term conditions. For these patients, it is especially challenging to assess the balance of benefits, risks and costs between prompt intervention with ES and 'optimisation', which could include reversing aspects of frailty as part of NES strategies (e.g. later surgery).[121] Although HES data have been used previously to identify multimorbidity 'clusters' in all elective and emergency hospital admissions,[122] clusters have not been identified for emergency admissions for acute conditions, and this is an important gap in evidence, as patients with long-term conditions that tend to be 'quicker to optimise' (e.g. atrial fibrillation)

may benefit more from ES than from NES strategies. By contrast, patients with underlying conditions that are 'slower to optimise' (e.g. kidney disease) or 'unlikely to optimise', (e.g. late-stage dementia) may benefit more from planned surgery or medical management, as this can allow the patient's underlying condition to be stabilised. Further research to identify clusters of patients with acute conditions who present for ES or NES could, therefore, be useful to inform how best to 'personalise' the choice of strategy. This proposed research aligns with NIHR strategy ,which emphasises that research on multiple long-term conditions (multimorbidity) is of high priority.[123]

- These findings suggest that RCTs for ES compared with NES strategies for patients with acute gastrointestinal conditions now warrants consideration. In particular, for patients with an abdominal wall hernia, the ESORT study found similar rates of ES and NES strategies, and that ES led to a small increase in average QALYs at a small additional cost. To the best of our knowledge, there is no previous RCT evidence for patients presenting with acute hernia in the emergency setting, and care pathways are not clearly defined. The ESORT study findings may, therefore, help motivate a future RCT comparing ES with NES strategies for patients with an abdominal wall hernia. This RCT could investigate the relative outcomes, costs and cost-effectiveness of ES compared with NES strategies, with eligibility criteria informed by the subgroup results of the ESORT study. The RCT could also draw on related research, which examines complications for patients following acute symptomatic hernia.[42]

- Further research is required on patient-reported outcomes following ES and NES strategies. For conditions such as acute diverticulitis, previous studies have measured HRQoL for only small samples of patients,[15] and this has not been used in comparing ES with NES strategies.

- Following the COVID-19 pandemic, NHS waiting lists in England are projected to reach 13 million by 2025,[4] with implications for hospitals' capacity for elective surgery and ES. The ESORT study approach of combining large-scale routine data with advanced analytical approaches to address confounding and heterogeneity could be extended to other acute conditions. This research design can provide timely evidence to help ensure the most cost-effective strategy (ES or NES) is provided according to the patients' prognosis, and to help reduce pressure on scarce resources from unnecessary ES or subsequent hospital re-admissions.

Acknowledgements

Contributions of authors

Richard Grieve (https://orcid.org/0000-0001-8899-1301) (Professor, Health Economics) designed the ESORT study; directed the acquisition and analysis of data and interpretation of the results; and drafted the report manuscript.

Andrew Hutchings (https://orcid.org/0000-0003-0215-9923) (Assistant Professor, Health Services Research) contributed to the study design; undertook the analysis and interpreted the results; helped draft the report manuscript.

Silvia Moler Zapata (https://orcid.org/0000-0003-4733-5601) (PhD student, Health Services Research) undertook the analysis and interpreted the results; helped draft the report manuscript.

Stephen O'Neill (https://orcid.org/0000-0002-0022-0500) (Associate Professor, Health Economics) contributed to the study design; undertook the analysis and interpreted the results; helped draft the report manuscript.

David G Lugo-Palacios (https://orcid.org/0000-0003-2621-5627) (Assistant Professor, Health Economics) contributed to the study design; undertook the analysis and interpreted the results; helped draft the report manuscript.

Richard Silverwood (https://orcid.org/0000-0002-2744-1194) (Associate Professor, Statistics) contributed to the study design; interpreted the results; helped draft the report manuscript.

David Cromwell (https://orcid.org/0000-0002-6516-8125) (Professor, Health Services Research) contributed to the study design; interpreted the results; helped draft the report manuscript.

Tommaso Kircheis (https://orcid.org/0000-0001-7162-2518) (Master's Degree Student, Public Health) assisted in interpreting the results; helped draft the report manuscript.

Publications

Hutchings A, Moler Zapata S, O'Neill S, Smart N, Cromwell D, Hinchliffe R, Grieve R. Variation in the rates of emergency surgery amongst emergency admissions to hospital for common acute conditions. *BJS Open* 2021;**5**:zrab094. https://doi.org/10.1093/bjsopen/zrab094

Hutchings A, O'Neill S, Lugo-Palacios D, Moler Zapata S, Silverwood R, Cromwell D, *et al.* Effectiveness of emergency surgery for five common acute conditions: an instrumental variable analysis of a national routine database. *Anaesthesia* 2022;**77**:865–81.

Moler-Zapata S, Grieve R, Lugo-Palacios D, Hutchings A, Silverwood R, Keele K, *et al.* Local instrumental variable methods to address confounding and heterogeneity when using electronic health records: an application to emergency surgery. *Med Decis Making* 2022;**42**:1010–26.

Data-sharing statement

All data requests should be submitted to the corresponding author for consideration. Access to available anonymised or aggregated data may be granted following review and is subject to Data Sharing Agreement NIC-185179, which expires on 1 November 2023.

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.

Information about the ESORT's study's use of data and privacy statements are available via URL: www.lshtm.ac.uk/research/centres-projects-groups/esort#privacy (accessed 13 September 2022).

References

1. Stewart B, Khanduri P, McCord C, Ohene-Yeboah M, Uranues S, Vega Rivera F, Mock C. Global disease burden of conditions requiring emergency surgery. *Br J Surg* 2014;**101**:e9–22. https://doi.org/10.1002/bjs.9329

2. Abercrombie J. *Getting it Right First Time (GiRFT) Report: General Surgery*. 2017. http://gettingitrightfirsttime.co.uk/national-general-surgery-report-published-2/ (accessed 20 April 2022).

3. Abbott TEF, Fowler AJ, Dobbs TD, Harrison EM, Gillies MA, Pearse RM. Frequency of surgical treatment and related hospital procedures in the UK: a national ecological study using hospital episode statistics. *Br J Anaesth* 2017;**119**:249–57. https://doi.org/10.1093/bja/aex137

4. Institute for Fiscal Studies. *Could NHS Waiting Lists Really Reach 13 Million?* URL: https://ifs.org.uk/publications/15557 (accessed 19 October 2022).

5. Watson R, Crump H, Imison C, Currie C, Gaskins M. *Emergency General Surgery: Challenges and Opportunities*. London: Nuffield Trust; 2016.

6. Wohlgemut JM, Ramsay G, Jansen JO. The changing face of emergency general surgery: a 20-year analysis of secular trends in demographics, diagnoses, operations, and outcomes. *Ann Surg* 2020;**271**:581–9. https://doi.org/10.1097/SLA.0000000000003066

7. Flum DR, Davidson GH, Monsell SE, Shapiro NI, Odom SR, Sanchez SE, *et al.* A randomized trial comparing antibiotics with appendectomy for appendicitis. *N Engl J Med* 2020;**383**:1907–19. https://doi.org/10.1056/NEJMoa2014320

8. O'Leary DP, Walsh SM, Bolger J, Baban C, Humphreys H, O'Grady S, *et al.* A randomized clinical trial evaluating the efficacy and quality of life of antibiotic-only treatment of acute uncomplicated appendicitis: results of the COMMA trial. *Ann Surg* 2021;**274**:240–7. https://doi.org/10.1097/SLA.0000000000004785

9. Siddiqui T, MacDonald A, Chong PS, Jenkins JT. Early versus delayed laparoscopic cholecystectomy for acute cholecystitis: a meta-analysis of randomized clinical trials. *Am J Surg* 2008;**195**:40–7. https://doi.org/10.1016/j.amjsurg.2007.03.004

10. Roulin D, Saadi A, Di Mare L, Demartines N, Halkic N. Early versus delayed cholecystectomy for acute cholecystitis, are the 72 hours still the rule?: a randomized trial. *Ann Surg* 2016;**264**:717–22. https://doi.org/10.1097/SLA.0000000000001886

11. Gutt CN, Encke J, Köninger J, Harnoss JC, Weigand K, Kipfmüller K, *et al.* Acute cholecystitis: early versus delayed cholecystectomy, a multicenter randomized trial (ACDC study, NCT00447304). *Ann Surg* 2013;**258**:385–93. https://doi.org/10.1097/SLA.0b013e3182a1599b

12. Gargya V, Smolarek S, Walsh TN. Concerns about acute cholecystitis: early versus delayed cholecystectomy – a multicenter randomized trial. *Ann Surg* 2015;**262**:e63–4. https://doi.org/10.1097/SLA.0000000000000554

13. Harnoss JC, Zelienka I, Probst P, Grummich K, Müller-Lantzsch C, Harnoss JM, *et al.* Antibiotics versus surgical therapy for uncomplicated appendicitis: systematic review and meta-analysis of controlled trials (PROSPERO 2015: CRD42015016882). *Ann Surg* 2017;**265**:889–900. https://doi.org/10.1097/SLA.0000000000002039

14. Thornell A, Angenete E, Bisgaard T, Bock D, Burcharth J, Heath J, *et al.* Laparoscopic lavage for perforated diverticulitis with purulent peritonitis: a randomized trial. *Ann Intern Med* 2016;**164**:137–45. https://doi.org/10.7326/M15-1210

15. Azhar N, Johanssen A, Sundström T, Folkesson J, Wallon C, Kørner H, *et al.* Laparoscopic lavage vs primary resection for acute perforated diverticulitis: long-term outcomes from the Scandinavian Diverticulitis (SCANDIV) randomized clinical trial. *JAMA Surg* 2021;**156**:121–7. https://doi.org/10.1001/jamasurg.2020.5618

16. EMSurg Collaborators. Methodological overview of systematic reviews to establish the evidence base for emergency general surgery. *Br J Surg* 2017;**104**:513–24. https://doi.org/10.1002/bjs.10476sd

17. Cao AM, Eslick GD, Cox MR. Early laparoscopic cholecystectomy is superior to delayed acute cholecystitis: a meta-analysis of case-control studies. *Surg Endosc* 2016;**30**:1172–82. https://doi.org/10.1007/s00464-015-4325-4

18. Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. *Stat Med* 2014;**33**:2297–340. https://doi.org/10.1002/sim.6128

19. Brookhart MA, Schneeweiss S. Preference-based instrumental variable methods for the estimation of treatment effects: assessing validity and interpreting results. *Int J Biostat* 2007;**3**:Article 14. https://doi.org/10.2202/1557-4679.1072

20. Keele L, Sharoky C, Sellers M, Wirtalla C, Kelz R. An instrumental variables design for the effect of emergency general surgery. *Epidemiol Methods* 2018;**7**:20170012. https://doi.org/10.1515/em-2017-0012

21. Javanmard-Emamghissi H, Hollyman M, Boyd-Carson H, Doleman B, Adiamah A, Lund JN, *et al.* Antibiotics as first-line alternative to appendicectomy in adult appendicitis: 90-day follow-up from a prospective, multicentre cohort study. *Br J Surg* 2021;**108**:1351–9. https://doi.org/10.1093/bjs/znab287

22. Sugiura K, Suzuki K, Umeyama T, Omagari K, Hashimoto T, Tamura A. Cost-effectiveness analysis of initial nonoperative management versus emergency laparoscopic appendectomy for acute complicated appendicitis. *BMC Health Serv Res* 2020;**20**:1019. https://doi.org/10.1186/s12913-020-05839-6

23. Sippola S, Grönroos J, Tuominen R, Paajanen H, Rautio T, Nordström P, *et al.* Economic evaluation of antibiotic therapy versus appendicectomy for the treatment of uncomplicated acute appendicitis from the APPAC randomized clinical trial. *Br J Surg* 2017;**104**:1355–61. https://doi.org/10.1002/bjs.10575

24. Wu JX, Dawes AJ, Sacks GD, Brunicardi FC, Keeler EB. Cost effectiveness of nonoperative management versus laparoscopic appendectomy for acute uncomplicated appendicitis. *Surgery* 2015;**158**:712–21. https://doi.org/10.1016/j.surg.2015.06.021

25. Sutton AJ, Vohra RS, Hollyman M, Marriott PJ, Buja A, Alderson D, *et al.* Cost-effectiveness of emergency versus delayed laparoscopic cholecystectomy for acute gallbladder pathology. *Br J Surg* 2017;**104**:98–107. https://doi.org/10.1002/bjs.10317

26. Morris S, Gurusamy KS, Patel N, Davidson BR. Cost-effectiveness of early laparoscopic cholecystectomy for mild acute gallstone pancreatitis. *Br J Surg* 2014;**101**:828–35. https://doi.org/10.1002/bjs.9501

27. Macafee DA, Humes DJ, Bouliotis G, Beckingham IJ, Whynes DK, Lobo DN. Prospective randomized trial using cost–utility analysis of early versus delayed laparoscopic cholecystectomy for acute gallbladder disease. *Br J Surg* 2009;**96**:1031–40. https://doi.org/10.1002/bjs.6685

28. Royal College of Surgeons of England, Age UK. *Access all Ages: Assessing the Impact of Age on Access to Surgical Treatment.* London: Royal College of Surgeons; 2012.

29. Royal College of Surgeons, Age UK. *Access All Ages 2: Exploring Variations in Access to Surgery Among Older Patients.* London: Royal College of Surgeons; 2014.

30. London School of Hygiene and Tropical Medicine. *ESORT*. URL: www.lshtm.ac.uk/research/centres-projects-groups/esort#welcome (accessed 20 April 2022).

31. Hernan MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol* 2016;**183**:758–64. https://doi.org/10.1093/aje/kwv254

32. García-Albéniz X, Hsu J, Hernán MA. The value of explicitly emulating a target trial when using real world evidence: an application to colorectal cancer screening. *Eur J Epidemiol* 2017;**32**:495–500. https://doi.org/10.1007/s10654-017-0287-2

33. Herbert A, Wijlaars L, Zylbersztejn A, Cromwell D, Hardelid P. Data resource profile: Hospital Episode Statistics Admitted Patient Care (HES APC). *Int J Epidemiol* 2017;**46**:1093–1093i. https://doi.org/10.1093/ije/dyx015

34. Armitage JN, van der Meulen JH, Royal College of Surgeons Co-morbidity Consensus Group. Identifying co-morbidity in surgical patients using administrative data with the Royal College of Surgeons Charlson Score. *Br J Surg* 2010;**97**:772–81. https://doi.org/10.1002/bjs.6930

35. Jauhari Y, Gannon MR, Dodwell D, Horgan K, Clements K, Medina J, *et al.* Construction of the Secondary Care Administrative Records Frailty (SCARF) index and validation on older women with operable invasive breast cancer in England and Wales: a cohort study. *BMJ Open* 2020;**10**:e035395. https://doi.org/10.1136/bmjopen-2019-035395

36. Hutchings A, Moler Zapata S, O'Neill S, Smart N, Cromwell D, Hinchliffe R, Grieve R. Variation in the rates of emergency surgery amongst emergency admissions to hospital for common acute conditions. *BJS Open* 2021;**5**:zrab094. https://doi.org/10.1093/bjsopen/zrab094

37. CholeS Study Group, West Midlands Research Collaborative. Population-based cohort study of variation in the use of emergency cholecystectomy for benign gallbladder diseases. *Br J Surg* 2016;**103**:1716–26. https://doi.org/10.1002/bjs.10288

38. National Institute for Health and Care Excellence. *Gallstone Disease: Diagnosis and Management*. Clinical Guideline [CG188]. 2014. URL: www.nice.org.uk/guidance/cg188 (accessed 6 May 2021).

39. Wu XD, Tian X, Liu MM, Wu L, Zhao S, Zhao L. Meta-analysis comparing early versus delayed laparoscopic cholecystectomy for acute cholecystitis. *Br J Surg* 2015;**102**:1302–13. https://doi.org/10.1002/bjs.9886

40. Bamber JR, Stephens TJ, Cromwell DA, Duncan E, Martin GP, Quiney NF, *et al.* Effectiveness of a quality improvement collaborative in reducing time to surgery for patients requiring emergency cholecystectomy. *BJS Open* 2019;**3**:802–11. https://doi.org/10.1002/bjs5.50221

41. Peacock O, Bassett MG, Kuryba A, Walker K, Davies E, Anderson I, Vohra RS, National Emergency Laparotomy Audit (NELA) Project Team. Thirty-day mortality in patients undergoing laparotomy for small bowel obstruction. *Br J Surg* 2018;**105**:1006–13. https://doi.org/10.1002/bjs.10812

42. British Hernia Society. *Management of Acutely Symptomatic Hernias (MASH) Study Protocol*. URL: www.britishherniasociety.org/management-of-acutely-symptomatic-hernias-mash-study (accessed 27 October 2021).

43. Hwang MJ, Bhangu A, Webster CE, Bowley DM, Gannon MX, Karandikar SS. Unintended consequences of policy change to watchful waiting for asymptomatic inguinal hernias. *Ann R Coll Surg Engl* 2014;**96**:343–7. https://doi.org/10.1308/003588414X13946184902000

44. Orchard MR, Wright JA, Kelly A, McCabe DJ, Hewes J. The impact of healthcare rationing on elective and emergency hernia repair. *Hernia* 2016;**20**:405–9. https://doi.org/10.1007/s10029-015-1441-y

45. Schultz JK, Yaqub S, Wallon C, Blecic L, Forsmo HM, Folkesson J, *et al.* Laparoscopic lavage vs. primary resection for acute perforated diverticulitis: the SCANDIV randomized clinical trial. *JAMA* 2015;**314**:1364–75. https://doi.org/10.1001/jama.2015.12076

46. Schultz JK, Wallon C, Blecic L, Forsmo HM, Folkesson J, Buchwald P, *et al.* One-year results of the SCANDIV randomized clinical trial of laparoscopic lavage versus primary resection for acute perforated diverticulitis. *Br J Surg* 2017;**104**:1382–92. https://doi.org/10.1002/bjs.10567

47. Paterson HM, Arnott ID, Nicholls RJ, Clark D, Bauer J, Bridger PC, *et al.* Diverticular disease in Scotland: 2000–2010. *Colorectal Dis* 2015;**17**:329–34. https://doi.org/10.1111/codi.12811

48. National Institute for Health and Care Excellence. *Diverticular Disease: Diagnosis and Management.* NICE Guideline [NG147]. 2019. URL: www.nice.org.uk/guidance/ng147 (accessed 6 May 2021).

49. Strong SA, Soper NJ. Minimally invasive approaches to rectal cancer and diverticulitis: does less mean more? *JAMA* 2015;**314**:1343–5. https://doi.org/10.1001/jama.2015.11454

50. Jerath A, Austin PC, Wijeysundera DN. Days alive and out of hospital: validation of a patient-centered outcome for perioperative medicine. *Anesthesiology* 2019;**131**:84–93. https://doi.org/10.1097/ALN.0000000000002701

51. Boney O, Moonesinghe SR, Myles PS, Grocott MPW, StEP-COMPAC group. Core Outcome Measures for Perioperative and Anaesthetic Care (COMPAC): a modified Delphi process to develop a core outcome set for trials in perioperative care and anaesthesia. *Br J Anaesth* 2022;**128**:174–85. https://doi.org/10.1016/j.bja.2021.09.027

52. Spurling LJ, Moonesinghe SR, Oliver CM. Validation of the days alive and out of hospital (DAOH) outcome measure in a retrospective cohort of patients having emergency laparotomy. *Br J Anaesth* 2022;**128**:449–56. https://doi.org/10.1016/j.bja.2021.12.006

53. Staiger D, Stock JH. Instrumental variables regression with weak instruments. *Econometrica* 1997;**65**:557–86. https://doi.org/10.2307/2171753

54. Branson Z, Keele L. Evaluating a key instrumental variable assumption using randomization tests. *Am J Epidemiol* 2020;**189**:1412–20. https://doi.org/10.1093/aje/kwaa089

55. Basu A. Estimating person-centered treatment (PeT) effects using instrumental variables: an application to evaluating prostate cancer treatments. *J Appl Econ* 2014;**29**:671–91. https://doi.org/10.1002/jae.2343

56. Basu A. Person-centered treatment (PeT) effects: individualized treatment effects using instrumental variables. *Stata J* 2015;**15**:397–410. https://doi.org/10.1177/1536867X1501500204

57. Basu A, Gore JL. Are elderly patients with clinically localized prostate cancer overtreated? Exploring heterogeneity in survival effects. *Med Care* 2015;**53**:79–86. https://doi.org/10.1097/MLR.0000000000000260

58. Grieve R, O'Neill S, Basu A, Keele L, Rowan KM, Harris S. Analysis of benefit of intensive care unit transfer for deteriorating ward patients: a patient-centered approach to clinical evaluation. *JAMA Netw Open* 2019;**2**:e187704. https://doi.org/10.1001/jamanetworkopen.2018.7704

59. National Emergency Laparotomy Audit (NELA) Project Team. *Second Patient Report of the National Emergency Laparotomy Audit.* 2016. URL: www.nela.org.uk/reports (accessed 20 April 2022).

60. National Emergency Laparotomy Audit (NELA) Project Team. *Third Patient Report of the National Emergency Laparotomy Audit.* 2017. URL: www.nela.org.uk/reports (accessed 20 April 2022).

61. National Emergency Laparotomy Audit (NELA) Project Team. *Fourth Patient Report of the National Emergency Laparotomy Audit.* 2018. URL: www.nela.org.uk/reports (accessed 20 April 2022).

62. Ariti CA, Cleland JG, Pocock SJ, Pfeffer MA, Swedberg K, Granger CB, *et al.* Days alive and out of hospital and the patient journey in patients with heart failure: Insights from the candesartan in heart failure: assessment of reduction in mortality and morbidity (CHARM) program. *Am Heart J* 2011;**162**:900–6. https://doi.org/10.1016/j.ahj.2011.08.003

63. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal 2013.* Process and Methods [PMG9]. URL: www.nice.org.uk/process/pmg9/chapter/foreword (accessed 5 September 2022).

64. Neumann PJ, Ganiats TG, Russell LB, Sanders GD, Siegel JE. *Cost-Effectiveness in Health and Medicine.* 2nd edn. New York, NY: Oxford University Press; 2016. https://doi.org/10.1093/acprof:oso/9780190492939.001.0001

65. Salminen P, Paajanen H, Rautio T, Nordström P, Aarnio M, Rantanen T, *et al.* Antibiotic therapy vs appendectomy for treatment of uncomplicated acute appendicitis: the APPAC randomized clinical trial. *JAMA* 2015;**313**:2340–8. https://doi.org/10.1001/jama.2015.6154

66. Salminen P, Tuominen R, Paajanen H, Rautio T, Nordström P, Aarnio M, *et al.* Five-year follow-up of antibiotic therapy for uncomplicated acute appendicitis in the APPAC randomized clinical trial. *JAMA* 2018;**320**:1259–65. https://doi.org/10.1001/jama.2018.13201

67. Sippola S, Haijanen J, Viinikainen L, Grönroos J, Paajanen H, Rautio T, *et al.* Quality of life and patient satisfaction at 7-year follow-up of antibiotic therapy vs appendectomy for uncomplicated acute appendicitis: a secondary analysis of a randomized clinical trial. *JAMA Surg* 2020;**155**:283–9. https://doi.org/10.1001/jamasurg.2019.6028

68. Bell H WA, Hernandez M, Grieve R, Faria R, Gibson L, Grimm S. *The Use of Real World Data for the Estimation of Treatment Effects in NICE Decision Making: Report by the Decision Support Unit.* Sheffield: University of Sheffield; 2016.

69. Faria R, Hernández Alava M, Manca A, Wailoo AJ. *The Use of Observational Data to Inform Estimates of Treatment Effectiveness in Technology Appraisal: Methods for Comparative Individual Patient Data.* Sheffield: University of Sheffield; 2015.

70. Raine RFR, Barratt H, Bevan G, Black N, Boaden R, Bower P, *et al.* Challenges, solutions and future directions in the evaluation of service innovations in health care and public health. *Health Serv Deliv Res* 2016;**4**(16). https://doi.org/10.3310/hsdr04160

71. Basu A. Economics of individualization in comparative effectiveness research and a basis for a patient-centered health care. *J Health Econ* 2011;**30**:549–59. https://doi.org/10.1016/j.jhealeco.2011.03.004

72. Pennington M, Grieve R, Sekhon JS, Gregg P, Black N, van der Meulen JH. Cemented, cementless, and hybrid prostheses for total hip replacement: cost effectiveness analysis. *BMJ* 2013;**346**:f1026. https://doi.org/10.1136/bmj.f1026

73. Kwong E, Abel G, Black N. Can Patient Reported Outcomes (PROs) from population surveys provide accurate estimates of pre-admission health status of emergency hospital admissions? *Patient Relat Outcome Meas* 2020;**11**:39–48 https://doi.org/10.2147/PROM.S215513

74. Haijanen J, Sippola S, Tuominen R, Grönroos J, Paajanen H, Rautio T, *et al.* Cost analysis of antibiotic therapy versus appendectomy for treatment of uncomplicated acute appendicitis: 5-year results of the APPAC randomized clinical trial. *PLOS ONE* 2019;**14**:e0220202. https://doi.org/10.1371/journal.pone.0220202

75. Saunders DI, Sinclair RCF, Griffiths B, Pugh E, Harji D, Salas B, *et al.* Emergency Laparotomy Follow-Up Study (ELFUS): prospective feasibility investigation into postoperative complications and quality of life using patient-reported outcome measures up to a year after emergency laparotomy. *Perioper Med* 2021;**10**:22. https://doi.org/10.1186/s13741-021-00193-5

76. Campbell HE, Stokes EA, Bargo D, Logan RF, Mora A, Hodge R, *et al.* Costs and quality of life associated with acute upper gastrointestinal bleeding in the UK: cohort analysis of patients in a cluster randomised trial. *BMJ Open* 2015;**5**:e007230. https://doi.org/10.1136/bmjopen-2014-007230

77. Investigators IT. Endovascular strategy or open repair for ruptured abdominal aortic aneurysm: one-year outcomes from the IMPROVE randomized trial. *Eur Heart J* 2015;**36**:2061–9. https://doi.org/10.1093/eurheartj/ehv125

78. NHS Digital. *Hospital Episode Statistics Dictionary*. URL: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics/hospital-episode-statistics-data-dictionary (accessed 1 September 2018).

79. Information Services Division (ISD) Scotland. *Theatres Costs – Detailed – SFR 5.10*. 2019. URL: www.isdscotland.org/Health-Topics/Finance/Costs/Detailed-Tables/Theatres.asp (accessed 20 April 2022).

80. Ismail I, Wolff S, Gronfier A, Mutter D, Swanström LL, Swantröm LL. A cost evaluation methodology for surgical technologies. *Surg Endosc* 2015;**29**:2423–32. https://doi.org/10.1007/s00464-014-3929-4

81. NHS Improvement. *NHS Reference Costs 2017/2018*. 2018. URL: https://webarchive.nationalarchives.gov.uk/ukgwa/20200501111106/https://improvement.nhs.uk/resources/reference-costs/ (accessed 20 April 2022).

82. HM Treasury Department. *Annex A: How to Use the GDP Deflator Series: Practical Examples*. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/205904/GDP_Deflators_User_Guide.pdf (accessed 19 August 2021).

83. Sutherland JM, Albanese CM, Crump T, Liu G, Karimuddin A. The minimally important difference of the Gastrointestinal Quality of Life Index for symptomatic gallstone surgery. *Surg Endosc* 2021;**35**:6938–48. https://doi.org/10.1007/s00464-020-08205-z

84. Rutegård M, Gümüşçü R, Stylianidis G, Nordin P, Nilsson E, Haapamäki MM. Chronic pain, discomfort, quality of life and impact on sex life after open inguinal hernia mesh repair: an expertise-based randomized clinical trial comparing lightweight and heavyweight mesh. *Hernia* 2018;**22**:411–18. https://doi.org/10.1007/s10029-018-1734-z

85. Young CJ, Zahid A. Randomized controlled trial of colonic stent insertion in non-curable large bowel obstruction: a post hoc cost analysis. *Colorectal Dis* 2018;**20**:288–95. https://doi.org/10.1111/codi.13951

86. McCormack K, Wake B, Perez J, Fraser C, Cook J, McIntosh E, *et al.* Laparoscopic surgery for inguinal hernia repair: systematic review of effectiveness and economic evaluation. *Health Technol Assess* 2005;**9**(14). https://doi.org/10.3310/hta9140

87. Sharma P, Boyers D, Scott N, Hernández R, Fraser C, Cruickshank M, *et al.* The clinical effectiveness and cost-effectiveness of open mesh repairs in adults presenting with a clinically diagnosed primary unilateral inguinal hernia who are operated in an elective setting: systematic review and economic evaluation. *Health Technol Assess* 2015;**19**(92). https://doi.org/10.3310/hta19920

88. Ara R, Brazier J, Zouraq IA. The use of health state utility values in decision models. *PharmacoEconomics* 2017;**35**:77–88. https://doi.org/10.1007/s40273-017-0550-0

89. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value Health* 2010;**13**:509–18. https://doi.org/10.1111/j.1524-4733.2010.00700.x

90. Widding-Havneraas T, Chaulagain A, Lyhmann I, Zachrisson HD, Elwert F, Markussen S, *et al.* Preference-based instrumental variables in health research rely on important and underreported assumptions: a systematic review. *J Clin Epidemiol* 2021;**139**:269–78. https://doi.org/10.1016/j.jclinepi.2021.06.006

91. Hosmer DW, Lemeshow S. *Applied Logistic Regression*. 2nd edn. New York, NY: Wiley; 2000. https://doi.org/10.1002/0471722146

92. Pregibon D. Goodness of link tests for generalized linear models. *J Roy Stat Soc Ser C (Applied Stat)* 1980;**29**:14–15. https://doi.org/10.2307/2346405

93. Wilson E, Gurusamy K, Gluud C, Davidson BR. Cost–utility and value-of-information analysis of early versus delayed laparoscopic cholecystectomy for acute cholecystitis. *Br J Surg* 2010;**97**:210–19. https://doi.org/10.1002/bjs.6872

94. Bolkenstein HE, de Wit GA, Consten ECJ, Van de Wall BJM, Broeders IAMJ, Draaisma WA. Cost-effectiveness analysis of a multicentre randomized clinical trial comparing surgery with conservative management for recurrent and ongoing diverticulitis (DIRECT trial). *Br J Surg* 2019;**106**:448–57. https://doi.org/10.1002/bjs.11024

95. Palmqvist E, Larsson K, Anell A, Hjalmarsson C. Prospective study of pain, quality of life and the economic impact of open inguinal hernia repair. *Br J Surg* 2013;**100**:1483–8. https://doi.org/10.1002/bjs.9232

96. Behman R, Nathens AB, Pechlivanoglou P, Karanicolas P, Jung J, Look Hong N. Early operative management in patients with adhesive small bowel obstruction: population-based cost analysis. *BJS Open* 2020;**4**:914–23. https://doi.org/10.1002/bjs5.50311

97. Kreif N, Grieve R, Sadique MZ. Statistical methods for cost-effectiveness analyses that use observational data: a critical appraisal tool and review of current practice. *Health Econ* 2013;**22**:486–500. https://doi.org/10.1002/hec.2806

98. Thorn JC, Turner EL, Hounsome L, Walsh E, Down L, Verne J, *et al.* Validating the use of Hospital Episode Statistics data and comparison of costing methodologies for economic evaluation: an end-of-life case study from the Cluster randomised triAl of PSA testing for Prostate cancer (CAP). *BMJ Open* 2016;**6**:e011063. https://doi.org/10.1136/bmjopen-2016-011063

99. Lee K, Bargagli-Stoffi FJ, Dominici F. *Causal Rule Ensemble: Interpretable Inference of Heterogeneous Treatment Effects*. 2020. URL: http://arxiv.org/abs/2009.09036 (accessed 5 September 2022).

100. Hettiaratchy S, Deakin D. *COVID-19: Guidance for Surgeons Working During the Pandemic*. 2020. URL: www.rcseng.ac.uk/coronavirus/joint-guidance-for-surgeons-v1/ (accessed 27 October 2021).

101. Javanmard-Emamghissi H, Boyd-Carson H, Hollyman M, Doleman B, Adiamah A, Lund JN, *et al.* The management of adult appendicitis during the COVID-19 pandemic: an interim analysis of a UK cohort study. *Tech Coloproctol* 2021;**25**:401–11.

102. Javanmard-Emamghissi H, Boyd-Carson H, Hollyman M, Doleman B, Adiamah A, Lund JN, *et al.* Correction to: The management of adult appendicitis during the COVID–19 pandemic: an interim analysis of a UK cohort study. *Tech Coloproctol* 2021;**25**:149.

103. Bhangu A, RIFT Study Group on behalf of the West Midlands Research Collaborative. Evaluation of appendicitis risk prediction models in adults with suspected appendicitis. *Br J Surg* 2020;**107**:73–86. https://doi.org/10.1002/bjs.11440

104. Vimalachandran D, Knowles C, Pinkney T. *Diverticulitis Management: A Snapshot Collaborative Audit Study. DAMASCUS Protocol*. 2020. URL: https://mcusercontent.com/99832f93daeaab6070e0f6946/files/9f1a3098-e318-4e46-8172-d4a209d2ca24/DAMASCUS_Protocol_V3.0_28_Sep_2020_.pdf (accessed 19 October 2021).

105. Fitzgibbons RJ, Giobbie-Hurder A, Gibbs JO, Dunlop DD, Reda DJ, McCarthy M, *et al.* Watchful waiting vs repair of inguinal hernia in minimally symptomatic men: a randomized clinical trial. *JAMA* 2006;**295**:285–92. https://doi.org/10.1001/jama.295.3.285

106. Fitzgibbons RJ, Ramanan B, Arya S, Turner SA, Li X, Gibbs JO, *et al.* Long-term results of a randomized controlled trial of a nonoperative strategy (watchful waiting) for men with minimally symptomatic inguinal hernias. *Ann Surg* 2013;**258**:508–15. https://doi.org/10.1097/SLA.0b013e3182a19725

107. Kokotovic D, Sjølander H, Gögenur I, Helgstrand F. Watchful waiting as a treatment strategy for patients with a ventral hernia appears to be safe. *Hernia* 2016;**20**:281–7. https://doi.org/10.1007/s10029-016-1464-z

108. Wolf LL, Ejiofor JI, Wang Y, Hunink MG, Losina E, Haider AH, Smink DS. Management of reducible ventral hernias: clinical outcomes and cost-effectiveness of repair at diagnosis versus watchful waiting. *Ann Surg* 2019;**269**:358–66. https://doi.org/10.1097/SLA.0000000000002507

109. Bellows CF, Robinson C, Fitzgibbons RJ, Webber LS, Berger DH. Watchful waiting for ventral hernias: a longitudinal study. *Am Surg* 2014;**80**:245–52. https://doi.org/10.1177/000313481408000319

110. Cirocchi R, Farinella E, Trastulli S, Desiderio J, Listorti C, Boselli C, *et al.* Safety and efficacy of endoscopic colonic stenting as a bridge to surgery in the management of intestinal obstruction due to left colon and rectal cancer: a systematic review and meta-analysis. *Surg Oncol* 2013;**22**:14–21. https://doi.org/10.1016/j.suronc.2012.10.003

111. Bauer J, Keeley B, Krieger B, Deliz J, Wallace K, Kruse D, *et al.* Adhesive small bowel obstruction: early operative versus observational management. *Am Surg* 2015;**81**:614–20. https://doi.org/10.1177/000313481508100627

112. Ten Broek RPG, Krielen P, Di Saverio S, Coccolini F, Biffl WL, Ansaloni L, *et al.* Bologna guidelines for diagnosis and management of adhesive small bowel obstruction (ASBO): 2017 update of the evidence-based guidelines from the world society of emergency surgery ASBO working group. *World J Emerg Surg* 2018;**13**:24. https://doi.org/10.1186/s13017-018-0185-2

113. O'Caoimh R, Sezgin D, O'Donovan MR, Molloy DW, Clegg A, Rockwood K, Liew A. Prevalence of frailty in 62 countries across the world: a systematic review and meta-analysis of population-level studies. *Age Ageing* 2021;**50**:96–104. https://doi.org/10.1093/ageing/afaa219

114. Xue QL. The frailty syndrome: definition and natural history. *Clin Geriatr Med* 2011;**27**:1–15. https://doi.org/10.1016/j.cger.2010.08.009

115. Rockwood K, Song X, MacKnight C, Bergman H, Hogan DB, McDowell I, Mitnitski A. A global clinical measure of fitness and frailty in elderly people. *CMAJ* 2005;**173**:489–95. https://doi.org/10.1503/cmaj.050051

116. Clegg A, Bates C, Young J, Ryan R, Nichols L, Ann Teale E, *et al.* Development and validation of an electronic frailty index using routine primary care electronic health record data. *Age Ageing* 2016;**45**:353–60. https://doi.org/10.1093/ageing/afw039

117. Parmar KL, Law J, Carter B, Hewitt J, Boyle JM, Casey P, *et al.* Frailty in older patients undergoing emergency laparotomy: results from the UK observational emergency laparotomy and frailty (ELF) study. *Ann Surg* 2021;**273**:709–18. https://doi.org/10.1097/SLA.0000000000003402

118. Royal College of Surgeons. *The High-Risk General Surgical Patient: Raising the Standard.* 2018. URL: www.rcseng.ac.uk/-/media/files/rcs/news-and-events/media-centre/2018-press-releases-documents/rcs-report-the-highrisk-general-surgical-patient–raising-the-standard–december-2018.pdf (accessed 18 October 2021).

119. Centre for Perioperative Care. *Guideline for Perioperative Care for People Living with Frailty Undergoing Elective and Emergency Surgery*. 2021. URL: www.bgs.org.uk/sites/default/files/content/attachment/2021-09-28/Guideline%20for%20Perioperative%20Care%20for%20People%20Living%20with%20Frailty%20Undergoing%20Elective%20and%20Emergency%20Surgery.pdf (accessed 6 October 2021).

120. Keele L, Small D. Instrumental variables: don't throw the baby out with the bathwater. *Health Serv Res* 2019;**54**(3). https://doi.org/10.1111/1475-6773.13130

121. Poulton T, Murray D, National Emergency Laparotomy Audit (NELA) Project Team. Pre-optimisation of patients undergoing emergency laparotomy: a review of best practice. *Anaesthesia* 2019;**74**:100–7. https://doi.org/10.1111/anae.14514

122. Stokes J, Guthrie B, Mercer SW, Rice N, Sutton M. Multimorbidity combinations, costs of hospital care and potentially preventable emergency admissions in England: a cohort study. *PLOS Med* 2021;**18**:e1003514. https://doi.org/10.1371/journal.pmed.1003514

123. National Institute for Health and Care Research. *NIHR Strategic Framework for Multiple Long-Term Conditions (Multimorbidity) MLTC-M Research*. 2021. URL: www.nihr.ac.uk/documents/research-on-multiple-long-term-conditions-multimorbidity-mltc-m/24639 (accessed 27 October 2021).

124. American College of Surgeons. *Cholecystectomy: Surgical Removal of the Gallbladder*. URL: www.facs.org/~/media/files/education/patient ed/cholesys.ashx (accessed 19 October 2022).

125. Silva GL, de Moura EG, Bernardo WM, Leite de Castro V, Morais C, Baba ER, Safatle-Ribeiro AV. Endoscopic versus surgical resection for early colorectal cancer-a systematic review and meta-analysis. *J Gastrointest Oncol* 2016;**7**:326–35. https://doi.org/10.21037/jgo.2015.10.02

126. Heah SM, Eu KW, Ho YH, Leong AF, Seow-Choen F. Hartmann's procedure vs. abdominoperineal resection for palliation of advanced low rectal cancer. *Dis Colon Rectum* 1997;**40**:1313–17. https://doi.org/10.1007/BF02050815

127. Teramoto A, Aoyama N, Ebisutani C, Matsumoto T, Machida H, Yoshida S, *et al.* Clinical importance of cold polypectomy during the insertion phase in the left side of the colon and rectum: a multicenter randomized controlled trial (PRESECT study). *Gastrointest Endosc* 2020;**91**:917–24. https://doi.org/10.1016/j.gie.2019.12.019

128. Wu CC, Chueh SC, Tsai YC. Is contralateral exploration justified in endoscopic total extraperitoneal repair of clinical unilateral groin hernias – a prospective cohort study. *Int J Surg* 2016;**36**:206–11. https://doi.org/10.1016/j.ijsu.2016.10.012

129. Wilson M. Urinary catheterisation in the community: exploring challenges and solutions. *Br J Community Nurs* 2016;**21**:492–6. https://doi.org/10.12968/bjcn.2016.21.10.492

130. Sallinen V, Di Saverio S, Haukijärvi E, Juusela R, Wikström H, Koivukangas V, *et al.* Laparoscopic versus open adhesiolysis for adhesive small bowel obstruction (LASSO): an international, multicentre, randomised, open-label trial. *Lancet Gastroenterol Hepatol* 2019;**4**:278–86. https://doi.org/10.1016/S2468-1253(19)30016-0

131. Curtis L, Burns A. *Unit Costs of Health and Social Care 2019*. Canterbury: PSSRU, University of Kent; 2019.

132. Clement KD, Emslie K, Maniam P, Wilson MSJ. What is the operative cost of managing acute appendicitis in the NHS: the impact of stump technique and perioperative imaging. *World J Surg* 2020;**44**:749–54. https://doi.org/10.1007/s00268-019-05306-2

133. Stellingwerf ME, Sahami S, Winter DC, Martin ST, D'Haens GR, Cullen G, *et al.* Prospective cohort study of appendicectomy for treatment of therapy-refractory ulcerative colitis. *Br J Surg* 2019;**106**:1697–704. https://doi.org/10.1002/bjs.11259

134. Koumarelas K, Theodoropoulos GE, Spyropoulos BG, Bramis K, Manouras A, Zografos G. A prospective longitudinal evaluation and affecting factors of health related quality of life after appendectomy. *Int J Surg* 2014;**12**:848–57. https://doi.org/10.1016/j.ijsu.2014.06.015

135. Barger J, Stenquist DS, Mohamadi A, Weaver MJ, Dyer GSM, von Keudell A. Acute versus delayed reverse total shoulder arthroplasty for the management of proximal humerus fractures. *Injury* 2021;**52**:2272–8. https://doi.org/10.1016/j.injury.2021.05.040

136. Karimuddin A, Albanese CM, Crump T, Liu G, Sutherland JM. Measuring the impact of delayed access to elective cholecystectomy through patient's cost–utility: an observational cohort study. *Int J Qual Health Care* 2021;**33**:mzab018. https://doi.org/10.1093/intqhc/mzab018

137. Holmberg ST, Salvesen ØO, Vangen-Lønne V, Hara S, Fredheim OM, Solberg TK, *et al.* Pain during sex before and after surgery for lumbar disc herniation: a multicenter observational study. *Spine* 2020;**45**:1751–7. https://doi.org/10.1097/BRS.0000000000003675

138. Ruiz-Romero MV, Fernández-Ojeda MDR, Castilla Yélamo J, García-Benítez JB, Calero-Bernal ML, Fernández-Moyano A. [Influence of early hip fracture surgery in the elderly on mortality, readmissions, dependence and quality of life.] *Rev Esp Salud Publica* 2020;**94**:e202011142.

139. Bisoyi S, Jagannathan U, Dash AK, Mohapatra R, Nayak D, Sahu S, Satyanarayan P. Decision making, management, and midterm outcomes of postinfarction ventricular septal rupture: our experience with 21 patients. *Ann Card Anaesth* 2020;**23**:471–6. https://doi.org/10.4103/aca.ACA_119_19

140. van Tol FR, Suijkerbuijk KPM, Choi D, Verkooijen HM, Oner FC, Verlaan JJ. The importance of timely treatment for quality of life and survival in patients with symptomatic spinal metastases. *Eur Spine J* 2020;**29**:3170–8. https://doi.org/10.1007/s00586-020-06599-x

141. Achit H, Guillemin F, Karam G, Ladrière M, Baumann C, Frimat L, *et al.* Cost-effectiveness of four living-donor nephrectomy techniques from a hospital perspective. *Nephrol Dial Transplant* 2020;**35**:2004–12. https://doi.org/10.1093/ndt/gfz143

142. Sutherland JM, Mok J, Liu G, Karimuddin A, Crump T. A cost–utility study of laparoscopic cholecystectomy for the treatment of symptomatic gallstones. *J Gastrointest Surg* 2020;**24**:1314–19. https://doi.org/10.1007/s11605-019-04268-z

143. Sinan H, Saydam M, Demir P, Ozer MT, Demirbas S. Comparison of single-incision and conventional laparoscopic cholecystectomy in terms of quality of life, body image, and cosmesis. *Niger J Clin Pract* 2019;**22**:521–6. https://doi.org/10.4103/njcp.njcp_218_18

144. Lombardo S, Rosenberg JS, Kim J, Erdene S, Sergelen O, Nellermoe J, *et al.* Cost and outcomes of open versus laparoscopic cholecystectomy in Mongolia. *J Surg Res* 2018;**229**:186–91. https://doi.org/10.1016/j.jss.2018.03.036

145. Ramakrishnan A, Webb KM, Cowperthwaite MC. One-year outcomes of early-crossover patients in a cohort receiving nonoperative care for lumbar disc herniation. *J Neurosurg Spine* 2017;**27**:391–6. https://doi.org/10.3171/2017.2.SPINE16760

146. Rystedt JML, Tingstedt B, Montgomery F, Montgomery AK. Routine intraoperative cholangiography during cholecystectomy is a cost-effective approach when analysing the cost of iatrogenic bile duct injuries. *HPB* 2017;**19**:881–8. https://doi.org/10.1016/j.hpb.2017.06.004

147. Rosenmüller MH, Nilsson E, Lindberg F, Åberg SO, Haapamäki MM. Costs and quality of life of small-incision open cholecystectomy and laparoscopic cholecystectomy – an expertise-based randomised controlled trial. *BMC Gastroenterol* 2017;**17**:48. https://doi.org/10.1186/s12876-017-0601-1

148. Ali TF, Warkentin LM, Gazala S, Wagg AS, Padwal RS, Khadaroo RG, Acute Care and Emergency Surgery (ACES) Group. Self-reported outcomes in individuals aged 65 and older admitted for treatment to an acute care surgical service: a 6-month prospective cohort study. *J Am Geriatr Soc* 2015;**63**:2388–94. https://doi.org/10.1111/jgs.13783

149. Loures FB, Chaoubah A, Oliveira VM, Almeida AM, Campos EM, Paiva EP. Economic analysis of surgical treatment of hip fracture in older adults. *Rev Saude Publica* 2015;**49**:12. https://doi.org/10.1590/S0034-8910.2015049005172

150. Iranmanesh P, Frossard JL, Mugnier-Konrad B, Morel P, Majno P, Nguyen-Tang T, *et al.* Initial cholecystectomy vs sequential common duct endoscopic assessment and subsequent cholecystectomy for suspected gallstone migration: a randomized clinical trial. *JAMA* 2014;**312**:137–44. https://doi.org/10.1001/jama.2014.7587

151. Chikuda H, Ohtsu H, Ogata T, Sugita S, Sumitani M, Koyama Y, *et al.* Optimal treatment for spinal cord injury associated with cervical canal stenosis (OSCIS): a study protocol for a randomized controlled trial comparing early versus delayed surgery. *Trials* 2013;**14**:245. https://doi.org/10.1186/1745-6215-14-245

152. Abd Ellatif ME, Askar WA, Abbas AE, Noaman N, Negm A, El-Morsy G, *et al.* Quality-of-life measures after single-access versus conventional laparoscopic cholecystectomy: a prospective randomized study. *Surg Endosc* 2013;**27**:1896–906. https://doi.org/10.1007/s00464-012-2625-5

153. Wagner MJ, Kern H, Hapfelmeier A, Mehler J, Schoenberg MH. Single-port cholecystectomy versus multi-port cholecystectomy: a prospective cohort study with 222 patients. *World J Surg* 2013;**37**:991–8. https://doi.org/10.1007/s00268-013-1946-4

154. Mendelow AD, Gregson BA, Mitchell PM, Murray GD, Rowan EN, Gholkar AR, STICH II Investigators. Surgical trial in lobar intracerebral haemorrhage (STICH II) protocol. *Trials* 2011;**12**:124. https://doi.org/10.1186/1745-6215-12-124

155. van den Hout WB, Peul WC, Koes BW, Brand R, Kievit J, Thomeer RT, Leiden-The Hague Spine Intervention Prognostic Study Group. Prolonged conservative care versus early surgery in patients with sciatica from lumbar disc herniation: cost utility analysis alongside a randomised controlled trial. *BMJ* 2008;**336**:1351–4. https://doi.org/10.1136/bmj.39583.709074.BE

156. van Hooft JE, Bemelman WA, Breumelhof R, Siersema PD, Kruyt PM, van der Linde K, *et al.* Colonic stenting as bridge to surgery versus emergency surgery for management of acute left-sided malignant colonic obstruction: a multicenter randomized trial (Stent-in 2 study). *BMC Surg* 2007;**7**:12. https://doi.org/10.1186/1471-2482-7-12

157. Nilsson E, Ros A, Rahmqvist M, Bäckman K, Carlsson P. Cholecystectomy: costs and health-related quality of life: a comparison of two techniques. *Int J Qual Health Care* 2004;**16**:473–82. https://doi.org/10.1093/intqhc/mzh077

158. Ainslie WG, Catton JA, Davides D, Dexter S, Gibson J, Larvin M, *et al.* Micropuncture cholecystectomy vs conventional laparoscopic cholecystectomy: a randomized controlled trial. *Surg Endosc* 2003;**17**:766–72. https://doi.org/10.1007/s00464-002-8568-5

159. LeBlanc KA, Gonzalez A, Dickens E, Olsofka J, Ortiz-Ortiz C, Verdeja JC, Pierce R, Prospective Hernia Study Group. Robotic-assisted, laparoscopic, and open incisional hernia repair: early outcomes from the prospective hernia study. *Hernia* 2021;**25**:1071–82. https://doi.org/10.1007/s10029-021-02381-0

160. Asencio F, Carbó J, Ferri R, Peiró S, Aguiló J, Torrijo I, *et al.* Laparoscopic versus open incisional hernia repair: long-term follow-up results of a randomized clinical trial. *World J Surg* 2021;**45**:2734–41. https://doi.org/10.1007/s00268-021-06164-7

161. Susmallian S, Barnea R, Ponomarenko O. Laparoscopic total extraperitoneal repair and open prolene hernia system for inguinal hernia repair have similar outcomes: a retrospective study. *Chirurgia* 2021;**116**:271–83. https://doi.org/10.21614/chirurgia.116.3.271

162. Singh S, Anand A, Kumar A, Pal AK, Agrawal MK, Kumar S, *et al.* A prospective randomised control trial to compare the perioperative outcomes and ergonomic challenges between triangular versus midline port placement in total extra-peritoneal repair of uncomplicated unilateral inguinal hernia. *Surg Endosc* 2021;**35**:1395–404. https://doi.org/10.1007/s00464-020-07525-4

163. A Guzman-Pruneda F, Huang LC, Collins C, Renshaw S, Narula V, K Poulose B. Abdominal core quality of life after ventral hernia repair: a comparison of open versus robotic-assisted retromuscular techniques. *Surg Endosc* 2021;**35**:241–8. https://doi.org/10.1007/s00464-020-07386-x

164. Greco D, Santori G, Brancato G, Gossetti F, Ipponi PL, Negro P, *et al.* A new semiresorbable mesh for primary inguinal repair: a preliminary observational study on quality of life and safety. *Hernia* 2020;**24**:1019–31. https://doi.org/10.1007/s10029-020-02276-6

165. Min L, Yong P, Yun L, Balde AI, Chang Z, Qian G, *et al.* Propensity score analysis of outcomes between the transabdominal preperitoneal and open Lichtenstein repair techniques for inguinal hernia repair: a single-center experience. *Surg Endosc* 2020;**34**:5338–45. https://doi.org/10.1007/s00464-019-07324-6

166. Kaufmann R, Isemer FE, Strey CW, Jeekel J, Lange JF, Woeste G. Non-cross-linked biological mesh in complex abdominal wall hernia: a cohort study. *Langenbecks Arch Surg* 2020;**405**:345–52. https://doi.org/10.1007/s00423-020-01881-4

167. He J, Xu YJ, Sun P, Wang J, Yang CG. The incidence and analysis of ipsilateral occult hernia in patients undergoing hernia repair: a single institution retrospective study of 1066 patients. *BMC Surg* 2021;**21**:182. https://doi.org/10.1186/s12893-021-01181-8

168. Prabhu AS, Carbonell A, Hope W, Warren J, Higgins R, Jacob B, *et al.* Robotic inguinal vs transabdominal laparoscopic inguinal hernia repair: the RIVAL randomized clinical trial. *JAMA Surg* 2020;**155**:380–7. https://doi.org/10.1001/jamasurg.2020.0034

169. van Rooijen MM, Jairam AP, Tollens T, Jørgensen LN, de Vries Reilingh TS, Piessen G, *et al.* Outcomes of a new slowly resorbable biosynthetic mesh (Phasix™) in potentially contaminated incisional hernias: a prospective, multi-center, single-arm trial. *Int J Surg* 2020;**83**:31–6. https://doi.org/10.1016/j.ijsu.2020.08.053

170. Langenbach MR, Enz D. Mesh fixation in open IPOM procedure with tackers or sutures? A randomized clinical trial with preliminary results. *Hernia* 2020;**24**:79–84. https://doi.org/10.1007/s10029-019-01991-z

171. Olavarria OA, Bernardi K, Shah SK, Wilson TD, Wei S, Pedroza C, *et al.* Robotic versus laparoscopic ventral hernia repair: multicenter, blinded randomized controlled trial. *BMJ* 2020;**370**:m2457. https://doi.org/10.1136/bmj.m2457

172. Wang D, Shen Y, Wang F, Chen J, Chen Y, Zhang Y. Mini-mesh and Lichtenstein repair compared with a modified Kugel technique for femoral hernia: a randomised controlled trial. *Ann R Coll Surg Engl* 2020;**102**:284–9. https://doi.org/10.1308/rcsann.2019.0181

173. Christoffersen MW, Westen M, Rosenberg J, Helgstrand F, Bisgaard T. Closure of the fascial defect during laparoscopic umbilical hernia repair: a randomized clinical trial [published online ahead of print]. *Br J Surg* 2020. https://doi.org/10.1002/bjs.11490

174. Bernardi K, Olavarria OA, Holihan JL, Kao LS, Ko TC, Roth JS, *et al.* Primary fascial closure during laparoscopic ventral hernia repair improves patient quality of life: a multicenter, blinded randomized controlled trial. *Ann Surg* 2020;**271**:434–9. https://doi.org/10.1097/SLA.0000000000003505

175. Wang D, Zhang H, Lei T, Chen J, Chen Y, Zhang Y, Qu P. Randomized trial comparing self-gripping mesh with polypropylene mesh in female Lichtenstein hernioplasty. *Am Surg* 2020;**86**:110–15. https://doi.org/10.1177/000313482008600229

176. Ahonen-Siirtola M, Nevala T, Vironen J, Kössi J, Pinta T, Niemeläinen S, *et al.* Laparoscopic versus hybrid approach for treatment of incisional ventral hernia: a prospective randomised multicentre study, 1-year results. *Surg Endosc* 2020;**34**:88–95. https://doi.org/10.1007/s00464-019-06735-9

177. Brans E, Reininga IHF, Balink H, Munzebrock AVE, Bessem B, de Graaf JS. Early recovery after endoscopic totally extraperitoneal (TEP) hernia repair in athletes with inguinal disruption: a prospective cohort study. *PLOS ONE* 2019;**14**:e0226011. https://doi.org/10.1371/journal.pone.0226011

178. Bernardi K, Martin AC, Holihan JL, Olavarria OA, Flores-Gonzalez JR, Cherla DV, *et al.* Is non-operative management warranted in ventral hernia patients with comorbidities? A case-matched, prospective 3 year follow-up, patient-centered study. *Am J Surg* 2019;**218**:1234–8. https://doi.org/10.1016/j.amjsurg.2019.07.044

179. Chen DC, Morrison J. State of the art: open mesh-based inguinal hernia repair. *Hernia* 2019;**23**:485–92. https://doi.org/10.1007/s10029-019-01983-z

180. Gutlic N, Gutlic A, Petersson U, Rogmark P, Montgomery A. Randomized clinical trial comparing total extraperitoneal with Lichtenstein inguinal hernia repair (TEPLICH trial). *Br J Surg* 2019;**106**:845–55. https://doi.org/10.1002/bjs.11230

181. Bullen NL, Massey LH, Antoniou SA, Smart NJ, Fortelny RH. Open versus laparoscopic mesh repair of primary unilateral uncomplicated inguinal hernia: a systematic review with meta-analysis and trial sequential analysis. *Hernia* 2019;**23**:461–72. https://doi.org/10.1007/s10029-019-01989-7

182. Peña ME, Dreifuss NH, Schlottmann F, Sadava EE. Could long-term follow-up modify the outcomes after laparoscopic TAPP? A 5-year retrospective cohort study. *Hernia* 2019;**23**:693–8. https://doi.org/10.1007/s10029-019-01953-5

183. Kaufmann R, Timmermans L, van Loon YT, Vroemen JPAM, Jeekel J, Lange JF. Repair of complex abdominal wall hernias with a cross-linked porcine acellular matrix: cross-sectional results of a Dutch cohort study. *Int J Surg* 2019;**65**:120–7. https://doi.org/10.1016/j.ijsu.2019.03.023

184. Thiyagarajan S, Velraj J, Hussain Ahmed MI, Murugesan R. Subarachnoid block with continuous TAP catheter analgesia produces less chronic pain and better functional outcome after inguinal hernioplasty: a randomized controlled observer-blinded study. *Reg Anesth Pain Med* 2019;**44**:228–33. https://doi.org/10.1136/rapm-2018-000029

185. Magnusson J, Gustafsson UO, Nygren J, Thorell A. Sustainability of the relationship between preoperative symptoms and postoperative improvement in quality of life after inguinal hernia repair. *Hernia* 2019;**23**:583–91. https://doi.org/10.1007/s10029-018-01875-8

186. Denham M, Johnson B, Leong M, Kuchta K, Conaty E, Ujiki MB, *et al.* An analysis of results in a single-blinded, prospective randomized controlled trial comparing non-fixating versus self-fixating mesh for laparoscopic inguinal hernia repair. *Surg Endosc* 2019;**33**:2670–9. https://doi.org/10.1007/s00464-018-6555-8

187. Feng MP, Baucom RB, Broman KK, Harris DA, Holzman MD, Huang LC, *et al.* Early repair of ventral incisional hernia may improve quality of life after surgery for abdominal malignancy: a prospective observational cohort study. *Hernia* 2019;**23**:81–90. https://doi.org/10.1007/s10029-018-1863-4

188. Schneeberger S, Phillips S, Huang LC, Pierce RA, Etemad SA, Poulose BK. Cost–utility analysis of biologic and biosynthetic mesh in ventral hernia repair: when are they worth it? *J Am Coll Surg* 2019;**228**:66–71. https://doi.org/10.1016/j.jamcollsurg.2018.10.009

189. Nouh T, Ali FS, Krause KJ, Zaimi I. Ventral hernia recurrence in women of childbearing age: a systematic review and meta-analysis. *Hernia* 2018;**22**:1067–75. https://doi.org/10.1007/s10029-018-1821-1

190. Georgiou E, Schoina E, Markantonis SL, Karalis V, Athanasopoulos PG, Chrysoheris P, *et al.* Laparoscopic total extraperitoneal inguinal hernia repair: retrospective study on prosthetic materials, postoperative management, and quality of life. *Medicine* 2018;**97**:e13974. https://doi.org/10.1097/MD.0000000000013974

191. van Rooijen MMJ, Jairam AP, Tollens T, Jørgensen LN, de Vries Reilingh TS, Piessen G, *et al.* A post-market, prospective, multi-center, single-arm clinical investigation of Phasix™ mesh for VHWG grade 3 midline incisional hernia repair: a research protocol. *BMC Surg* 2018;**18**:104. https://doi.org/10.1186/s12893-018-0439-7

192. Muysoms F, Van Cleven S, Pletinckx P, Ballecer C, Ramaswamy A. Robotic transabdominal retromuscular umbilical prosthetic hernia repair (TARUP): observational study on the operative time during the learning curve. *Hernia* 2018;**22**:1101–11. https://doi.org/10.1007/s10029-018-1825-x

193. Ielpo B, Nuñez-Alfonsel J, Duran H, Diaz E, Fabra I, Caruso R, *et al.* Cost-effectiveness of randomized study of laparoscopic versus open bilateral inguinal hernia repair. *Ann Surg* 2018;**268**:725–30. https://doi.org/10.1097/SLA.0000000000002894

194. Sheen AJ, Pilkington JJ, Baltatzis M, Tyurkylmaz A, Stathakis P, Jamdar S, Siriwardena AK. Comparison of mesh fixation techniques in elective laparoscopic repair of incisional Hernia-ReliaTack™ v ProTack™ (TACKoMesh) – a double-blind randomised controlled trial. *BMC Surg* 2018;**18**:46. https://doi.org/10.1186/s12893-018-0378-3

195. Roos MM, Verleisdonk EMM, Sanders FBM, Hoes AW, Stellato RK, Frederix GWJ, *et al.* Effectiveness of endoscopic totally extraperitoneal (TEP) hernia correction for clinically occult inguinal hernia (EFFECT): study protocol for a randomized controlled trial. *Trials* 2018;**19**:322. https://doi.org/10.1186/s13063-018-2711-7

196. Harsløf S, Krum-Møller P, Sommer T, Zinther N, Wara P, Friis-Andersen H. Effect of fixation devices on postoperative pain after laparoscopic ventral hernia repair: a randomized clinical trial of permanent tacks, absorbable tacks, and synthetic glue. *Langenbecks Arch Surg* 2018;**403**:529–37. https://doi.org/10.1007/s00423-018-1676-z

197. Mitura K, Garnysz K, Wyrzykowska D, Michałek I. The change in groin pain perception after transabdominal preperitoneal inguinal hernia repair with glue fixation: a prospective trial of a single surgeon's experience. *Surg Endosc* 2018;**32**:4284–9. https://doi.org/10.1007/s00464-018-6178-0

198. Rouet J, Bwelle G, Cauchy F, Masso-Misse P, Gaujoux S, Dousset B. Polyester mosquito net mesh for inguinal hernia repair: a feasible option in resource limited settings in Cameroon? *J Visc Surg* 2018;**155**:111–16. https://doi.org/10.1016/j.jviscsurg.2017.10.006

199. Roth JS, Anthone GJ, Selzer DJ, Poulose BK, Bittner JG, Hope WW, *et al*. Prospective evaluation of poly-4-hydroxybutyrate mesh in CDC class I/high-risk ventral and incisional hernia repair: 18-month follow-up. *Surg Endosc* 2018;**32**:1929–36. https://doi.org/10.1007/s00464-017-5886-1

200. Slooter GD, Zwaans WAR, Perquin CW, Roumen RMH, Scheltinga MRM. Laparoscopic mesh removal for otherwise intractable inguinal pain following endoscopic hernia repair is feasible, safe and may be effective in selected patients. *Surg Endosc* 2018;**32**:1613–19. https://doi.org/10.1007/s00464-017-5824-2

201. Cherla DV, Moses ML, Viso CP, Holihan JL, Flores-Gonzalez JR, Kao LS, *et al*. Impact of abdominal wall hernias and repair on patient quality of life. *World J Surg* 2018;**42**:19–25. https://doi.org/10.1007/s00268-017-4173-6

202. Bona S, Rosati R, Opocher E, Fiore B, Montorsi M, SUPERMESH Study Group. Pain and quality of life after inguinal hernia surgery: a multicenter randomized controlled trial comparing lightweight vs heavyweight mesh (Supermesh Study). *Updates Surg* 2018;**70**:77–83. https://doi.org/10.1007/s13304-017-0483-3

203. Ielpo B, Duran H, Diaz E, Fabra I, Caruso R, Malavé L, *et al*. A prospective randomized study comparing laparoscopic transabdominal preperitoneal (TAPP) versus Lichtenstein repair for bilateral inguinal hernias. *Am J Surg* 2018;**216**:78–83. https://doi.org/10.1016/j.amjsurg.2017.07.016

204. de Goede B, Wijsmuller AR, van Ramshorst GH, van Kempen BJ, Hop WCJ, Klitsie PJ, *et al*. Watchful waiting versus surgery of mildly symptomatic or asymptomatic inguinal hernia in men aged 50 years and older: a randomized controlled trial. *Ann Surg* 2018;**267**:42–9. https://doi.org/10.1097/SLA.0000000000002243

205. Sun P, Cheng X, Deng S, Hu Q, Sun Y, Zheng Q. Mesh fixation with glue versus suture for chronic pain and recurrence in Lichtenstein inguinal hernioplasty. *Cochrane Database Syst Rev* 2017;**2**:CD010814. https://doi.org/10.1002/14651858.CD010814.pub2

206. Lederhuber H, Stiede F, Axer S, Dahlstrand U. Mesh fixation in endoscopic inguinal hernia repair: evaluation of methodology based on a systematic review of randomised clinical trials. *Surg Endosc* 2017;**31**:4370–81. https://doi.org/10.1007/s00464-017-5509-x

207. Koju R, Koju RB, Malla B, Dongol Y, Thapa LB. Transabdominal pre-peritoneal mesh repair versus Lichtenstein's hernioplasty. *J Nepal Health Res Counc* 2017;**15**:135–40. https://doi.org/10.3126/jnhrc.v15i2.18202

208. Kushwaha JK, Enny LE, Anand A, Sonkar AA, Kumar A, Pahwa HS. A prospective randomized controlled trial comparing quality of life following endoscopic totally extraperitoneal (TEP) versus open stoppa inguinal hernioplasty. *Surg Laparosc Endosc Percutan Tech* 2017;**27**:257–61. https://doi.org/10.1097/SLE.0000000000000450

209. Jensen KK. Recovery after abdominal wall reconstruction. *Dan Med J* 2017;**64**:B5349.

210. Molegraaf MJ, Grotenhuis B, Torensma B, de Ridder V, Lange JF, Swank DJ. The HIPPO trial, a randomized double-blind trial comparing self-gripping parietex progrip mesh and sutured parietex mesh in lichtenstein hernioplasty: a long-term follow-up study. *Ann Surg* 2017;**266**:939–45. https://doi.org/10.1097/SLA.0000000000002169

211. Löfgren J, Matovu A, Wladis A, Ibingira C, Nordin P, Galiwango E, Forsberg BC. Cost-effectiveness of groin hernia repair from a randomized clinical trial comparing commercial versus low-cost mesh in a low-income country. *Br J Surg* 2017;**104**:695–703. https://doi.org/10.1002/bjs.10483

REFERENCES

212. Čadanová D, van Dijk JP, Mollen RMHG. The transinguinal preperitoneal technique (TIPP) in inguinal hernia repair does not cause less chronic pain in relation to the ProGrip technique: a prospective double-blind randomized clinical trial comparing the TIPP technique, using the PolySoft mesh, with the ProGrip self-fixing semi-resorbable mesh. *Hernia* 2017;**21**:17–27. https://doi.org/10.1007/s10029-016-1522-6

213. Bansal VK, Krishna A, Manek P, Kumar S, Prajapati O, Subramaniam R, *et al.* A prospective randomized comparison of testicular functions, sexual functions and quality of life following laparoscopic totally extra-peritoneal (TEP) and trans-abdominal pre-peritoneal (TAPP) inguinal hernia repairs. *Surg Endosc* 2017;**31**:1478–86. https://doi.org/10.1007/s00464-016-5142-0

214. Fischer JP, Basta MN, Krishnan NM, Wink JD, Kovach SJ. A cost–utility assessment of mesh selection in clean-contaminated ventral hernia repair. *Plast Reconstr Surg* 2016;**137**:647–59. https://doi.org/10.1097/01.prs.0000475775.44891.56

215. Bansal VK, Asuri K, Panaiyadiyan S, Kumar S, Subramaniam R, Ramachandran R, *et al.* Comparison of absorbable versus nonabsorbable tackers in terms of long-term outcomes, chronic pain, and quality of life after laparoscopic incisional hernia repair: a randomized study. *Surg Laparosc Endosc Percutan Tech* 2016;**26**:476–83. https://doi.org/10.1097/SLE.0000000000000347

216. Emanuelsson P, Gunnarsson U, Dahlstrand U, Strigård K, Stark B. Operative correction of abdominal rectus diastasis (ARD) reduces pain and improves abdominal wall muscle strength: a randomized, prospective trial comparing retromuscular mesh repair to double-row, self-retaining sutures. *Surgery* 2016;**160**:1367–75. https://doi.org/10.1016/j.surg.2016.05.035

217. Muysoms FE, Vanlander A, Ceulemans R, Kyle-Leinhase I, Michiels M, Jacobs I, *et al.* A prospective, multicenter, observational study on quality of life after laparoscopic inguinal hernia repair with ProGrip laparoscopic, self-fixating mesh according to the European Registry for Abdominal Wall Hernias Quality of Life Instrument. *Surgery* 2016;**160**:1344–57. https://doi.org/10.1016/j.surg.2016.04.026

218. Gillion JF, Fromont G, Lepère M, Letoux N, Dabrowski A, Zaranis C, Barrat C, Hernia-Club Members. Laparoscopic ventral hernia repair using a novel intraperitoneal lightweight mesh coated with hyaluronic acid: 1-year follow-up from a case-control study using the Hernia-Club registry. *Hernia* 2016;**20**:711–22. https://doi.org/10.1007/s10029-016-1501-y

219. Magnusson J, Nygren J, Gustafsson UO, Thorell A. UltraPro Hernia System, Prolene Hernia System and Lichtenstein for primary inguinal hernia repair: 3-year outcomes of a prospective randomized controlled trial. *Hernia* 2016;**20**:641–8. https://doi.org/10.1007/s10029-016-1507-5

220. Choi BJ, Jeong WJ, Lee IK, Lee SC. Single-port versus conventional three-port laparoscopic totally extraperitoneal inguinal hernia repair: a randomized controlled trial. *Hernia* 2016;**20**:789–95. https://doi.org/10.1007/s10029-016-1499-1

221. Handojo K, Meylemans D, Devroe K, Vermeiren K, Aelvoet C, Tollens T. Initial experience with a new macroporous partially absorbable mesh: introducing Ultrapro® Advanced™. *Surg Technol Int* 2016;**28**:125–30.

222. Tobler WD, Itani KM. Current Status and challenges of laparoscopy in ventral hernia repair. *J Laparoendosc Adv Surg Tech A* 2016;**26**:281–9. https://doi.org/10.1089/lap.2016.0095

223. Berney CR, Descallar J. Review of 1000 fibrin glue mesh fixation during endoscopic totally extraperitoneal (TEP) inguinal hernia repair. *Surg Endosc* 2016;**30**:4544–52. https://doi.org/10.1007/s00464-016-4791-3

224. Burgmans JP, Voorbrood CE, Simmermacher RK, Schouten N, Smakman N, Clevers G, *et al.* Long-term results of a randomized double-blinded prospective trial of a lightweight (Ultrapro) versus a heavyweight mesh (Prolene) in laparoscopic total extraperitoneal inguinal hernia repair (TULP-trial). *Ann Surg* 2016;**263**:862–6. https://doi.org/10.1097/SLA.0000000000001579

225. Nedelcu M, Verhaeghe P, Skalli M, Champault G, Barrat C, Sebbag H, *et al.* Multicenter prospective randomized study comparing the technique of using a bovine pericardium biological prosthesis reinforcement in parietal herniorrhaphy (Tutomesh TUTOGEN) with simple parietal herniorrhaphy, in a potentially contaminated setting. *Wound Repair Regen* 2016;**24**:427–33. https://doi.org/10.1111/wrr.12386

226. Yazicioğlu D, Caparlar C, Akkaya T, Mercan U, Kulaçoğlu H. Tizanidine for the management of acute postoperative pain after inguinal hernia repair: a placebo-controlled double-blind trial. *Eur J Anaesthesiol* 2016;**33**:215–22. https://doi.org/10.1097/EJA.0000000000000371

227. Wang Y, Zhang X. Short-term results of open inguinal hernia repair with self-gripping Parietex ProGrip mesh in China: a retrospective study of 90 cases. *Asian J Surg* 2016;**39**:218–24. https://doi.org/10.1016/j.asjsur.2015.05.001

228. Rogmark P, Petersson U, Bringman S, Ezra E, Österberg J, Montgomery A. Quality of life and surgical outcome 1 year after open and laparoscopic incisional hernia repair: PROLOVE: a randomized controlled trial. *Ann Surg* 2016;**263**:244–50. https://doi.org/10.1097/SLA.0000000000001305

229. Rönkä K, Vironen J, Kössi J, Hulmi T, Silvasti S, Hakala T, *et al.* Randomized multicenter trial comparing glue fixation, self-gripping mesh, and suture fixation of mesh in lichtenstein hernia repair (FinnMesh study). *Ann Surg* 2015;**262**:714–19. https://doi.org/10.1097/SLA.0000000000001458

230. Christoffersen MW, Olsen BH, Rosenberg J, Bisgaard T. Randomized clinical trial on the postoperative use of an abdominal binder after laparoscopic umbilical and epigastric hernia repair. *Hernia* 2015;**19**:147–53. https://doi.org/10.1007/s10029-014-1289-6

231. Chatterjee A, Krishnan NM, Rosen JM. Complex ventral hernia repair using components separation with or without biologic mesh: a cost–utility analysis. *Ann Plast Surg* 2015;**74**:471–8. https://doi.org/10.1097/SAP.0b013e31829fd306

232. Chatterjee A, Krishnan NM, Rosen JM. Complex ventral hernia repair using components separation with or without synthetic mesh: a cost–utility analysis. *Plast Reconstr Surg* 2014;**133**:137–46. https://doi.org/10.1097/01.prs.0000436835.96194.79

233. Berrevoet F, Tollens T, Berwouts L, Bertrand C, Muysoms F, De Gols J, *et al.* A Belgian multicenter prospective observational cohort study shows safe and efficient use of a composite mesh with incorporated oxidized regenerated cellulose in laparoscopic ventral hernia repair. *Acta Chir Belg* 2014;**114**:233–8. https://doi.org/10.1080/00015458.2014.11681018

234. Greco DP, Fei L, Guerriero L, Pradella P, Mazzola M, Magistro C, *et al.* Feasibility and effectiveness of primary umbilical hernia repair with biologic graft: preliminary study. *Acta Chir Belg* 2014;**114**:125–30. https://doi.org/10.1080/00015458.2014.11680994

235. Christoffersen MW, Westen M, Assadzadeh S, Deigaard SL, Rosenberg J, Bisgaard T. The clinical effects of closure of the hernia gap after laparoscopic ventral hernia repair: protocol for a randomised controlled trial. *Dan Med J* 2014;**61**:A4865.

236. Stey AM, Danzig M, Qiu S, Yin S, Divino CM. Cost–utility analysis of repair of reducible ventral hernia. *Surgery* 2014;**155**:1081–9. https://doi.org/10.1016/j.surg.2014.03.041

237. Nikkolo C, Vaasna T, Murruste M, Seepter H, Kirsimägi Ü, Lepner U. Randomized clinical study evaluating the impact of mesh pore size on chronic pain after Lichtenstein hernioplasty. *J Surg Res* 2014;**191**:311–17. https://doi.org/10.1016/j.jss.2014.04.022

238. Abdalla RZ, Garcia RB, Said DF, Abdalla BM. Quality of life of in patients submitted to anterior abdominal wall laparoscopic hernioplasty. *Arq Bras Cir Dig* 2014;**27**:30–3. https://doi.org/10.1590/s0102-67202014000100008

239. Bensaadi H, Paolino L, Valenti A, Polliand C, Barrat C, Champault G. Intraperitoneal tension-free repair of a small midline ventral abdominal wall hernia: randomized study with a mean follow-up of 3 years. *Am Surg* 2014;**80**:57–65. https://doi.org/10.1177/000313481408000125

240. Dhankhar DS, Sharma N, Mishra T, Kaur N, Singh S, Gupta S. Totally extraperitoneal repair under general anesthesia versus Lichtenstein repair under local anesthesia for unilateral inguinal hernia: a prospective randomized controlled trial. *Surg Endosc* 2014;**28**:996–1002. https://doi.org/10.1007/s00464-013-3269-9

241. Chan MS, Melissa CS, Teoh AY, Bun TA, Chan KW, Wing CK, *et al.* Randomized double-blinded prospective trial of fibrin sealant spray versus mechanical stapling in laparoscopic total extraperitoneal hernioplasty. *Ann Surg* 2014;**259**:432–7. https://doi.org/10.1097/SLA.0b013e3182a6c513

242. Bellows CF, Shadduck P, Helton WS, Martindale R, Stouch BC, Fitzgibbons R. Early report of a randomized comparative clinical trial of Strattice™ reconstructive tissue matrix to lightweight synthetic mesh in the repair of inguinal hernias. *Hernia* 2014;**18**:221–30. https://doi.org/10.1007/s10029-013-1076-9

243. Peeters E, Spiessens C, Oyen R, De Wever L, Vanderschueren D, Penninckx F, Miserez M. Sperm motility after laparoscopic inguinal hernia repair with lightweight meshes: 3-year follow-up of a randomised clinical trial. *Hernia* 2014;**18**:361–7. https://doi.org/10.1007/s10029-012-1028-9

244. Stabilini C, Bracale U, Pignata G, *et al.* Laparoscopic bridging vs. anatomic open reconstruction for midline abdominal hernia mesh repair [LABOR]: single-blinded, multicenter, randomized, controlled trial on long-term functional results. *Trials* 2013;**14**:357. https://doi.org/10.1186/1745-6215-14-357

245. Calò PG, Pittau MR, Contu P, D'Aloja E, Nicolosi A, Demontis R. Chronic pain following inguinal hernia repair: assessment of quality of life and medico-legal aspects. *Ann Ital Chir* 2013;**84**:357–63.

246. Mariette C, Briez N, Denies F, Dervaux B, Duhamel A, Guilbert M, *et al.* Use of biological mesh versus standard wound care in infected incisional ventral hernias, the SIMBIOSE study: a study protocol for a randomized multicenter controlled trial. *Trials* 2013;**14**:131. https://doi.org/10.1186/1745-6215-14-131

247. Coronini-Cronberg S, Appleby J, Thompson J. Application of patient-reported outcome measures (PROMs) data to estimate cost-effectiveness of hernia surgery in England. *J R Soc Med* 2013;**106**:278–87. https://doi.org/10.1177/0141076813489679

248. Eriksen JR, Bisgaard T, Assaadzadeh S, Jorgensen LN, Rosenberg J. Fibrin sealant for mesh fixation in laparoscopic umbilical hernia repair: 1-year results of a randomized controlled double-blinded study. *Hernia* 2013;**17**:511–14. https://doi.org/10.1007/s10029-013-1101-z

249. Rogmark P, Petersson U, Bringman S, Eklund A, Ezra E, Sevonius D, *et al.* Short-term outcomes for open and laparoscopic midline incisional hernia repair: a randomized multicenter controlled trial: the ProLOVE (prospective randomized trial on open versus laparoscopic operation of ventral eventrations) trial. *Ann Surg* 2013;**258**:37–45. https://doi.org/10.1097/SLA.0b013e31828fe1b2

250. Muysoms F, Vander Mijnsbrugge G, Pletinckx P, Boldo E, Jacobs I, Michiels M, Ceulemans R. Randomized clinical trial of mesh fixation with 'double crown' versus 'sutures and tackers' in laparoscopic ventral hernia repair. *Hernia* 2013;**17**:603–12. https://doi.org/10.1007/s10029-013-1084-9

251. Bansal VK, Misra MC, Babu D, Victor J, Kumar S, Sagar R, *et al.* A prospective, randomized comparison of long-term outcomes: chronic groin pain and quality of life following totally extraperitoneal (TEP) and transabdominal preperitoneal (TAPP) laparoscopic inguinal hernia repair. *Surg Endosc* 2013;**27**:2373–82. https://doi.org/10.1007/s00464-013-2797-7

252. Koning GG, Adang EM, Stalmeier PF, Keus F, Vriens PW, van Laarhoven CJ. TIPP and Lichtenstein modalities for inguinal hernia repair: a cost minimisation analysis alongside a randomised trial. *Eur J Health Econ* 2013;**14**:1027–34. https://doi.org/10.1007/s10198-012-0453-0

253. Jorgensen LN, Sommer T, Assaadzadeh S, Strand L, Dorfelt A, Hensler M, Rosenberg J, Danish Multicentre DANGRIP Study Group. Randomized clinical trial of self-gripping mesh versus sutured mesh for Lichtenstein hernia repair. *Br J Surg* 2013;**100**:474–81. https://doi.org/10.1002/bjs.9006

254. Yazdankhah Kenary A, Afshin SN, Ahmadi Amoli H, Yagoobi Notash A, Borjian A, Yagoobi Notash A, *et al.* Randomized clinical trial comparing lightweight mesh with heavyweight mesh for primary inguinal hernia repair. *Hernia* 2013;**17**:471–7. https://doi.org/10.1007/s10029-012-1009-z

255. Lepski G, Vahedi P, Tatagiba MS, Morgalla M. Combined spinal cord and peripheral nerve field stimulation for persistent post-herniorrhaphy pain. *Neuromodulation* 2013;**16**:84–8. https://doi.org/10.1111/j.1525-1403.2012.00463.x

256. Bradley JF, Williams KB, Wormer BA, Tsirline VB, Walters AL, Sing RF, *et al.* Preliminary results of surgical and quality of life outcomes of Physiomesh in an international, prospective study. *Surg Technol Int* 2012;**22**:113–19.

257. Lermite E, Arnaud JP. Prospective randomized study comparing quality of life after shoudice or mesh plug repair for inguinal hernia: short-term results. *Surg Technol Int* 2012;**22**:101–6.

258. Bignell M, Partridge G, Mahon D, Rhodes M. Prospective randomized trial of laparoscopic (transabdominal preperitoneal-TAPP) versus open (mesh) repair for bilateral and recurrent inguinal hernia: incidence of chronic groin pain and impact on quality of life: results of 10 year follow-up. *Hernia* 2012;**16**:635–40. https://doi.org/10.1007/s10029-012-0940-3

259. Mutter D, Champault G, Binot D, Vix M, Leroy J, Marescaux J. PerFix TM plug versus 4DDOME® implants for inguinal hernia repair: prospective multicentric randomised controlled trial. *Hernia* 2012;**16**:561–6. https://doi.org/10.1007/s10029-012-0943-0

260. Bansal VK, Misra MC, Babu D, Singhal P, Rao K, Sagar R, *et al.* Comparison of long-term outcome and quality of life after laparoscopic repair of incisional and ventral hernias with suture fixation with and without tacks: a prospective, randomized, controlled study. *Surg Endosc* 2012;**26**:3476–85. https://doi.org/10.1007/s00464-012-2390-5

261. Schouten N, van Dalen T, Smakman N, Elias SG, Clevers GJ, Verleisdonk EJ, *et al.* The effect of ultrapro or prolene mesh on postoperative pain and well-being following endoscopic totally extraperitoneal (TEP) hernia repair (TULP): study protocol for a randomized controlled trial. *Trials* 2012;**13**:76. https://doi.org/10.1186/1745-6215-13-76

262. Eker HH, Langeveld HR, Klitsie PJ, van't Riet M, Stassen LP, Weidema WF, *et al.* Randomized clinical trial of total extraperitoneal inguinal hernioplasty vs Lichtenstein repair: a long-term follow-up study. *Arch Surg* 2012;**147**:256–60. https://doi.org/10.1001/archsurg.2011.2023

263. Magnusson J, Nygren J, Thorell A. Lichtenstein, prolene hernia system, and UltraPro Hernia System for primary inguinal hernia repair: one-year outcome of a prospective randomized controlled trial. *Hernia* 2012;**16**:277–85. https://doi.org/10.1007/s10029-012-0903-8

264. Fortelny RH, Petter-Puchner AH, May C, Jaksch W, Benesch T, Khakpour Z, *et al.* The impact of atraumatic fibrin sealant vs. staple mesh fixation in TAPP hernia repair on chronic pain and quality of life: results of a randomized controlled study. *Surg Endosc* 2012;**26**:249–54. https://doi.org/10.1007/s00464-011-1862-3

265. Yener O, Aksoy F, Güzel P, Bölük S, Dağ E, Atak T. Long-term quality of life after hernioplasty using a Prolene hernia system in adult inguinal hernia. *Hernia* 2012;**16**:29–32. https://doi.org/10.1007/s10029-011-0855-4

266. Sadowski B, Rodriguez J, Symmonds R, Roberts J, Song J, Rajab MH, *et al.* Comparison of polypropylene versus polyester mesh in the Lichtenstein hernia repair with respect to chronic pain and discomfort. *Hernia* 2011;**15**:643–54. https://doi.org/10.1007/s10029-011-0841-x

267. Tollens T, Den Hondt M, Devroe K, Terry C, Speybroeck S, Aelvoet C, *et al.* Retrospective analysis of umbilical, epigastric, and small incisional hernia repair using the Ventralex TM hernia patch. *Hernia* 2011;**15**:531–40. https://doi.org/10.1007/s10029-011-0816-y

268. van den Heuvel B, Dwars BJ, Klassen DR, Bonjer HJ. Is surgical repair of an asymptomatic groin hernia appropriate? A review. *Hernia* 2011;**15**:251–9. https://doi.org/10.1007/s10029-011-0796-y

269. Nikkolo C, Lepner U, Murruste M, Vaasna T, Seepter H, Tikk T. Randomised clinical trial comparing lightweight mesh with heavyweight mesh for inguinal hernioplasty. *Hernia* 2010;**14**:253–8. https://doi.org/10.1007/s10029-010-0630-y

270. śmietański M, Polish Hernia Study Group. Randomized clinical trial comparing a polypropylene with a poliglecaprone and polypropylene composite mesh for inguinal hernioplasty. *Br J Surg* 2008;**95**:1462–8. https://doi.org/10.1002/bjs.6383

271. Turaga K, Wright A, Lee R, Dias WP, Destache C, Christian R, Fitzgibbons RJ. A randomized trial of the peri-operative use of COX-2 inhibitors in Lichtenstein herniorrhaphy. *Hernia* 2008;**12**:515–19. https://doi.org/10.1007/s10029-008-0379-8

272. Paajanen H. A single-surgeon randomized trial comparing three composite meshes on chronic pain after Lichtenstein hernia repair in local anesthesia. *Hernia* 2007;**11**:335–9. https://doi.org/10.1007/s10029-007-0236-1

273. Hynes DM, Stroupe KT, Luo P, Giobbie-Hurder A, Reda D, Kraft M, *et al.* Cost effectiveness of laparoscopic versus open mesh hernia operation: results of a Department of Veterans Affairs randomized clinical trial. *J Am Coll Surg* 2006;**203**:447–57. https://doi.org/10.1016/j.jamcollsurg.2006.05.019

274. Pokorny H, Klingler A, Scheyer M, Függer R, Bischof G. Postoperative pain and quality of life after laparoscopic and open inguinal hernia repair: results of a prospective randomized trial. *Hernia* 2006;**10**:331–7. https://doi.org/10.1007/s10029-006-0105-3

275. Heikkinen T, Wollert S, Osterberg J, Smedberg S, Bringman S. Early results of a randomised trial comparing Prolene and VyproII-mesh in endoscopic extraperitoneal inguinal hernia repair (TEP) of recurrent unilateral hernias. *Hernia* 2006;**10**:34–40. https://doi.org/10.1007/s10029-005-0026-6

276. Bringman S, Wollert S, Osterberg J, Heikkinen T. Early results of a randomized multicenter trial comparing Prolene and VyproII mesh in bilateral endoscopic extraperitoneal hernioplasty (TEP). *Surg Endosc* 2005;**19**:536–40. https://doi.org/10.1007/s00464-004-9100-x

277. Nienhuijs SW, van Oort I, Keemers-Gels ME, Strobbe LJ, Rosman C. Randomized trial comparing the Prolene hernia system, mesh plug repair and Lichtenstein method for open inguinal hernia repair. *Br J Surg* 2005;**92**:33–8. https://doi.org/10.1002/bjs.4702

278. Post S, Weiss B, Willer M, Neufang T, Lorenz D. Randomized clinical trial of lightweight composite mesh for Lichtenstein inguinal hernia repair. *Br J Surg* 2004;**91**:44–8. https://doi.org/10.1002/bjs.4387

279. Fitzgibbons RJ, Jonasson O, Gibbs J, Dunlop DD, Henderson W, Reda D, *et al.* The development of a clinical trial to determine if watchful waiting is an acceptable alternative to routine herniorrhaphy for patients with minimal or no hernia symptoms. *J Am Coll Surg* 2003;**196**:737–42. https://doi.org/10.1016/S1072-7515(03)00003-6

280. Sarli L, Iusco DR, Sansebastiano G, Costi R. Simultaneous repair of bilateral inguinal hernias: a prospective, randomized study of open, tension-free versus laparoscopic approach. *Surg Laparosc Endosc Percutan Tech* 2001;**11**:262–7. https://doi.org/10.1097/00129689-200108000-00007

281. Liem MS, Halsema JA, van der Graaf Y, Schrijvers AJ, van Vroonhoven TJ, Coala trial group. Cost-effectiveness of extraperitoneal laparoscopic inguinal hernia repair: a randomized comparison with conventional herniorrhaphy. *Ann Surg* 1997;**226**:668–75. https://doi.org/10.1097/00000658-199712000-00004

282. Kawabata R, Fujitani K, Sakamaki K, Ando M, Ito Y, Tanizawa Y, *et al.* Survival analysis of a prospective multicenter observational study on surgical palliation among patients with malignant bowel obstruction caused by peritoneal dissemination of gastric cancer. *Gastric Cancer* 2022;**25**:422–9. https://doi.org/10.1007/s10120-021-01251-z

283. Ito Y, Fujitani K, Sakamaki K, Ando M, Kawabata R, Tanizawa Y, *et al.* QOL assessment after palliative surgery for malignant bowel obstruction caused by peritoneal dissemination of gastric cancer: a prospective multicenter observational study. *Gastric Cancer* 2021;**24**:1131–9. https://doi.org/10.1007/s10120-021-01179-4

284. Tabusa H, Blazeby JM, Blencowe N, Callaway M, Daniels IR, Gunning A, *et al.* Protocol for the UK cohort study to investigate the prevention of parastomal hernia (the CIPHER study). *Colorectal Dis* 2021;**23**:1900–8. https://doi.org/10.1111/codi.15621

Appendix 1 Diagnostic subcategories included in the cohorts

TABLE 21 Diagnostic subcategories included in the cohorts

Condition	ICD-10 code	Description	n (%)
Appendicitis	K35	Acute appendicitis	14,794 (5.5)
	K35.2	Acute appendicitis with generalized peritonitis	10,475 (3.9)
	K35.3	Acute appendicitis with localized peritonitis	59,434 (22.2)
	K35.8	Acute appendicitis, other and unspecified	147,081 (54.9)
	K37	Unspecified appendicitis	36,360 (13.6)
Diverticular disease	K57.2	Diverticular disease of large intestine with perforation and abscess	32,657 (23.5)
	K57.3	Diverticular disease of large intestine without perforation or abscess	106,212 (76.5)
Cholelithiasis	K80.0	Calculus of gallbladder with acute cholecystitis	86,878 (36.1)
	K80.1	Calculus of gallbladder with other cholecystitis	67,503 (28.0)
	K80.2	Calculus of gallbladder without cholecystitis	86,596 (35.9)
Hernia	K40.0	Bilateral inguinal hernia, with obstruction, without gangrene	957 (0.1)
	K40.1	Bilateral inguinal hernia, with gangrene	52 (0.0)
	K40.2	Bilateral inguinal hernia, without obstruction or gangrene	1987 (1.9)
	K40.3	Unilateral or unspecified inguinal hernia, with obstruction, without gangrene	17,177 (16.1)
	K40.4	Unilateral or unspecified inguinal hernia, with gangrene	730 (0.7)
	K40.9	Unilateral or unspecified inguinal hernia, without obstruction or gangrene	30,108 (28.3)
	K41.0	Bilateral femoral hernia, with obstruction, without gangrene	263 (0.2)
	K41.1	Bilateral femoral hernia, with gangrene	37 (0.0)
	K41.2	Bilateral femoral hernia, without obstruction or gangrene	55 (0.1)
	K41.3	Unilateral or unspecified femoral hernia, with obstruction, without gangrene	8491 (8.0)
	K41.4	Unilateral or unspecified femoral hernia, with gangrene	1148 (1.1)
	K41.9	Unilateral or unspecified femoral hernia, without obstruction or gangrene	3460 (3.3)
	K42.0	Umbilical hernia with obstruction, without gangrene	18,324 (17.2)
	K42.1	Umbilical hernia with gangrene	1245 (1.2)
	K42.9	Umbilical hernia without obstruction or gangrene	20,235 (19.0)
	K43.6	Other and unspecified ventral hernia with obstruction, without gangrene	2066 (1.9)
	K43.7	Other and unspecified ventral hernia with gangrene	97 (0.1)
Intestinal obstruction	K56.1	Intussusception	1465 (1.1)
	K56.2	Volvulus	16,126 (12.1)
	K56.3	Gallstone ileus	2027 (1.5)
	K56.5	Intestinal adhesions [bands] with obstruction	46,061 (34.6)
	K56.6	Other and unspecified intestinal obstruction	67,394 (50.6)

Appendix 2 Additional notes on the instrumental variable (tendency to operate) and statistical analysis methods

Tendency to operate as an instrumental variable

The ESORT study uses a preference-based IV design, with the IV as the hospital's TTO.[19] The TTO is defined for each qualifying emergency admission, as the proportion of eligible emergency admissions in the 12 months prior to each admission in the specific hospital, who received ES rather than NES strategies (*Figure 28*). A precedent study by Keele *et al.* took a similar approach, using TTO at the surgeon level, to evaluate ES compared with NES strategies using US claims data.[20] A recent systematic review identified 185 studies that applied preference-based IV methods within health research and reported that such methods were most commonly applied to cancer, cardiovascular diseases and mental health.[90] As noted by the authors, a valid IV must (1) strongly predict treatment status ('relevance'), (2) affect outcomes only through exposure ('exclusion'), (3) not share any unmeasured causes with the outcome ('unconfoundedness') and (4) affect treatment status in only one direction ('monotonocity').[90]
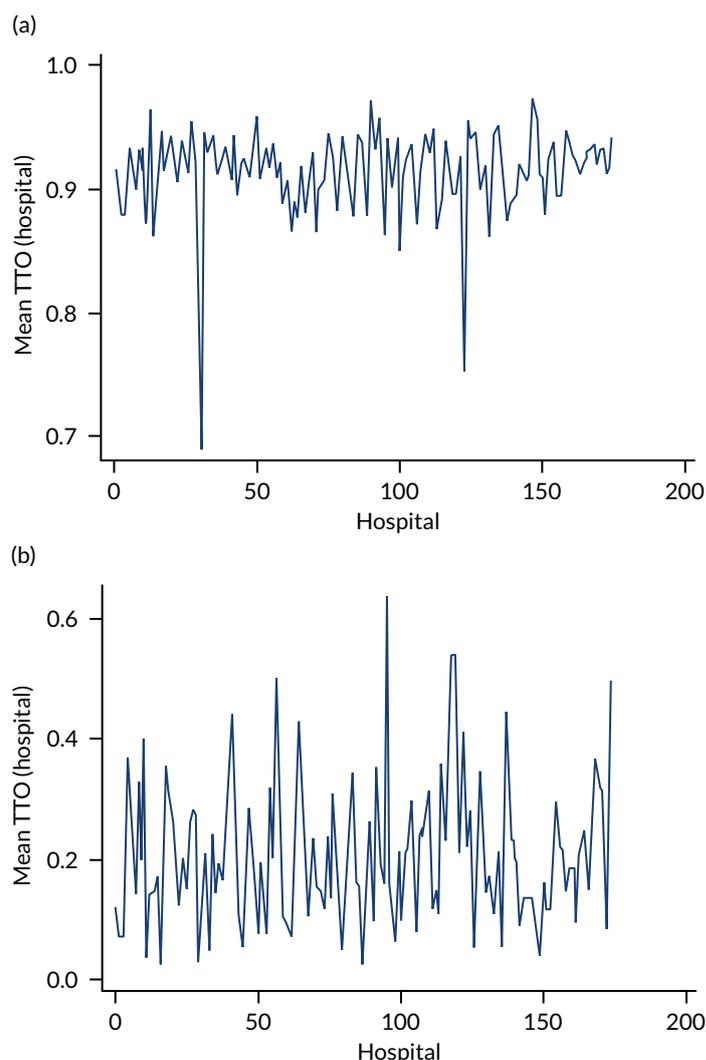
FIGURE 28 Variation in the TTO across 175 NHS hospitals in the 1 year prior to emergency hospitals admissions that meet the inclusion criteria for each of the five acute conditions (2009–19). (a) Acute appendicitis; (b) cholelithiasis; (c) diverticular disease; (d) abdominal wall hernia; and (e) intestinal obstruction. (*continued*)

(c)

(d)

(e)
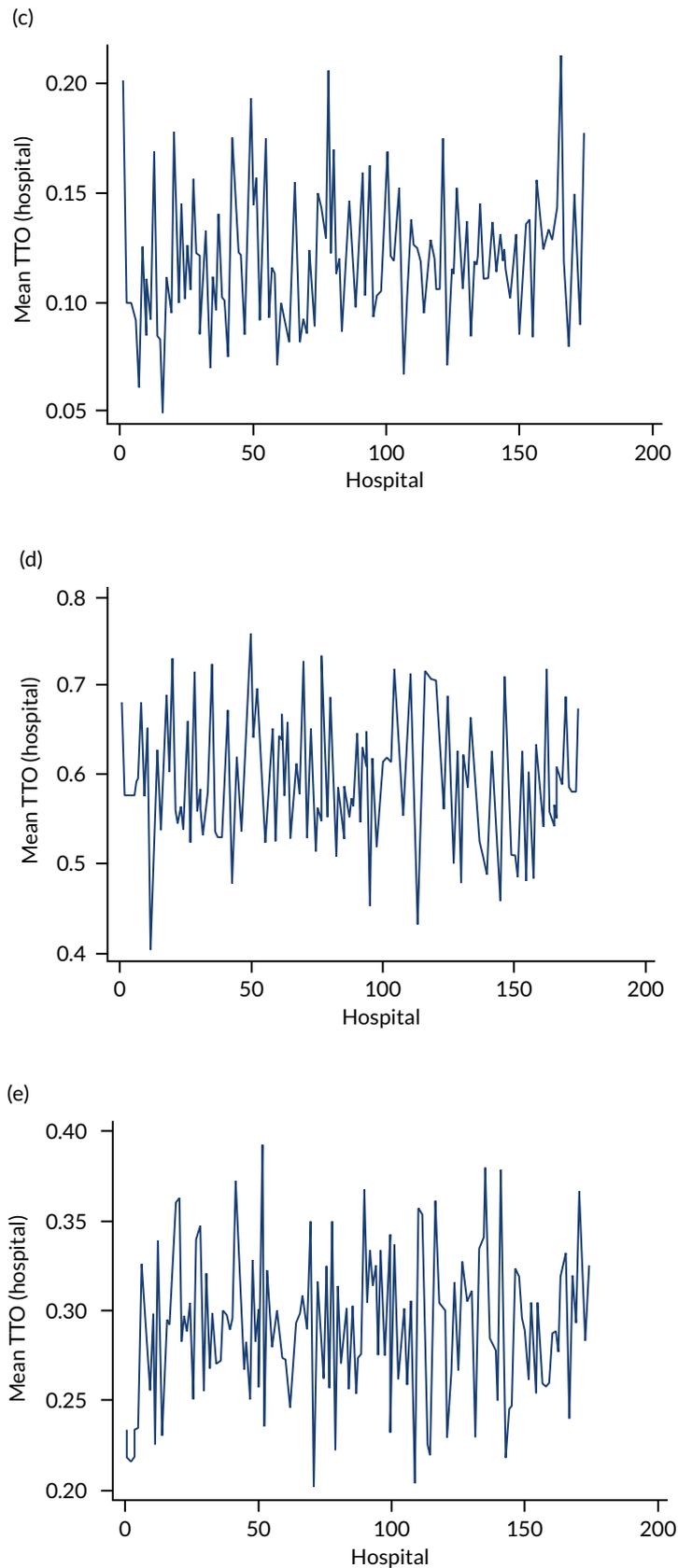
FIGURE 28 Variation in the TTO across 175 NHS hospitals in the 1 year prior to emergency hospitals admissions that meet the inclusion criteria for each of the five acute conditions (2009–19). (a) Acute appendicitis; (b) cholelithiasis; (c) diverticular disease; (d) abdominal wall hernia; and (e) intestinal obstruction.

Hospitals that differ in their TTO may be expected to differ in the probability that they will provide ES to a particular patient. We find considerable variation across hospitals in their TTO, even after adjusting for observable characteristics of the patients. The hospital's TTO was strongly correlated with ES receipt for each of the five conditions, even after controlling for a rich set of covariates, with F-statistics that ranged from 450 (diverticular disease) to 24,517 (cholelithiasis) (*Table 22*). A commonly applied threshold for 'weak' instruments is an F-statistic less than 10.[53] Therefore, the hospital's past preference for ES strongly predicts treatment choice for the current patient.

The validity of the IV analyses rests on the assumption that, conditional on the included variables, the TTO does not influence the outcomes except through changing the uptake of ES [i.e. assumptions 1 and 2. See *Chapter 3*, *Sensitivity analyses* for further information]. To increase the plausibility of this assumption, we adjust for fixed effects (indicators) for each financial year, observable patient characteristics and proxies for hospital quality, a key unobservable confounder. Including hospital fixed effects would result in greatly inflated standard errors, as they would remove all between-hospital variation. We did not include these fixed effects, as they aim to control for time-invariant confounders, captured by the proxies for hospital quality (i.e. past mortality or re-admissions either at baseline or over the preceding year). Moreover, the approach taken, of controlling for baseline and a moving window of mortality and re-admissions, reduces the concern that bias may be introduced if sicker patients tend to attend hospitals with a higher (or a lower) TTO, as we anticipate that this is captured in the measured covariates, including hospital quality. Although the assumption that the IV is not correlated with outcomes or unobserved confounders, conditional on the variables controlled for, is fundamentally untestable, but some reassurance is provided if observed covariates do not vary across levels of the instrument. *Figure 29* illustrates that observed confounders are similar across levels of the TTO, which provides support for the underlying assumptions.

Person-level instrumental variable approach

The IV approach begins by estimating treatment effects for 'marginal' patients, that is patients for whom there was equipoise about the ES decision according to these measured characteristics (e.g. age, TTO), as well as characteristics unmeasured in the data (e.g. physiology). For these (hypothetical) marginal patients, the IV approach estimates treatment effects for patients for whom a small change (or nudge) in the TTO (the instrument) can 'tip the balance' towards ES, but does not change the level of any risk factors, including factors that are not unmeasured. Comparing outcomes for patients defined according to small differences in the TTO, therefore, provides an estimate of the causal effect of ES compared with NES for similar patients. By repeating this contrast across different levels of TTO, the study can estimate treatment effects for sets of marginal patients with different combinations of confounders (e.g. frailty levels).

Each individual in the data set, given their observed and unobserved confounders, would be a marginal patient at some level of the TTO. In the absence of further information, we might simply calculate individual-level treatment effects as the average across all marginal patients with similar characteristics

TABLE 22 Instrumental variable strength for the hospital-level TTO within the HES data (2009–19) for emergency admissions that met the ESORT study inclusion criteria for each of the five conditions

Condition	Montiel–Pflueger robust weak instrument test: F-statistic
Appendicitis	6207
Cholelithiasis	24,517
Diverticular disease	450
Hernia	1664
Intestinal obstruction	526
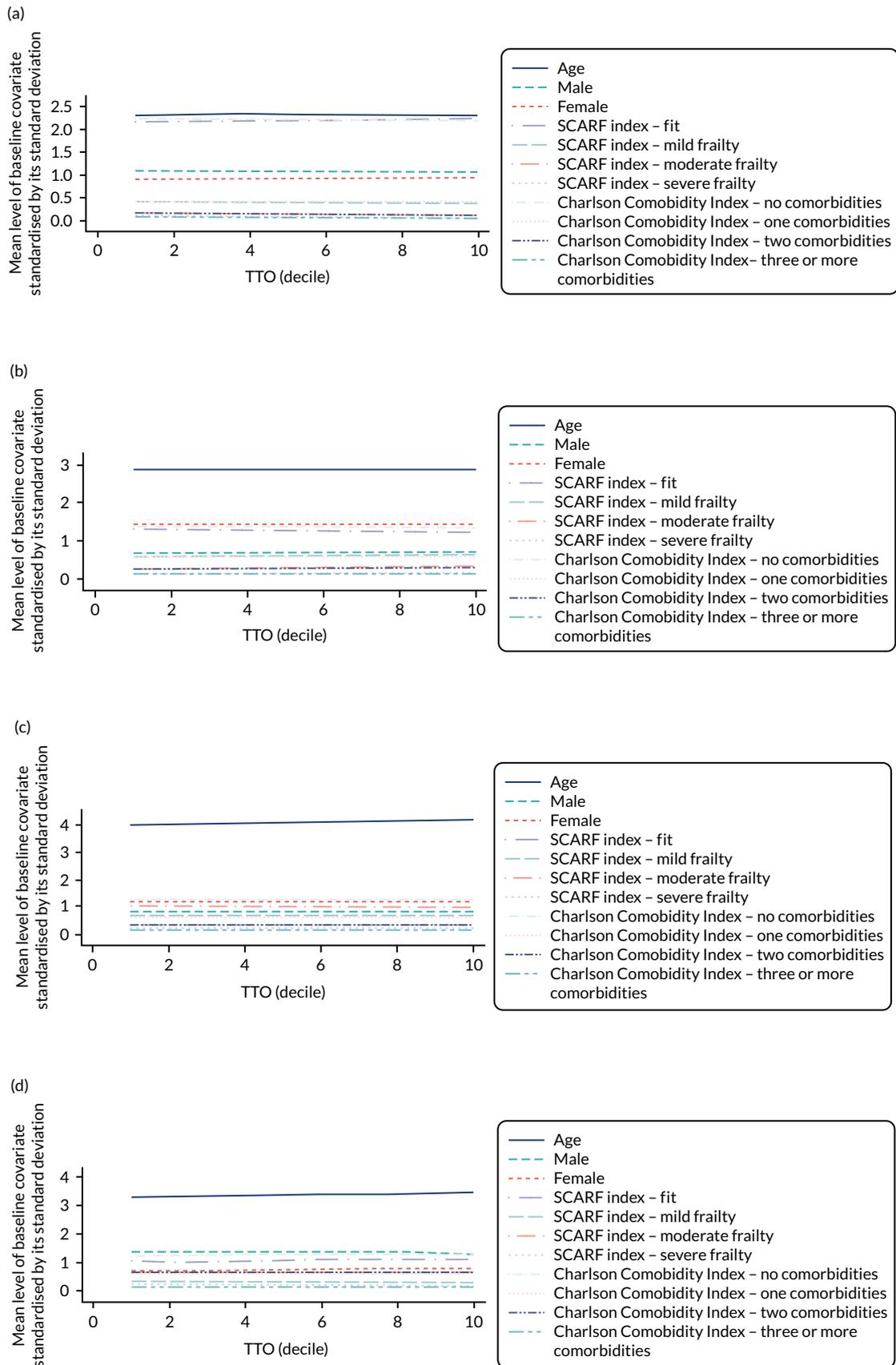
(a)

(b)

(c)

(d)

FIGURE 29 Mean level of rescaled baseline covariates according to the level of the IV. (a) Acute appendicitis; (b) cholelithiasis; (c) diverticular disease; (d) abdominal wall hernia; and (e) intestinal obstruction. (*continued*)
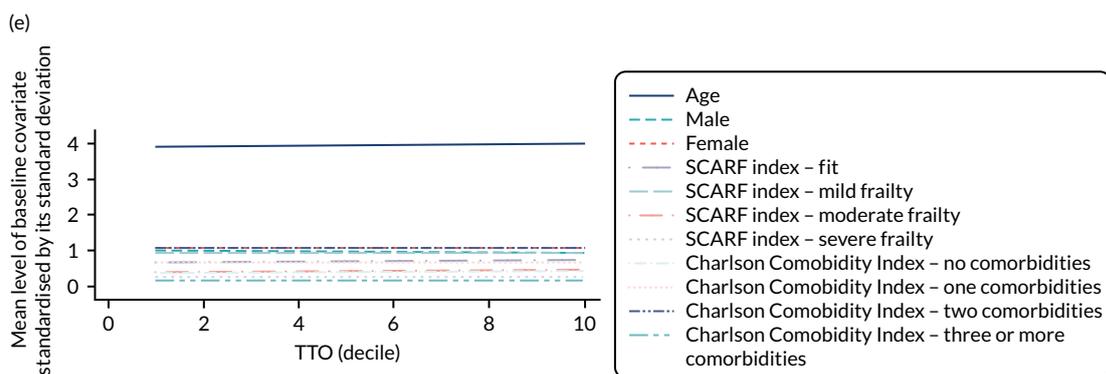
FIGURE 29 Mean level of rescaled baseline covariates according to the level of the IV. (a) Acute appendicitis; (b) cholelithiasis; (c) diverticular disease; (d) abdominal wall hernia; and (e) intestinal obstruction.

to the individual to obtain an effect estimate for each individual. However, the observed decision on whether or not an individual patient has ES provides some information on the extent to which the levels of unobserved confounders for that individual either encouraged or discouraged ES. For instance, if an individual who, according to the observed baseline measures, is at low risk and presents to a hospital with a low TTO does actually receive ES, then we can infer that their unobserved characteristics were such that they influenced the decision to have ES. Therefore, for each individual, a treatment effect is obtained by averaging the treatment effects for those marginal patients who share the same observed characteristics, and who has estimated levels of unobserved confounders that are consistent with the observed ES decision, for that individual, given the TTO in the hospital that they were admitted to, at the time they were first seen by the surgical team (i.e. time zero). The estimated individual-level treatment effects can then be averaged over any sample characteristics to report the effectiveness of ES at the subgroup level, or for the full sample, to obtain an average treatment effect estimate.

This LIV approach was implemented as follows. First, each patient's propensity for ES was estimated according to their observed characteristics and the TTO. Second, an outcome model was estimated relating the observed outcome to the individuals' observed characteristics and their propensity for ES, along with interactions between them. After estimating the model, the marginal treatment effect was obtained by considering the impact of a marginal change in the propensity for ES on outcomes. Third, numerical integration was used to obtain individual-level treatment effects. The steps were bootstrapped 300 times (200 times for sensitivity analyses because of computational complexity) to obtain standard errors and CIs. (For further details on the estimation steps, see Basu.[55])

Proxies for the quality of acute care

The validity of an IV analysis assumes that, conditional on the observed baseline covariates, the IV does not have a direct effect on the outcome, except through influencing the receipt of treatment. The requisite assumptions could be violated if the quality of the acute care, which is an unmeasured variable, was associated with the hospital's TTO. Information was, therefore, collated to proxy the quality of acute care for emergency admissions with each acute condition, by extracting from the HES data rates of all-cause mortality and emergency re-admissions up to 90 days for each hospital. The proxy measures for quality were chosen to adjust for both time-constant differences in quality across hospitals, and those that differed over time. This information was reported for each condition for the 2009–10 financial year to provide a baseline, time-invariant proxies for care quality in each hospital. The information was also reported for the one year preceding each qualifying emergency hospital admission, to provide time-varying proxies for care quality.

We chose to include these proxies for care quality rather than hospital fixed effects to provide a more specific proxy for care quality and to avoid removing all between-hospital variation in rates of ES, which would lead to inflated standard errors in the estimated effectiveness of ES.

As part of the sensitivity analyses, we considered external measures of hospital quality using data from NELA.[59-61] As data were not available for all years of the study, and definitions changed over time, we constructed an average (weighted by volume) using data from 2016–18 for the following seven indicators of quality of perioperative management for emergency laparotomy patients, which we anticipate would capture the influence of any potential time-invariant observed confounders associated with hospital quality:

1. adjusted mortality rate
2. proportion of patients in whom a risk assessment was documented preoperatively
3. proportion of patients arriving in theatre within a time appropriate for the urgency of surgery
4. proportion of patients with a calculated preoperative risk of death > 5% for whom a consultant surgeon and anaesthetist were present in theatre
5. admission to critical care when risk of death is $\geq$ 5%
6. unplanned returns to theatre
7. unplanned returns to critical care.

Sensitivity analyses

We assessed the extent to which the findings from the main analyses were sensitive to alternative definitions and assumptions. We consider alternative definitions of ES by considering a more conservative panel definition of ES (i.e. SA1), and a reduced time window for the procedure to qualify as ES by considering only procedures that occurred before the 75th percentile of the time of ES used in the main analysis (i.e. SA2). We explored the sensitivity of the findings to alternative proxies for hospital quality (i.e. SA3) by using an external proxy for hospital quality (i.e. the NELA quality measures). In addition, we considered the impact of removing hospitals with low volume for the procedures of interest (i.e. SA4). We excluded hospitals whose volume of eligible procedures was less that thresholds of one IQR below the median, which, in this SA, led to the exclusion of 22 , 23, 27, 25 and 16 hospitals for the appendicitis, cholelithiasis, diverticular disease, hernia and intestinal obstruction cohorts, respectively, and to between 2.84% (Intestinal obstruction) and 6.25% (diverticular disease) of observations. We considered an alternative definition of the primary outcome of DAOH (i.e. SA5), which weights the number of DAOH prior to death for patients who died before day 90.[62] Finally, we used regression adjustment to report estimates of relative effectiveness under the assumption of no unobserved confounding (i.e. SA6).

Appendix 3 Additional notes on resource use and cost calculation

TABLE 23 Resource use categories for operative procedures in ES window

Resource use category	Source	Appendicitis ES	Appendicitis NES strategies	Cholelithiasis ES	Cholelithiasis NES strategies	Diverticular disease ES	Diverticular disease NES strategies	Hernia ES	Hernia NES strategies	Intestinal obstruction ES	Intestinal obstruction NES strategies
Most common operative procedures in the ES window in each arm		Emergency excision of abnormal appendix	Interval appendicectomy	Total cholecystectomy	Endoscopic sphincterotomy of sphincter of Oddi and removal of calculus	Rectosigmoidectomy and closure of rectal stump and exteriorisation of bowel	Fibreoptic endoscopic snare resection of lesion of colon	Primary repair of inguinal hernia using insert of prosthetic material	Unspecified urethral catheterisation of bladder	Freeing of adhesions of peritoneum	Other specified other therapeutic endoscopic operations on lower bowel using fibreoptic sigmoidoscope
Duration of surgery (minutes) (source)	Literature	70 (Javanmard-Emamghissi et al.[21])	70 (Javanmard-Emamghissi et al.[21])	90 (American College of Surgeons[124])	90 (Silva et al.[125])	135[a] (Heah et al.[126])	25 (Teramoto et al.[127])	60 (Wu et al.[128])	15 (Wilson[129])	50 (Sallinen et al.[130])	50 (Sallinen et al.[130])
Staffing levels[b]	Expert opinion	S1	S1	S1	S3	S1	S1	S1	S2	S1	S3
Instruments	Expert opinion	Laparoscopic set	Laparoscopic set	Laparoscopic set		Major general set	Polypectomy set	Minor general set		Major general set	
Equipment	Expert opinion	Laparoscope, cable and tray	Laparoscope, cable and tray	Laparoscope, cable and tray	Laparoscope, cable and tray		Laparoscope, cable and tray				Laparoscope, cable and tray
Main disposables	Expert opinion	Loops for stump closure	Loops for stump closure			Linear stapler and reload		Biosynthetic mesh	Foyle catheterisation kit		

a If the procedure appeared with operative codes for loop colostomy, other specified other exteriorisation of colon, or unspecified other exteriorisation of colon, then the duration was assumed to be 205 minutes.

b S1 considered as one consultant surgeon, one registrar surgeon, two band 5 nurses, one band 6 nurse, one operating department practitioner, one consultant anaesthetist and one registrar anaesthetist. S2 considered as one band 5 nurse. S3 considered as one consultant anaesthetist, one endoscopist and two band 5 nurses.

Note
Table includes exemplar data for the most common operative procedures in ES window.

TABLE 24 Unit costs (£ 2019/20) for potential cost drivers

Item	Unit	Unit cost (£)	Source, definitions and assumptions
Inpatient stay			
General ward	Day	347.00	NHS reference costs 2017/18:[81] weighted average of codes FD05A (Abdominal Pain with Interventions) and FD05B (Abdominal Pain without Interventions) in non-elective excess bed-days sheet
ICU ward			
Level 2 ICU	Day	1190.00	NHS reference costs 2017/18:[81] XC06Z (Adult Critical Care, 1 organ supported)
Level 3 ICU	Day	1890.00	NHS reference costs 2017/18:[81] Weighted average of codes XC01Z (Adult Critical Care, 6 or more Organs Supported) to XC05Z (Adult Critical Care, 2 Organs Supported)
Diagnostic procedures			
More common diagnostic procedures for acute appendicitis			
Computed tomography	Procedure	83.00	NHS reference costs 2017/18:[81] RD20A [Computerised Tomography Scan of One Area, without Contrast, 19 years and over (IMAG)]
Unspecified diagnostic endoscopic examination of colon	Procedure	206.00	NHS reference costs 2017/18:[81] FE31Z [Diagnostic Colonoscopy with Biopsy, 19 years and over (NES)]. Mean bed-day costs of general ward subtracted to avoid double-counting
Fibreoptic endoscopic examination of upper gastrointestinal tract and biopsy of lesion of upper gastrointestinal tract	Procedure	197.00	NHS reference costs 2017/18:[81] FE21Z [Diagnostic Endoscopic Upper Gastrointestinal Tract Procedures with Biopsy, 19 years and over (NES)]. Mean bed-day costs of general ward subtracted to avoid double-counting
Diagnostic fibreoptic endoscopic examination of colon and biopsy of lesion of colon	Procedure	277.00	NHS reference costs 2017/18:[81] FE31Z [Diagnostic Colonoscopy with Biopsy, 19 years and over (NES)]. Mean bed-day costs of general ward subtracted to avoid double-counting
Computed tomography of head	Procedure	83.00	NHS reference costs 2017/18:[81] RD20 [Computerised Tomography Scan of One Area, without Contrast, 19 years and over (IMAG)]
More common diagnostic procedures for diverticular disease			
Computed tomography	Procedure	83.00	NHS reference costs 2017/18:[81] RD20 [Computerised Tomography Scan of One Area, without Contrast, 19 years and over (IMAG)]
Unspecified diagnostic endoscopic examination of lower bowel using fibreoptic sigmoidoscope	Procedure	143.00	NHS reference costs 2017/18:[81] [FE35Z: Diagnostic Flexible Sigmoidoscopy, 19 years and over (NES)]. Mean bed-day costs of general ward subtracted to avoid double-counting
Unspecified diagnostic endoscopic examination of colon	Procedure	206.00	NHS reference costs 2017/18:[81] FE32Z [Diagnostic Colonoscopy, 19 years and over (NES)]. Mean bed-day costs of general ward subtracted to avoid double-counting

TABLE 24 Unit costs (£ 2019/20) for potential cost drivers (*continued*)

Item	Unit	Unit cost (£)	Source, definitions and assumptions
Unspecified diagnostic fibreoptic endoscopic examination of upper gastrointestinal tract	Procedure	277.00	NHS reference costs 2017/18:[81] FE31Z [Diagnostic Colonoscopy with Biopsy, 19 years and over (NES)]. Mean bed-day costs of general ward subtracted to avoid double-counting
Diagnostic endoscopic examination of lower bowel and biopsy of lesion of lower bowel using fibreoptic sigmoidoscope	Procedure	205.00	NHS reference costs 2017/18:[81] FE34Z [Diagnostic Flexible Sigmoidoscopy with Biopsy, 19 years and over (NES)]. Mean bed-day costs of general ward (see above) subtracted to avoid double-counting
More common diagnostic procedures for a hernia			
Computed tomography	Procedure	83.00	NHS reference costs 2017/18:[81] RD20A [Computerised Tomography Scan of One Area, without Contrast, 19 years and over (IMAG)]
Transthoracic echocardiography	Procedure	101.00	NHS reference costs 2017/18:[81] RD51C [Simple Echocardiogram, 5 years and under (IMAG)]
Computed tomography of abdomen	Procedure	83.00	NHS reference costs 2017/18:[81] RD20 [Computerised Tomography Scan of One Area, without Contrast, 19 years and over (IMAG)]
Computed tomography of head	Procedure	83.00	NHS reference costs 2017/18:[81] RD20A [Computerised Tomography Scan of One Area, without Contrast, 19 years and over (IMAG)]
Diagnostic endoscopic examination of peritoneum	Procedure	404.00	NHS reference costs 2017/18:[81] FE31Z [Diagnostic Colonoscopy with Biopsy, 19 years and over (NES)]. Mean bed-day costs of general ward subtracted to avoid double-counting
Operative procedures			
Staff input			
Consultant surgeon	Minute	1.80	PSSRU 2019:[131] section 14. Cost per working hour: consultant – surgical
Anaesthesiologist	Minute	1.80	PSSRU 2019:[131] section 14. Cost per working hour: consultant – medical
Consultant radiologist	Minute	1.80	PSSRU 2019:[131] section 14. Cost per working hour: consultant – medical
Registrar: surgery	Minute	0.80	PSSRU 2019:[131] section 14. Cost per working hour: registrar
Registrar: anaesthesiology	Minute	0.80	PSSRU 2019:[131] section 14. Cost per working hour: registrar
Registrar: radiology	Minute	0.80	PSSRU 2019:[131] section 14. Cost per working hour: registrar
Nurse: band 5	Minute	0.60	PSSRU 2019:[131] section 13. Cost per working hour: band 5 – hospital-based nurse
Nurse: band 6	Minute	0.80	PSSRU 2019:[131] section 13. Cost per working hour: band 6 – hospital-based nurse
Operating department practitioner	Minute	0.80	PSSRU 2019:[131] section 13. Assumed same cost as cost per working hour of band 6 hospital-based nurse

TABLE 24 Unit costs (£ 2019/20) for potential cost drivers (*continued*)

Item	Unit	Unit cost (£)	Source, definitions and assumptions
Overhead costs			
Operating room	Minute	5.40	Includes direct drug and CSSD costs, as well allocated costs [other staff; property and equipment maintenance; domestics and cleaning; heat, light and power; rent and rates; purchases of furniture, fittings and equipment (non-capital charge) and others]. Weighted average of 43 hospitals in Scotland[79]
Reusable instruments and equipment			
Laparoscopic colorectal set	Procedure	39.20	Set lists and indicative unit costs informed by the procurement office of the Royal Devon and Exeter NHS Foundation Trust. See *Appendix 3, Table 25*, for the full list of components. Total purchase cost is £3112. Number of uses is 2750. Final cost includes sterilisation cost following at £0.80 per instrument used[80]
Main laparoscopic set	Procedure	36.80	Set lists and indicative unit costs informed by the procurement office of the Royal Devon and Exeter NHS Foundation Trust. See *Appendix 3, Table 25*, for the full list of components. Total purchase cost is £2511. Assumed number of uses is 2750. Final cost includes sterilisation cost following at £0.80 per instrument used[80]
Major general set	Procedure	39.20	Set lists and indicative unit costs informed by the procurement office of the Royal Devon and Exeter NHS Foundation Trust. See *Appendix 3, Table 25*, for the full list of components. Total purchase cost is £2744. Assumed number of uses is 2750. Final cost includes sterilisation cost following at £0.80 per instrument used[80]
Minor general set	Procedure	32.80	Set lists and indicative unit costs informed by the procurement office of the Royal Devon and Exeter NHS Foundation Trust. See *Appendix 3, Table 25*, for the full list of components. Total purchase cost is £1417. Assumed number of uses is 2750. Final cost includes sterilisation cost following at £0.80 per instrument used[80]
Endoscopic polypectomy set	Procedure	16.20	Set lists and indicative unit costs informed by the procurement office of the Royal Devon and Exeter NHS Foundation Trust. Includes endoscopic forceps, snare and endoscopic clips. Final cost includes sterilisation cost following at £0.80 per instrument used.[80] Unit cost calculated assuming number of uses is 4400 (except for snare and clips, which are assumed to be disposable)

continued

TABLE 24 Unit costs (£ 2019/20) for potential cost drivers *(continued)*

Item	Unit	Unit cost (£)	Source, definitions and assumptions
Telescope and stack	Procedure	15.20	Set lists and indicative unit costs informed by the procurement office of the Royal Devon and Exeter NHS Foundation Trust. Includes stack, scope (Precision ideal eyes 10 mm 30°, HD autoclavable Laparoscope 33 cm), tray and cable [fibreoptic cable 5.0 mm × 10 ft. (3.05 m)]. Purchase cost of stack and scope, tray and table are £68,760 and £2334, respectively. Unit cost calculated assuming expected number of uses is 4400
Ultrasound system	Procedure	1.50	Set lists and indicative unit costs informed by the procurement office of the Royal Devon and Exeter NHS Foundation Trust. Purchase cost of ultrasound system is £7132. Unit cost calculated assuming expected number of uses is 4400
Disposables			
Laparoscopic linear stapler	Procedure	262.00	Set lists and indicative unit costs informed by the procurement office of the Royal Devon and Exeter NHS Foundation Trust. Linear cutter 75 mm. One is assumed to be used per procedure
Stapler reload	Procedure	36.50	Set lists and indicative unit costs informed by the procurement office of the Royal Devon and Exeter NHS Foundation Trust. Reload linear cutter, blue, 75 mm. Purchase cost of £465.61 per box of 12. One is assumed to be used per procedure
Endoloop ligature	Procedure	56.90	Set lists and indicative unit costs informed by the procurement office of the Royal Devon and Exeter NHS Foundation Trust. Endoloop Ethicon. Three are assumed to be used per procedure[132]
Biosynthetic mesh	Procedure	61.10	Set lists and indicative unit costs informed by the procurement office of the Royal Devon and Exeter NHS Foundation Trust. Sutumed Polipropilene Non-absorbable Hernia Mesh 12 × 12 inches. One is assumed to be used per procedure
Abdominal drain set	Procedure	18.20	Set lists and indicative unit costs informed by the procurement office of the Royal Devon and Exeter NHS Foundation Trust. Set includes 1000 ml drainage bag, catheter valve cap, slide clamp, tape strips and wipe. Purchase cost £36.50 per box of two
Foyle catheterisation kit	Procedure	9.40	Set lists and indicative unit costs informed by the procurement office of the Royal Devon and Exeter NHS Foundation Trust. Catheterisation Set 16fr Foley and extras. Includes a 16fr Foley catheter, a 500-ml leg-bag, a 2000-ml bedside drainage bag, sterile syringe and lube

CSSD, Central Sterile Services Department; ICU, intensive care unit; PSSRU, Personal Social Services Research Unit.

TABLE 25 Full list of components of surgical sets considered in cost analysis

Cost analysis	Surgical set			
	Laparoscopic colorectal set	Main laparoscopic set	Major general set	Minor general set
Component	Aesculap dorsey forceps four parts	Anti-tamper tags	BP handle no. 4	Artery forceps mosquito curved
	Anti-tamper tags	BP handle no. 3	BP handle no. 5	BP handle no. 3
	BP handle no. 3	BP handle no. 4	Babcock tissue forceps 6 1/2 inch	BP handle no. 4
	BP handle no. 4	Bottom tray	Babcock tissue forceps 8 inch	Babcock tissue forceps
	Babcock tissue forceps long	Container	Balfour self-retaining retractor	Catspaw retractor
	Babcock tissue forceps short	Container identification label	Deaver retractor, broad	Diathermy dissecting forceps mcindoe
	Bottom tray	De-jardin stone forceps	Deaver retractor, narrow	Diathermy quiver
	Container	Diathermy dissecting forceps mcindoe	Diathermy dissecting forceps mcindoe	Disposable blue tray wrap 120 × 150
	Container identification label	Diathermy quiver long plus black end cap	Diathermy quiver	Disposable green tray wrap 120 × 150
	Diathermy dissecting forceps mcindoe	Dissecting forceps debakey 6 inch	Disposable green tray wrap 120 × 150	Dissecting forceps debakey 6 inch
	Diathermy quiver	Dissecting forceps gillies toothed	Dissecting forceps debakey 6 inch	Dissecting forceps gillies toothed
	Diathermy quiver long plus black end cap	Dunhill artery forceps	Dissecting forceps debakey 8 inch	Dunhill artery forceps
	Dissecting forceps debakey 6 inch	Eragon ratchet handle – do not assemble to forceps	Dissecting forceps debakey 9 1/2 inch	Heiss artery forceps
	Dissecting forceps debakey 8 inch	Filter and retaining clip	Dissecting forceps gillies toothed	Lahey artery forceps
	Dissecting forceps gillies toothed	Grasping forceps plus ratchet with connector (pm 109)	Dissecting forceps non toothed 5 inch	Lanes dissecting forceps (1/2 teeth)
	Doyen intestinal clamp curved	Hassan 10 mm (two parts plus 10-mm clear seal)	Doyen intestinal clamp curved	Littlewoods tissue forceps
	Dunhill artery forceps	Insulated hook with connector	Doyen intestinal clamp straight	Mayo pin holding next two items
	Dyball retractor	Lanes dissecting forceps (1/2 teeth)	Dunhill artery forceps	Mayo pin holding next three items
	Filter and retaining clip	Laparoscopic diathermy lead (8 mm bovie)	Dyball retractor	Mayo pin holding next four items
	Grasping forceps plus ratchet with connector (pm 109)	Littlewoods tissue forceps	Heiss artery forceps	Meyarding finger retractor
	Hasson 12 mm (3 parts) ea12nh send disassembled	Maryland forceps no ratchet with connector (pm 102)	Lahey artery forceps	Monopolar diathermy lead pin fitting
	Heiss artery forceps		Lanes dissecting forceps (1/2 teeth)	Needle holder crilewood
			Lang stevenson intestinal clamps	Needle holder mayo hegar

TABLE 25 Full list of components of surgical sets considered in cost analysis (*continued*)

Cost analysis	Surgical set			
	Laparoscopic colorectal set	**Main laparoscopic set**	**Major general set**	**Minor general set**
	Insulated hook with connector	Mesh basket with lid	Littlewoods tissue forceps	Poirers/allis tissue forceps
	Ireusable cannula 12 mm	Modular monopolar forceps (johan) sn 8393.184 2 parts	Massons needle holder	Retractor langenbeck medium
	Lanes dissecting forceps (1/2 teeth)	Monopolar diathermy lead pin fitting	Mayo pin holding next one item	Retractor langenbeck small
	Laparoscopic diathermy lead (8 mm bovie)	Myoma forceps plus ratchet with connector (pm 117)	Mayo pin holding next three items	Retractor morris medium
	Littlewoods tissue forceps	Needle holder crilewood	Mayo pin holding next four items	Retractor self-retaining travers
	Maryland f/cep no ratchet with connector (pm 102)	Needle holder mayo hegar	Mayo pin holding next six items	Retractor self-retaining west
	Massons needle holder	Pike mouth forceps plus ratchet with connector (pm 107)	Monopolar diathermy lead pin fitting	Scissors kilner curved
	Monopolar diathermy lead pin fitting	Retractor langenbeck medium	Moynihan cholecystectomy clamp	Scissors mayo curved
	Needle holder mayo hegar	Retractor langenbeck small	Needle holder mayo hegar 7 1/4 inch	Scissors mayo straight 5 3/4 inch
	Nelson robert scissors	Reusable cannula 10 mm	Needle holder mayo hegar 8 1/2 inch	Scissors mcindoe curved
	Parker kerr intestinal clamp straight	Reusable cannula 12 mm	Nelson robert scissors	SH/SH scissors
	Retractor langenbeck medium	Scissors mayo straight	Parker kerr intestinal clamp curved	Soaker sheet to be placed under basket/tray
	Retractor langenbeck small	Scissors mcindoe curved	Parker kerr intestinal clamp straight	Spencer wells artery forceps 7 inch curved
	Retractor morris medium	SH/SH scissors	Retractor langenbeck medium	Spencer wells artery forceps 8 inch straight
	Roberts artery forceps	Spencer wells artery forceps 7 inch curved	Retractor morris large	Sponge holder rampley
	Scissors mayo straight	Sponge holder rampley	Roberts artery forceps	T.o.e. dissecting forceps
	Scissors mcindoe curved	Threaded cannula 5 mm (two parts)	Scissors mayo curved	Trayliner
	SH/SH scissors	Top tray	Scissors mayo straight 5 3/4 inch	Wash basket
	Sponge holder rampley	Towel clip small	Scissors mcindoe curved	
	Threaded cannula 5 mm (two parts)	Trayliner	SH/SH scissors	
		Trocar blunt tip 10 mm	Soaker sheet to be placed under basket/tray	
		Trocar pencil point 12 mm	Sponge holder rampley	
			Styles tissue forceps	
			Trayliner	

TABLE 25 Full list of components of surgical sets considered in cost analysis (*continued*)

| Cost analysis | Surgical set | | | |
	Laparoscopic colorectal set	**Main laparoscopic set**	**Major general set**	**Minor general set**
	Top tray	Trocar pencil point 5 mm	Wash basket	
	Trayliner	Trocar sharp tip 5 mm	Waughs diathermy dissecting forceps	
	Trocar blunt tip 12 mm	Wash basket		
	Trocar pencil point 12 mm			
	Trocar pencil point 5 mm			
	Trocar sharp tip 5 mm			
	Waughs diathermy dissecting forceps			

Appendix 4 Additional notes on calculations of costs and quality-adjusted life-years

Search for appropriate health-related quality-of-life scores and adjustment

Search strategies and criteria for selection of studies

The approach to estimating QALYs required that appropriate HRQoL values were identified from a literature review. We carried out separate search strategies for each condition in MEDLINE (*Box 1*). The criteria used to select the most appropriate source of HRQoL recognised the specific requirements of the ESORT study (see study protocol on project webpage[30]) and were prioritised according to the following:

- The study measured HRQoL following an emergency admission for patients with at least one of the diagnostic subcategories described in *Table 21*.
- The study considered at least one intervention regarded as ES66 by the clinical panel.
- The intervention was performed in the emergency (non-elective) setting.
- The study evaluated HRQoL using the tool recommended by NICE in its methodological guidance (i.e. the EuroQol-5 Dimensions, three-level version).[63]
- The study evaluated HRQoL at baseline (i.e. preoperatively) and at 1 year from baseline.
- The study was conducted in the UK or in a country with a similar health-care system and similar demographics.
- The study was conducted no earlier than 10 years before the start date of the ESORT study (i.e. 2010).

BOX 1 Search strategies for HRQoL data

Appendicitis

Database: Ovid MEDLINE(R) ALL

Date range searched: 1946 to 19 August 2021

Date searched: 20 August 2021

Search strategy

1. *appendicitis/ (16,067)
2. *appendicectomy/ (6399)
3. appendic*.ti,ab. (33,073)
4. appendec*.ti,ab. (10,129)
5. emergency+surgery*.ti,ab. (9412)
6. emergency+appendicectomy*.ti,ab. (153)
7. non-operative+manag*.mp. (1888)
8. conservative+manag*.mp. (16,648)
9. antibiotic*.ti,ab. (360,099)
10. antibiotic+adj+therapy.ti,ab. (0)
11. Anti-Bacterial+Agents/tu (135,940)
12. Watchful+wait$.tu. (0)
13. delayed+surg$.ti,ab. (2186)
14. trial.ti,ab. (657,219)
15. RCT.ti,ab. (24,987)
16. randomi#ed+controlled+trial.pt. (541,163)
17. controlled+clinical+trial.pt. (94,345)

BOX 1 Search strategies for HRQoL data (*continued*)

18. case+control+stud$.ti,ab. (113,048)
19. cross-sectional+stud$.ti,ab. (197,037)
20. cohort+stud$.ti,ab. (244,943)
21. observational+stud$.ti,ab. (126,477)
22. Economic+evaluation.ti,ab. (9983)
23. EuroQol-5+Dimension.ti,ab. (670)
24. "EQ-5D".ab. (9451)
25. or/1-2 (19,106)
26. or/3-13 (496,587)
27. and/27-28 (16,860)
28. or/14-24 (1,682,168)
29. and/29-30 (1350)
30. or/25-26 (9720)
31. and/31-32 (4)

Cholelithiasis

Database: Ovid MEDLINE(R)

Date range searched: 1946 to week 2 April 2022

Date searched: 3 April 2022

Search strategy

1. cholelithiasis.ti,ab. (7694)
2. *cholecystitis/ (9171)
3. gallston*.ti,ab. (14,648)
4. Gall* disease.ti,ab. (4221)
5. *cholecystectomy/ (8810)
6. cholecystectom$.ti,ab. (27,011)
7. ((excis* or remov* or surg* of) adj4 gallbladder).ti,ab. (2201)
8. ((early or emergency) adj (drain* or surg* or cholec*)).ti,ab. (20,616)
9. trial.ti,ab. (606,312)
10. RCT.ti,ab. (22,408)
11. randomi#ed+controlled+trial.pt. (563,728)
12. controlled+clinical+trial.pt. (94,782)
13. case+control+stud$.ti,ab. (103,937)
14. cross-sectional+stud$.ti,ab. (171,512)
15. cohort+stud$.ti,ab. (230,818)
16. retrospective stud$.ti,ab. (162,780)
17. retrospective analy$.ti,ab. (81,357)
18. observational+stud$.ti,ab. (114,701)
19. (cost adj (utility or effectiv*)).ti,ab. (129,042)
20. Economic+evaluation.ti,ab. (9060)
21. EuroQol.mp. (5920)
22. EQ-5D*.af. (8893)
23. or/1-8 (71,556)
24. or/9-20 (1,905,866)
25. and/23-24 (9231)
26. or/21-22 (11,568)
27. and/25-26 (27)

BOX 1 Search strategies for HRQoL data (*continued*)

Diverticular disease

Database: Ovid MEDLINE(R) ALL

Date range searched: 1946 to 24 September 2021

Date searched: 26 September 2021

Search strategy

1. *diverticulitis/ (2667)
2. *Diverticulum/ (8202)
3. Diverticul*.mp. (33,728)
4. emergency+surgery*.ti,ab. (9470)
5. Drainage*.ti,ab. (97,292)
6. Lavage*.ti,ab. (52,937)
7. Percutaneous+drainage*.ti,ab. (4232)
8. sigmoidectomy*.ti,ab. (1089)
9. colectomy*.mp. (24,998)
10. conservative+manag*.mp. (16,769)
11. antibiotic*.ti,ab. (362,483)
12. antibiotic+adj+therapy.ti,ab. (0)
13. Anti-Bacterial+Agents/tu (136,787)
14. Watchful+wait$.tu. (0)
15. delayed+surg$.ti,ab. (2207)
16. trial.ti,ab. (662,690)
17. RCT.ti,ab. (25,317)
18. randomi#ed+controlled+trial.pt. (544,498)
19. controlled+clinical+trial.pt. (94,426)
20. case+control+stud$.ti,ab. (113,845)
21. cross-sectional+stud$.ti,ab. (200,110)
22. cohort+stud$.ti,ab. (248,537)
23. observational+stud$.ti,ab. (128,258)
24. Economic+evaluation.ti,ab. (10,055)
25. EuroQol-5+Dimension.ti,ab. (690)
26. "EQ-5D".ti,ab. (9680)
27. or/1-3 (10,504)
28. or/4-15 (651,442)
29. and/27-28 (10,504)
30. or/16-24 (1,697,513)
31. and/29-30 (200)
32. or/25-26 (9957)
33. and/31-32 (2)

Hernia

Database: Ovid MEDLINE(R) ALL

Date range searched: 1946 to 24 September 2021

Date searched: 26 September 2021

Search strategy

1. (inguinal or femoral or ventral or umbilical or abdominal wall).ti,ab. (335,807)
2. hernia.ti,ab. (52,281)

BOX 1 Search strategies for HRQoL data (*continued*)

3. hernioplasty/ (9400)
4. herniorrhaphy/ (9400)
5. hernioplasty.ti,ab. (1602)
6. herniorrhaphy.ti,ab. (2372)
7. repair+or+surg*.ti,ab. (22)
8. hernia+adj+repair.ti,ab. (0)
9. (early adj3 (surg* or repair)).ti,ab. (28,362)
10. trial.ti,ab. (657,219)
11. RCT.ti,ab. (24,987)
12. randomi#ed+controlled+trial.pt. (541,163)
13. controlled+clinical+trial.pt. (94,345)
14. case+control+stud$.ti,ab. (113,048)
15. cross-sectional+stud$.ti,ab. (197,037)
16. cohort+stud$.ti,ab. (244,943)
17. retrospective+stud$.ti,ab. (179,855)
18. observational+stud$.ti,ab. (126,477)
19. (cost adj (utility or effectiv*)).ti,ab. (149,693)
20. Economic+evaluation.ti,ab. (9983)
21. (quality of life or QoL or HRQoL).ti,ab. (312,433)
22. EuroQol.af. (6571)
23. EQ-5D*.af. (9689)
24. and/1-2 (20,870)
25. or/3-9 (40,553)
26. or/10-20 (1,972,540)
27. or/21-23 (315,435)
28. and/24-27 (129)

Intestinal obstruction

Database: Ovid MEDLINE(R) ALL

Date range searched: 1946 to 24 September 2021

Date searched: 26 September 2021

Search strategy

1. *Intestinal+Obstruction/di (1521)
2. ((intestin* or bowel) adj4 obstruction).ti,ab. (24,747)
3. ((intestin* or bowel) adj4 blockage).ti,ab. (80)
4. ((intestin* or bowel) adj4 adhesion).ti,ab. (1139)
5. ((early or emergency) adj (drain* or surg* or relief)).ti,ab. (23,075)
6. intussusception.ti,ab. (8815)
7. volvulus.ti,ab. (8826)
8. gallstone+ileus.ti,ab. (969)
9. excision.ti,ab. (123,989)
10. anastomosis.ti,ab. (60,672)
11. ileostomy.ti,ab. (6729)
12. bypass.ti,ab. (138,349)
13. hemicolectomy.ti,ab. (4336)
14. stenting.ti,ab. (33,397)
15. colectomy.ti,ab. (12,406)

BOX 1 Search strategies for HRQoL data (*continued*)

16. resection.ti,ab. (308,640)
17. non-operative+manag*.mp. (1935)
18. conservative+manag*.mp. (16,821)
19. antibiotic*.ti,ab. (364,019)
20. Watchful+wait$.tu. (0)
21. delayed+surgery.ti,ab. (1401)
22. trial.ti,ab. (665,581)
23. RCT.ti,ab. (25,491)
24. randomi#ed+controlled+trial.pt. (546,615)
25. controlled+clinical+trial.pt. (94,462)
26. case+control+stud$.ti,ab. (114,292)
27. cross-sectional+stud$.ti,ab. (201,749)
28. cohort+stud$.ti,ab. (250,595)
29. ((cohort or retros* or observat*) adj4 stud*).ti,ab. (738,276)
30. (cost adj (utility or effectiv*)).ti,ab. (151,653)
31. Economic+evaluation.ti,ab. (10,101)
32. EuroQol.mp. (6624)
33. EQ-5D*.af. (9916)
34. or/1-4 (26,616)
35. or/5-21 (1,033,513)
36. and/34-35 (8489)
37. or/22-31 (2,155,159)
38. and/36-37 (855)
39. or/32-33 (12,853)
40. and/38-39 (4)

Findings of the literature review

The literature review (up to September 2021) identified five studies for appendicitis,[7,8,67,133,134] one study for diverticular disease,[94] 24 studies for cholelithiasis,[83,135–158] 124 studies for hernia,[84,108,159–281] and four studies for intestinal obstruction.[85,282–284] In addition to these studies, one additional study that met the inclusion criteria for diverticular disease was identified by the team on a separate search[14] and, therefore, the total number of studies reviewed for diverticular disease was two.[14,94]

The main reasons why studies above failed to meet the inclusion criteria were that (1) the study had a different definition of study population compared with the ESORT study,[135] (2) the study had an insufficient duration of follow-up period,[7] (3) the study interventions considered were not performed in the emergency setting[247] and (4) the instrument used to measure HRQoL was not the EQ-5D index.[160]

Of the studies identified, only two studies[8,67] met the criteria for appendicitis, one study met the criteria for diverticular disease,[14] three studies met the criteria for cholelithiasis,[83,136,147] one study met the criteria for hernia[84] and one study met the criteria for intestinal obstruction.[85] For appendicitis, the study by O'Leary *et al.*[8] was chosen over the study by Sippola *et al.*[67] because the latter study reported outcomes for a median follow-up of 7 years, whereas O'Leary *et al.*[8] reported HRQoL at baseline and at 1 year. None of the studies identified for cholelithiasis and abdominal wall hernia was restricted to patients admitted into hospital as an emergency. The studies by Sutherland *et al.*[83] for cholelithiasis and Rutegård *et al.*[84] for abdominal wall hernia were favoured over other studies (e.g. Karimuddin *et al.*[136] and Rosenmüller *et al.*[147]) because they restricted the population to symptomatic patients. Although it

was not possible to find studies conducted in the UK, the studies that met the selection criteria were from similar health-care systems and from populations with similar demographics and cultural norms to the UK. The study by O'Leary *et al.*[8] was undertaken in Ireland, the study by Thornell *et al.*[14] in Sweden and Denmark, the study by Sutherland *et al.*[83] in Canada, the study by Rutegård *et al.*[84] in Sweden and the study by Young and Zahid[85] in Australia.

Four[14,83–85] out of the five studies selected compared different forms of ES, rather than ES compared with NES strategies. For this reason, as a unified approach across all five conditions, we applied the same HRQoL scores at baseline and at 1 year to both comparison groups (see *Chapter 4, Outcomes*), in keeping with the assumption noted above, that is any differences in HRQoL between the comparison groups would be captured by differences in 1-year mortality and in the rate and duration of emergency re-admissions (*Figure 30*).
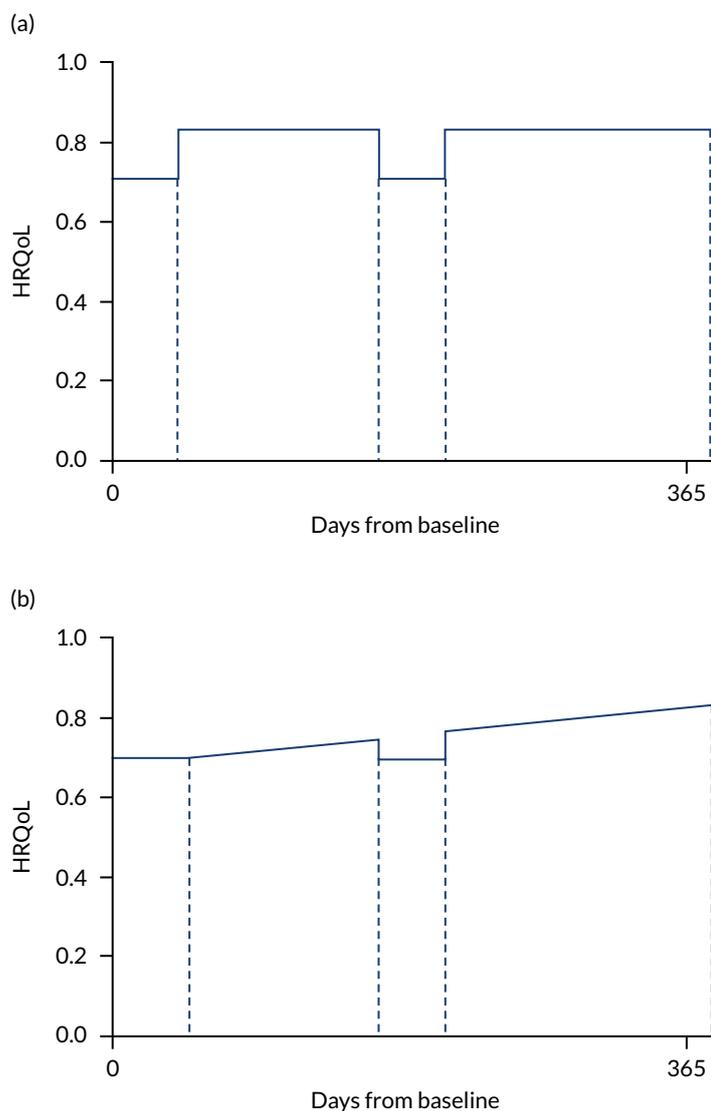
FIGURE 30 Health-related quality of life trajectory following initial (index) emergency admission and emergency re-admission for the base case, which assumes that HRQoL reaches follow-up levels following hospital discharge (panel A) and linear interpolation (B) (SA4). (a) Immediate interpolation. Baseline HRQoL is assumed to apply constantly for the duration of the index admission and any emergency re-admission. Following the index admission, the HRQoL is assumed to apply constantly for the duration of the period before the final (1-year) end point, which is accrued immediately after discharge. (b) Linear interpolation. Baseline HRQoL is assumed to apply constantly for the duration of the index admission and any emergency re-admission. HRQoL between the end points is assumed to increase linearly.

Appendix 5 Generalised linear models for life-years, quality-adjusted life-years and costs, assessment of model fit according to root mean squared error

TABLE 26 Generalised linear models for life-years, QALYs and costs, assessment of model fit according to RMSE

			Condition				
Family	**Link**	**Degree**[a]	**Appendicitis**	**Cholelithiasis**	**Diverticular Disease**	**Hernia**	**Intestinal obstruction**
Life-years							
Binomial	Logit	1	[0.050]	[0.116]	[0.181]	[0.170]	[0.288]
Binomial	Logit	2	0.050	0.116	0.181	0.170	0.288
Binomial	Logit	3	0.050	0.116	0.181	0.170	0.288
QALYs							
Binomial	Logit	1	[0.059]	[0.137]	[0.192]	[0.204]	[0.074]
Binomial	Logit	2	0.059	0.137	0.192	0.204	0.074
Binomial	Logit	3	0.059	0.137	0.192	0.204	0.074
Costs							
Gaussian	Identity	1	3530.330	6556.831	8980.985	[8975.387]	15,450.896
Inverse Gaussian	Identity	1	3533.875	6560.378	8993.261	8988.748	15,462.394
Gamma	Identity	1	3532.422	6558.867	8987.824	8983.084	15,456.593
Gaussian	Log	1	3525.947	6553.477	8977.639	8976.490	15,447.349
Inverse Gaussian	Log	1	3530.944	6558.233	9002.912	9009.775	15,461.177
Poisson	Log	1	3526.854	6554.225	8980.842	8980.611	[15,448.806]
Gamma	Log	1	3528.611	6555.931	8988.305	8990.651	15,453.142
Gaussian	Identity	2	3530.321	6556.845	8981.013	8975.399	15,450.947
Inverse Gaussian	Identity	2	3533.866	6560.396	8993.318	8988.820	15,462.888
Gamma	Identity	2	3532.425	6558.884	8987.953	8983.125	15,456.791
Gaussian	Log	2	[3525.900]	[6553.471]	[8975.837]	8976.386	15,447.391
Inverse Gaussian	Log	2	3530.938	6558.258	8999.703	9009.797	15,461.917
Gamma	Log	2	3528.617	6555.944	8985.700	8990.591	15,453.323

a Degree refers to the polynomial order of the propensity score.

Note
The most appropriate GLMs for costs and QALYs were selected looking at RMSEs (in square brackets).

Appendix 6 Forest plots of estimated incremental costs, quality-adjusted life-years and life-years of emergency surgery compared with non-emergency surgery from the local instrumental variables approach

Category and subgroup	Difference in means (95% CI)	Difference in means (95% CI)	Difference in means (95% CI)
Full sample			
All (*n* = 262,371)	−109 (−1139 to 913)	−0.010 (−0.023 to 0.003)	−0.003 (0.006 to −0.001)
Age (years)			
< 45 (*n* = 175,403)	−382 (−1452 to 689)	0.006 (−0.003 to 0.0019)	0.000 (−0.002 to 0.002)
45–49 (*n* = 19,417)	−52 (−1823 to 1719)	−0.026 (−0043 to −0.006)	0.002 (−0.013 to 0.017)
50–54 (*n* = 17,079)	35 (−1623 to 1893)	−0.036 (−0059 to −0.014)	−0.000 (−0.009 to 0.006)
55–59 (*n* = 13,538)	145 (−1775 to 2065)	−0.054 (−0076 to −0.032)	−0.003 (−0.007 to 0.001)
60–64 (*n* = 10,944)	1023 (−609 to 2655)	−0.040 (−0076 to −0.006)	−0.005 (−0.012 to 0.002)
65–69 (*n* = 9258)	5 (−2080 to 2091)	−0.046 (−0089 to −0.002)	−0.006 (−0.017 to 0.006)
70–74 (*n* = 6656)	946 (−1654 to 3550)	−0.070 (−0.119 to −0.021)	−0.022 (−0.035 to 0.009)
75–79 (*n* = 4649)	891 (−2800 to 4583)	−0.061 (−0.141 to −0.021)	−0.039 (−0.060 to 0.019)
80–84 (*n* = 2978)	2295 (−2078 to 6667)	−0.130 (−0.206 to −0.054)	−0.099 (−0.147 to 0.052)
84+ (*n* = 2247)	3619 (−1251 to 6459)	−0.011 (−0.120 to −0.097)	−0.076 (−0.161 to 0.009)
Sex			
Male (*n* = 141,234)	−577 (−1727 to 572)	0.025 (0.006 to 0.044)	−0.004 (−0.006 to −0.001)
Female (*n* = 121,137)	437 (−512 to 1356)	−0.050 (−0.057 to −0.043)	−0.003 (−0.006 to −0.001)
SCARF index			
Fit (*n* = 216,833)	−483 (−1525 to 559)	−0.006 (−0.018 to −0.007)	0.000 (−0.002 to 0.003)
Mild frailty (*n* = 37,914)	531 (−684 to 1747)	−0.025 (−0.045 to −0.005)	−0.007 (−0.012 to −0.002)
Moderate frailty (*n* = 6103)	5266 (3456 to 7077)	−0.024 (−0.072 to 0.024)	−0.049 (−0.077 to −0.021)
Severe frailty (*n* = 1521)	15,715 (11,181 to 20,249)	−0.150 (−0.243 to 0.058)	−0.223 (−0.311 to −0.135)
Charlson Comorbidity Index			
No comorbidities (*n* = 219,005)	−316 (−1357 to 725)	−0.007 (−0.020 to 0.005)	−0.000 (−0.003 to 0.002)
One comorbidity (*n* = 38,904)	244 (−1003 to 1490)	−0.013 (−0.033 to 0.006)	−0.006 (−0.012 to −0.003)
Two comorbidities (*n* = 4640)	5306 (3649 to 6968)	−0.056 (−0.103 to 0.008)	−0.063 (−0.097 to −0.029)
Three or more comorbidities (*n* = 822)	7570 (2007 to 13,133)	−0.212 (−0.334 to 0.069)	−0.291 (−0.399 to −0.164)
Subdiagnosis			
K350 – Acute appendicitis with generalized peritonitis (*n* = 11,966)	−738 (−1893 to 417)	−0.006 (−0.022 to 0.010)	−0.006 (−0.011 to −0.001)
K351 – Acute appendicitis with peritoneal abscess (*n* = 2451)	−724 (−1612 to 164)	−0.015 (−0.029 to 0.000)	−0.006 (−0.011 to 0.000)
K352 – Acute appendicitis with generalized peritonitis (*n* = 10,251)	11 (−1274 to 1295)	−0.019 (−0.037 to −0.000)	−0.008 (−0.011 to −0.004)
K353 – Acute appendicitis with localized peritonitis (*n* = 56,143)	327 (−772 to 1426)	−0.021 (−0.038 to −0.004)	−0.006 (−0.009 to −0.003)
K358 – Acute appendicitis, other and unspecified (*n* = 115,611)	−4 (−1096 to 1066)	−0.009 (−0.022 to 0.004)	−0.002 (−0.005 to −0.000)
K359 – Acute appendicitis, unspecified (*n* = 28,156)	−882 (−1959 to 195)	0.002 (−0.010 to 0.014)	−0.001 (−0.005 to 0.002)
K37 – Unspecified appendicitis (*n* = 35,741)	−329 (−1139 to 482)	−0.002 (−0.009 to 0.006)	−0.002 (−0.005 to 0.001)
Year			
2010/11 (*n* = 24,787)	−1559 (−3190 to 73)	−0.005 (−0.022 to 0.011)	−0.003 (−0.005 to −0.001)
2011/12 (*n* = 25,383)	−171 (−1054 to 712)	0.005 (−0.010 to 0.020)	−0.002 (−0.008 to 0.006)
2012/13 (*n* = 25,369)	−349 (−1516 to 818)	0.001 (−0.015 to 0.016)	−0.003 (−0.005 to 0.001)
2013/14 (*n* = 26,656)	−633 (−2666 to 1399)	−0.017 (−0.035 to 0.002)	−0.003 (−0.006 to 0.000)
2014/15 (*n* = 26,642)	−565 (−2221 to 1091)	−0.019 (−0.038 to 0.000)	−0.004 (−0.007 to −0.001)
2015/16 (*n* = 27,349)	−1097 (−3339 to 1146)	−0.002 (−0.027 to 0.022)	−0.003 (−0.005 to 0.001)
2016/17 (*n* = 27,465)	322 (−688 to 1532)	−0.007 (−0.020 to 0.033)	−0.002 (−0.009 to 0.005)
2017/18 (*n* = 27,646)	997 (−60 to 2053)	−0.025 (−0.041 to 0.010)	−0.004 (−0.006 to −0.001)
2018/19 (*n* = 28,746)	926 (−118 to 1969)	−0.019 (−0.038 to −0.000)	−0.005 (−0.009 to −0.002)
2019/20 (*n* = 21,928)	1010 (62 to 1958)	−0.021 (−0.037 to −0.005)	−0.005 (−0.009 to −0.001)

−10,000   0   10,000   20,000   30,000   −0.3 −0.2 −0.1 0.0 0.1 0.2   −0.4 −0.3 −0.2 −0.1 0.0 0.1

Favours ES   Favours NES   Favours NES   Favours ES   Favours ES   Favours NES

FIGURE 31 Forest plots of ES vs. NES from the LIVs approach for appendicitis. (a) Estimated incremental costs; (b) QALYs; and (c) life-years.

Category and subgroup	Difference in means (95% CI)	Difference in means (95% CI)	Difference in means (95% CI)
Full sample			
All (*n* = 237,042)	−77 (−702 to 548)	0.007 (−0.001 to 0.015)	−0.009 (−0.022 to 0.005)
Age			
< 45 (*n* = 75,453)	−109 (−537 to 318)	0.000 (−0.002 to 0.003)	−0.001 (−0.003 to 0.001)
45–49 (*n* = 20,253)	−533 (−1020 to −46)	−0.001 (−0.003 to 0.005)	−0.001 (−0.006 to 0.005)
50–54 (*n* = 20,757)	−760 (−1333 to −187)	0.009 (0.001 to 0.016)	0.007 (−0.003 to 0.011)
55–59 (*n* = 19,363)	−844 (−1830 to 142)	0.010 (−0.000 to 0.020)	0.008 (0.002 to 0.015)
60–64 (*n* = 18,515)	−170 (−1186 to −846)	0.000 (−0.014 to 0.014)	0.001 (−0.010 to 0.012)
65–69 (*n* = 19,367)	−790 (−1776 to 196)	0.012 (−0.001 to 0.025)	0.007 (−0.004 to 0.018)
70–74 (*n* = 18,574)	−181 (−1589 to 1227)	0.013 (−0.007 to 0.034)	−0.002 (−0.026 to 0.022)
75–79 (*n* = 16,534)	−309 (−2028 to 1410)	0.016 (−0.010 to 0.042)	−0.028 (−0.066 to 0.009)
80–84 (*n* = 13,957)	1735 (−731 to 4201)	0.005 (−0.034 to 0.043)	−0.042 (−0.106 to 0.021)
84+ (*n* = 13,254)	2697 (61 to 5332)	0.039 (−0.017 to 0.094)	−0.102 (−0.213 to 0.008)
Sex			
Male (*n* = 76,721)	29 (−809 to 857)	0.018 (0.007 to 0.029)	−0.018 (−0.038 to 0.003)
Female (*n* = 160,321)	−127 (−714 to 459)	0.002 (−0.006 to 0.010)	−0.004 (−0.015 to 0.007)
SCARF index			
Fit (*n* = 146,432)	−672 (−1063 to −280)	0.002 (−0.003 to 0.008)	−0.003 (−0.001 to 0.007)
Mild frailty (*n* = 55,218)	−179 (−976 to 619)	0.003 (−0.009 to 0.015)	−0.011 (−0.031 to 0.009)
Moderate frailty (*n* = 19,308)	2050 (49 to 4052)	0.045 (0.019 to 0.074)	−0.035 (−0.091 to 0.022)
Severe frailty (*n* = 6084)	8580 (3700 to 13,459)	0.044 (−0.019 to 0.107)	−0.185 (−0.304 to −0.066)
Charlson Comorbidity Index			
No comorbidities (*n* = 154,393)	−485 (−979 to 10)	0.000 (−0.006 to 0.005)	0.002 (−0.004 to −0.007)
One comorbidity (*n* = 61,1135)	139 (−647 to 925)	0.010 (−0.003 to 0.023)	−0.019 (−0.043 to 0.005)
Two comorbidities (*n* = 16,784)	2223 (−157 to 4504)	0.043 (0.013 to 0.072)	−0.049 (−0.106 to 0.007)
Three or more comorbidities (*n* = 4230)	2507 (−1450 to 6673)	0.080 (−0.004 to 0.164)	−0.068 (−0.184 to 0.048)
Subdiagnosis			
K800 – Calculus of gallbladder with acute cholecystis (*n* = 85,597)	9 (−595 to 514)	0.008 (−0.000 to 0.016)	−0.010 (−0.024 to 0.004)
K801 – Calculus of gallbladder with other cholecystis (*n* = 65,253)	−106 (−675 to 462)	0.007 (−0.001 to 0.015)	−0.008 (−0.022 to 0.005)
K802 – Calculus of gallbladder without other cholecystis (*n* = 85,192)	−141 (−842 to 560)	0.007 (−0.002 to 0.015)	−0.007 (−0.020 to 0.006)
Year			
2010/11 (*n* = 21,321)	77 (−1011 to 1165)	0.017 (0.001 to 0.033)	−0.003 (−0.025 to 0.020)
2011/12 (*n* = 22,339)	−993 (−2037 to 51)	0.028 (0.015 to 0.042)	0.010 (−0.004 to 0.024)
2012/13 (*n* = 22,084)	8 (−1289 to 1304)	0.002 (−0.013 to 0.017)	−0.015 (−0.040 to 0.007)
2013/14 (*n* = 23,395)	−1088 (−1866 to −311)	0.008 (−0.005 to 0.021)	−0.003 (−0.021 to 0.014)
2014/15 (*n* = 23,778)	331 (−748 to 1410)	−0.001 (−0.016 to 0.014)	−0.014 (−0.038 to 0.010)
2015/16 (*n* = 25,109)	−738 (−1516 to 39)	0.009 (−0.003 to 0.021)	−0.006 (−0.022 to 0.011)
2016/17 (*n* = 25,203)	8 (−1059 to 1074)	0.011 (−0.001 to 0.023)	−0.007 (−0.024 to 0.011)
2017/18 (*n* = 25,315)	−223 (−963 to 517)	0.008 (−0.004 to 0.019)	−0.006 (−0.022 to 0.009)
2018/19 (*n* = 25,995)	884 (−44 to 1813)	−0.005 (−0.017 to 0.008)	−0.014 (−0.032 to 0.004)
2019/20 (*n* = 21,453)	929 (89 to 1768)	−0.002 (−0.016 to 0.012)	−0.027 (−0.051 to −0.004)

−10,000   0   10,000   30,000
Favours ES   Favours NES

−0.1   0.0   0.1   0.2   0.3
Favours NES   Favours ES

−0.4 −0.3 −0.2 −0.1   0.0   0.1
Favours NES   Favours ES

FIGURE 32 Forest plots of ES vs. NES from the LIVs approach for cholelithiasis. (a) Estimated incremental costs; (b) QALYs; and (c) life-years.

FIGURE 33 Forest plots of ES vs. NES from the LIVs approach for diverticular disease. (a) Estimated incremental costs; (b) QALYs; and (c) life-years.

Category and subgroup	Difference in means (95% CI)	Difference in means (95% CI)	Difference in means (95% CI)
Full sample			
All (n = 104,913)	691 (21 to 1762)	0.039 (0.004 to 0.073)	0.046 (0.004 to 0.068)
Age (years)			
< 45 (n = 19,645)	−504 (−1956 to 756)	−0.001 (−0.014 to 0.012)	0.018 (−0.003 to 0.038)
45–49 (n = 7828)	−1345 (−2570 to −121)	0.037 (0.003 to 0.072)	0.033 (−0.013 to 0.079)
50–54 (n = 8061)	−1454 (−2635 to −73)	0.047 (0.009 to 0.065)	0.105 (0.029 to 0.162)
55–59 (n = 7811)	−102 (−1727 to 1522)	0.005 (−0.025 to 0.035)	0.039 (−0.005 to 0.083)
60–64 (n = 8194)	1423 (−331 to 3176)	0.007 (−0.036 to 0.050)	0.012 (−0.025 to 0.050)
65–69 (n = 9036)	−336 (−1567 to 1196)	0.013 (−0.037 to 0.063)	0.022 (−0.026 to 0.070)
70–74 (n = 10,167)	1024 (−553 to 2602)	0.046 (−0.014 to 0.107)	0.050 (−0.012 to 0.112)
75–79 (n = 10,624)	2757 (1200 to 4315)	0.056 (−0.012 to 0.125)	0.056 (−0.016 to 0.128)
80–84 (n = 10,716)	2379 (404 to 3954)	0.076 (−0.002 to 0.155)	0.066 (−0.017 to 0.149)
84 + (n = 12,831)	4226 (2420 to 6037)	0.102 (0.025 to 0.178)	0.076 (−0.016 to 0.167)
Sex			
Male (n = 67,815)	113 (−790 to 1918)	0.043 (0.014 to 0.072)	0.041 (0.007 to 0.074)
Female (n = 37,096)	2314 (1210 to 3416)	0.030 (−0.022 to 0.062)	0.056 (−0.008 to 0.121)
SCARF index			
Fit (n = 56,996)	−1224 (−2061 to −387)	0.041 (0.015 to 0.067)	0.047 (0.014 to 0.079)
Mild frailty (n = 32,264)	609 (−202 to 1660)	0.066 (0.019 to 0.113)	0.061 (0.023 to 0.139)
Moderate frailty (n = 12,205)	5456 (3456 to 7457)	−0.009 (−0.072 to 0.054)	−0.010 (−0.078 to 0.059)
Severe frailty (n = 4441)	15,369 (11,326 to 19,451)	−0.059 (−0.156 to 0.038)	−0.061 (−0.173 to 0.060)
Charlson Comorbidity Index			
No comorbidities (n = 64,570)	113 (−741 to 967)	0.026 (−0.002 to 0.057)	0.029 (−0.004 to 0.063)
One comorbidity (n = 29,262)	1206 (−50 to 2462)	0.064 (0.013 to 0.115)	0.077 (0.011 to 0.143)
Two comorbidities (n = 5625)	4536 (2523 to 6553)	0.053 (−0.015 to 0.121)	0.083 (0.005 to 0.162)
Three or more comorbidities (n = 2256)	4825 (1517 to 6130)	−0.037 (−0.162 to 0.068)	−0.022 (−0.143 to 0.100)
Subdiagnosis			
Inguinal (n = 50,261)	592 (−295 to 1460)	0.046 (0.015 to 0.076)	0.040 (0.007 to 0.074)
Femoral (n = 13,260)	2250 (1261 to 3299)	0.072 (−0.029 to 0.172)	0.085 (−0.027 to 0.195)
Umbilical (n = 39,227)	743 (−176 to 1663)	0.021 (−0.003 to 0.045)	0.041 (0.008 to 0.075)
Ventral (n = 2145)	2006 (792 to 3225)	−0.001 (−0.036 to 0.034)	0.024 (−0.014 to 0.061)
Bilateral (n = 3290)	1176 (219 to 2130)	0.041 (0.013 to 0.069)	0.034 (0.005 to 0.063)
Obstruction (n = 46,563)	1453 (536 to 2370)	0.054 (−0.010 to 0.116)	0.063 (−0.010 to 0.135)
Gangrene (n = 3279)	2297 (1278 to 3315)	0.059 (−0.025 to 0.143)	0.069 (−0.26 to 0.163)
Year			
2010/11 (n = 9022)	−945 (−2759 to 666)	0.057 (−0.010 to 0.125)	0.076 (0.005 to 0.158)
2011/12 (n = 9415)	1229 (−442 to 2901)	0.096 (0.025 to 0.166)	0.102 (0.016 to 0.168)
2012/13 (n = 9649)	1189 (−655 to 3234)	0.100 (0.006 to 0.167)	0.097 (0.011 to 0.184)
2013/14 (n = 10,018)	−370 (−1861 to 1142)	0.056 (0.007 to 0.109)	0.045 (−0.010 to 0.100)
2014/15 (n = 10,116)	1160 (−287 to 2607)	0.045 (−0.006 to 0.097)	0.045 (−0.012 to 0.102)
2015/16 (n = 10,327)	811 (−675 to 2297)	0.046 (−0.002 to 0.094)	0.053 (−0.002 to 0.107)
2016/17 (n = 11,380)	966 (−415 to 2346)	0.001 (−0.042 to 0.043)	0.011 (−0.032 to 0.054)
2017/18 (n = 11,753)	689 (−930 to 2306)	0.006 (−0.030 to 0.045)	0.025 (−0.017 to 0.067)
2018/19 (n = 12,948)	963 (−515 to 2441)	0.028 (−0.011 to 0.067)	0.039 (−0.002 to 0.061)
2019/20 (n = 10,285)	3017 (361 to 5673)	−0.033 (−0.069 to 0.003)	−0.012 (−0.044 to 0.020)

−10,000 0 10,000 20,000 30,000	−0.2 −0.1 0.0 0.1 0.2 0.3	−0.2 −0.1 0.0 0.1 0.2 0.3
Favours NES Favours ES	Favours NES Favours ES	Favours NES Favours ES

FIGURE 34 Forest plots of ES vs. NES from the LIVs approach for hernia. (a) Estimated incremental costs; (b) QALYs; and (c) life-years.

Category and subgroup	Difference in means (95% CI)	Difference in means (95% CI)	Difference in means (95% CI)
Full sample			
All (*n* = 130,921)	−536 (−3448 to 2376)	0.010 (−0.015 to 0.035)	0.007 (−0.058 to 0.072)
Age			
< 45 (*n* = 14,476)	824 (−3107 to 4754)	0.016 (−0.019 to 0.054)	0.076 (0.037 to 0.116)
45–49 (*n* = 6401)	−3511 (−8383 to 1361)	0.026 (−0.010 to 0.067)	0.202 (0.140 to 0.265)
50–54 (*n* = 8052)	−7862 (−16,124 to 400)	0.032 (−0.010 to 0.073)	0.203 (0.147 to 0.258)
55–59 (*n* = 9067)	−4502 (−12,553 to 3548)	0.047 (0.010 to 0.063)	0.213 (0.159 to 0.266)
60–64 (*n* = 10,868)	−3325 (−7899 to 1249)	0.031 (−0.004 to 0.065)	0.209 (0.152 to 0.266)
65–69 (*n* = 13,678)	−3094 (−7384 to 1197)	0.032 (0.001 to 0.063)	0.188 (0.135 to 0.241)
70–74 (*n* = 15,591)	−1818 (−5705 to 2069)	0.001 (−0.028 to 0.030)	0.066 (−0.017 to 0.148)
75–79 (*n* = 16,588)	−464 (−5812 to 4885)	0.004 (−0.026 to 0.033)	0.019 (−0.067 to 0.126)
80–84 (*n* = 16,183)	1745 (−2171 to 5661)	−0.032 (0.058 to 0.005)	−0.174 (−0.321 to −0.027)
84+ (*n* = 19,997)	6546 (1901 to 11,191)	−0.009 (−0.038 to 0.019)	−0.421 (−0.615 to −0.228)
Sex			
Male (*n* = 62,035)	615 (−2912 to 4142)	−0.023 (−0.051 to 0.005)	−0.068 (−0.150 to 0.013)
Female (*n* = 68,886)	−1571 (−4495 to 1352)	0.039 (0.013 to 0.064)	0.075 (0.019 to 0.131)
SCARF index			
Fit (*n* = 59,168)	−7350 (−10,195 to −4508)	0.025 (−0.007 to 0.057)	0.196 (0.169 to 0.227)
Mild frailty (*n* = 43,401)	221 (−2918 to 3359)	−0.011 (−0.035 to 0.013)	0.017 (−0.080 to 0.114)
Moderate frailty (*n* = 19,738)	9203 (3562 to 14,844)	−0.006 (−0.031 to 0.026)	−0.373 (−0.523 to −0.223)
Severe frailty (*n* = 8614)	20,149 (13,681 to 26,617)	0.042 (0.007 to 0.077)	−0.486 (−0.643 to −0.328)
Charlson Comorbidity Index			
No comorbidities (*n* = 68,426)	78 (−3076 to 3232)	−0.026 (−0.050 to 0.003)	−0.016 (−0.049 to 0.018)
One comorbidity (*n* = 41,505)	−1282 (−4603 to 2038)	0.011 (−0.015 to 0.037)	−0.003 (−0.094 to 0.087)
Two comorbidities (*n* = 16,424)	−85 (−4542 to 4373)	0.109 (0.064 to 0.154)	0.086 (−0.066 to 0.238)
Three or more comorbidities (*n* = 4566)	−4563 (−14,691 to 5564)	0.180 (0.093 to 0.267)	0.153 (−0.094 to 0.400)
Subdiagnosis			
K561 – Intussusception (*n* = 1435)	−1822 (−4986 to 1342)	0.010 (−0.011 to 0.031)	0.128 (0.067 to 0.190)
K562 – Volvulus (*n* = 15,843)	1442 (−1727 to 4610)	−0.001 (−0.026 to 0.012)	−0.080 (−0.184 to 0.025)
K563 – Gallstone ileus (*n* = 1998)	2728 (−582 to 6038)	−0.014 (−0.029 to 0.002)	−0.126 (−0.165 to 0.067)
K565 – Intestinal adhesions (bands) with obstruction (*n* = 45,287)	−1148 (−4020 to 1724)	−0.002 (−0.016 to 0.012)	0.036 (−0.003 to 0.076)
K566 – Other and unspecified intestinal obstruction (*n* = 66,358)	−660 (−3568 to 2248)	0.021 (−0.017 to 0.059)	0.009 (−0.066 to 0.104)
C18 – Malignant neoplasm of colon (*n* = 3473)	−5372 (−8985 to 1780)	0.053 (0.035 to 0.070)	0.195 (0.119 to 0.272)
C19 – Malignant neoplasm of rectosigmoid junction (*n* = 319)	−4502 (−8307 to −697)	0.049 (0.030 to 0.068)	0.152 (0.068 to 0.235)
C20 – Malignant neoplasm of rectum (*n* = 558)	−6934 (−10,642 to −3226)	0.056 (0.045 to 0.068)	0.256 (0.185 to 0.327)
Year			
2010/11 (*n* = 10,844)	159 (−4127 to 4446)	0.027 (−0.009 to 0.062)	0.046 (−0.043 to 0.134)
2011/12 (*n* = 11,713)	476 (−3672 to 4624)	0.016 (−0.015 to 0.050)	−0.010 (−0.094 to 0.074)
2012/13 (*n* = 11,933)	768 (−3320 to 4853)	0.044 (0.009 to 0.079)	0.105 (0.024 to 0.186)
2013/14 (*n* = 12,722)	−1722 (−5939 to 2495)	0.003 (−0.028 to 0.033)	−0.026 (−0.108 to 0.055)
2014/15 (*n* = 13,263)	−3079 (−7103 to 945)	−0.001 (−0.032 to 0.031)	−0.011 (−0.089 to 0.067)
2015/16 (*n* = 13,726)	−329 (−4434 to 3776)	−0.010 (−0.041 to 0.026)	−0.029 (−0.110 to 0.052)
2016/17 (*n* = 14,058)	−1554 (−5009 to 1902)	0.009 (−0.023 to 0.042)	0.005 (−0.075 to 0.085)
2017/18 (*n* = 14,660)	1170 (−2346 to 4686)	0.019 (−0.011 to 0.049)	0.022 (−0.050 to 0.094)
2018/19 (*n* = 15,706)	−4467 (−10,684 to 1750)	0.012 (−0.042 to 0.018)	−0.024 (−0.102 to 0.055)
2019/20 (*n* = 12,296)	4517 (413 to 8621)	0.010 (−0.022 to 0.042)	0.011 (−0.68 to 0.090)

−30,000 −20,000 −10,000 0 10,000 20,000 30,000 40,000
Favours ES | Favours NES

−0.2 −0.1 0.0 0.1 0.2 0.3 0.4
Favours NES | Favours ES

−0.7 −0.6 −0.5 −0.4 −0.3 −0.2 −0.1 0.0 0.1 0.2 0.3 0.4 0.5
Favours NES | Favours ES

FIGURE 35 Forest plots of ES vs. NES from the LIVs approach for intestinal obstruction. (a) Estimated incremental costs; (b) QALYs; and (c) life-years.