NIHR | National Institute for Health and Care Research

Journals Library

Check for updates

Health Technology Assessment

Volume 28 • Issue 26 • June 2024

ISSN 2046-4924

Laparoscopic cholecystectomy versus conservative management for adults with uncomplicated symptomatic gallstones: the C-GALL RCT

Karen Innes, Irfan Ahmed, Jemma Hudson, Rodolfo Hernández, Katie Gillies, Rebecca Bruce, Victoria Bell, Alison Avenell, Jane Blazeby, Miriam Brazzelli, Seonaidh Cotton, Bernard Croal, Mark Forrest, Graeme MacLennan, Peter Murchie, Samantha Wileman and Craig Ramsay

DOI 10.3310/MNBY3104

Laparoscopic cholecystectomy versus conservative management for adults with uncomplicated symptomatic gallstones: the C-GALL RCT

Karen Innes,[1*] Irfan Ahmed,[2] Jemma Hudson,[1] Rodolfo Hernández,[3] Katie Gillies,[1] Rebecca Bruce,[1] Victoria Bell,[1] Alison Avenell,[1] Jane Blazeby,[4] Miriam Brazzelli,[1] Seonaidh Cotton,[1] Bernard Croal,[5] Mark Forrest,[1] Graeme MacLennan,[1] Peter Murchie,[6] Samantha Wileman[1] and Craig Ramsay[1]

[1]Health Services Research Unit, University of Aberdeen, Aberdeen, UK
[2]Department of Surgery, NHS Grampian, Aberdeen, UK
[3]Health Economics Research Unit, University of Aberdeen, Aberdeen, UK
[4]Center for Surgical Research, NIHR Bristol and Western Biomedical Research Centre, University of Bristol, Bristol, UK
[5]Clinical Biochemistry, NHS Grampian, Aberdeen, UK
[6]Academic Primary Care, Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, UK

[*]Corresponding author

Published June 2024
DOI: 10.3310/MNBY3104

This report should be referenced as follows:

Innes K, Ahmed I, Hudson J, Hernández R, Gillies K, Bruce R, *et al.* Laparoscopic cholecystectomy versus conservative management for adults with uncomplicated symptomatic gallstones: the C-GALL RCT. *Health Technol Assess* 2024;**28**(26). https://doi.org/10.3310/MNBY3104

Health Technology Assessment

ISSN 2046-4924 (Online)

Impact factor: 3.6

A list of Journals Library editors can be found on the NIHR Journals Library website

Launched in 1997, *Health Technology Assessment* (HTA) has an impact factor of 3.6 and is ranked 32nd (out of 105 titles) in the 'Health Care Sciences & Services' category of the Clarivate 2022 Journal Citation Reports (Science Edition). It is also indexed by MEDLINE, CINAHL (EBSCO Information Services, Ipswich, MA, USA), EMBASE (Elsevier, Amsterdam, the Netherlands), NCBI Bookshelf, DOAJ, Europe PMC, the Cochrane Library (John Wiley & Sons, Inc., Hoboken, NJ, USA), INAHTA, the British Nursing Index (ProQuest LLC, Ann Arbor, MI, USA), Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and the Science Citation Index Expanded™ (Clarivate™, Philadelphia, PA, USA).

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

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

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta.

Criteria for inclusion in the *Health Technology Assessment* journal
Manuscripts are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

This article

The research reported in this issue of the journal was funded by the HTA programme as award number 14/192/71. The contractual start date was in April 2016. The draft manuscript began editorial review in April 2022 and was accepted for publication in February 2023. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' manuscript and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this article.

This article presents independent research funded by the National Institute for Health and Care Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care.

This article was published based on current knowledge at the time and date of publication. NIHR is committed to being inclusive and will continually monitor best practice and guidance in relation to terminology and language to ensure that we remain relevant to our stakeholders.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Newgen Digitalworks Pvt Ltd, Chennai, India (www.newgen.co).

Abstract

Laparoscopic cholecystectomy versus conservative management for adults with uncomplicated symptomatic gallstones: the C-GALL RCT

Karen Innes,[1*] Irfan Ahmed,[2] Jemma Hudson,[1] Rodolfo Hernández,[3] Katie Gillies,[1] Rebecca Bruce,[1] Victoria Bell,[1] Alison Avenell,[1] Jane Blazeby,[4] Miriam Brazzelli,[1] Seonaidh Cotton,[1] Bernard Croal,[5] Mark Forrest,[1] Graeme MacLennan,[1] Peter Murchie,[6] Samantha Wileman[1] and Craig Ramsay[1]

[1]Health Services Research Unit, University of Aberdeen, Aberdeen, UK
[2]Department of Surgery, NHS Grampian, Aberdeen, UK
[3]Health Economics Research Unit, University of Aberdeen, Aberdeen, UK
[4]Center for Surgical Research, NIHR Bristol and Western Biomedical Research Centre, University of Bristol, Bristol, UK
[5]Clinical Biochemistry, NHS Grampian, Aberdeen, UK
[6]Academic Primary Care, Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, UK

[*]Corresponding author chart@abdn.ac.uk

Background: Gallstone disease is a common gastrointestinal disorder in industrialised societies. The prevalence of gallstones in the adult population is estimated to be approximately 10–15%, and around 80% remain asymptomatic. At present, cholecystectomy is the default option for people with symptomatic gallstone disease.

Objectives: To assess the clinical and cost-effectiveness of observation/conservative management compared with laparoscopic cholecystectomy for preventing recurrent symptoms and complications in adults presenting with uncomplicated symptomatic gallstones in secondary care.

Design: Parallel group, multicentre patient randomised superiority pragmatic trial with up to 24 months follow-up and embedded qualitative research. Within-trial cost–utility and 10-year Markov model analyses. Development of a core outcome set for uncomplicated symptomatic gallstone disease.

Setting: Secondary care elective settings.

Participants: Adults with symptomatic uncomplicated gallstone disease referred to a secondary care setting were considered for inclusion.

Interventions: Participants were randomised 1 : 1 at clinic to receive either laparoscopic cholecystectomy or observation/conservative management.

Main outcome measures: The primary outcome was quality of life measured by area under the curve over 18 months using the Short Form-36 bodily pain domain. Secondary outcomes included the Otago gallstones' condition-specific questionnaire, Short Form-36 domains (excluding bodily pain), area under the curve over 24 months for Short Form-36 bodily pain domain, persistent symptoms, complications and need for further treatment. No outcomes were blinded to allocation.

Results: Between August 2016 and November 2019, 434 participants were randomised (217 in each group) from 20 United Kingdom centres. By 24 months, 64 (29.5%) in the observation/conservative management group and 153 (70.5%) in the laparoscopic cholecystectomy group had received surgery, median time to surgery of 9.0 months (interquartile range, 5.6–15.0) and 4.7 months (interquartile range 2.6–7.9), respectively.

At 18 months, the mean Short Form-36 norm-based bodily pain score was 49.4 (standard deviation 11.7) in the observation/conservative management group and 50.4 (standard deviation 11.6) in the laparoscopic cholecystectomy group. The mean area under the curve over 18 months was 46.8 for both groups with no difference: mean difference −0.0, 95% confidence interval (−1.7 to 1.7); *p*-value 0.996; *n* = 203 observation/conservative, *n* = 205 cholecystectomy.

There was no evidence of differences in quality of life, complications or need for further treatment at up to 24 months follow-up. Condition-specific quality of life at 24 months favoured cholecystectomy: mean difference 9.0, 95% confidence interval (4.1 to 14.0), $p < 0.001$ with a similar pattern for the persistent symptoms score.

Within-trial cost–utility analysis found observation/conservative management over 24 months was less costly than cholecystectomy (mean difference −£1033). A non-significant quality-adjusted life-year difference of −0.019 favouring cholecystectomy resulted in an incremental cost-effectiveness ratio of £55,235. The Markov model continued to favour observation/conservative management, but some scenarios reversed the findings due to uncertainties in longer-term quality of life.

The core outcome set included 11 critically important outcomes from both patients and healthcare professionals.

Conclusions: The results suggested that in the short term (up to 24 months) observation/conservative management may be a cost-effective use of National Health Service resources in selected patients, but subsequent surgeries in the randomised groups and differences in quality of life beyond 24 months could reverse this finding. Future research should focus on longer-term follow-up data and identification of the cohort of patients that should be routinely offered surgery.

Trial registration: This trial is registered as ISRCTN55215960.

Funding: This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: 14/192/71) and is published in full in *Health Technology Assessment*; Vol. 28, No. 26. See the NIHR Funding and Awards website for further award information.

Contents

List of tables

List of figures

List of boxes

List of abbreviations

A&E	accident and emergency
AE	adverse event
AUC	area under the curve
BCT	behaviour change techniques
BNF	British National Formulary
CEAC	cost-effectiveness acceptability curve
CHaRT	Centre for Healthcare Randomised Trials
CI	confidence interval
COMET	Core Outcome Measures in Effectiveness Trials
CONSORT	Consolidated Standards of Reporting Trials
COVID-19	coronavirus disease 2019
CRF	case report form
CrI	credible interval
CSQ	condition-specific quality of life
CT	computerised tomography
DMC	Data Monitoring Committee
ECR	environmental context and resources
ERCP	endoscopic retrograde cholangiopancreatography
GCP	good clinical practice
GP	general practitioner
HCRU	healthcare resource use
HDU	high dependency unit
HEAP	health economics analysis plan
HRG	healthcare resource group
HTA	Health Technology Assessment
ICER	incremental cost-effectiveness ratio
ICU	intensive care unit
IQR	interquartile range
ISD	information statistics division
ISRCTN	International Standard Randomised Controlled Trial Number
ITT	intention to treat
KM	Kaplan–Meier
MCS	mental component summary
MD	mean difference
MI	multiple imputation
MRC	Medical Research Council
MRI	magnetic resonance imaging
NIHR	National Institute Health and Care Research
NSAIDs	non-steroidal anti-inflammatory drugs
PCS	physical component summary
PI	principal investigator
PIL	patient information leaflet
PMG	Project Management Group
PPI	patient and public involvement
PQ	participant questionnaire
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROM	patient-reported outcome measure
PSA	probabilistic sensitivity analysis
QALY	quality-adjusted life-year
QoL	quality of life
QRI	QuinteT Recruitment Intervention
Q-QAT	quanti-qualitative appointment timing
RCT	randomised controlled trial
REC	Research Ethics Committee
RN	research nurse
RR	relative risk

RUQ	right upper quadrant
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SEAR	screened, eligible, approached and randomised
SeHCAT	23-seleno-25-homotaurocholic acid, selenium homocholic acid taurine, or tauroselcholic acid
SF-36	Short Form-36 items
SF-6D	Short Form-6 Dimensions
SHARE	Scottish Health Research Register
SPRI	social professional role and identity
SWAT	Study Within a Trial
TDF	theoretical domains framework
TSC	Trial Steering Committee

Plain language summary

Gallstones are common but only around 20% cause symptoms. For patients with symptomatic gallstone disease, having your gallbladder removed by surgery (cholecystectomy) or observation (conservative management) are the usual treatment options to consider.

What did the C-GALL study do?

The C-GALL study assessed the benefits, in terms of symptoms, quality of life and costs, of cholecystectomy versus observation (conservative management: by the patient and general practitioner that might include dietary advice and pain management and surgery if needed).

Four hundred and thirty-four patients with symptomatic gallstones were randomly allocated surgery or conservative management. The main symptom of ongoing bodily pain and some other quality-of-life measures were assessed over the next 2 years using postal questionnaires.

What did the C-GALL study find?

After 2 years, 70% of those allocated to surgery had been operated on and 37% of the observation group either had an operation or were waiting for one. There was no difference in bodily pain or overall quality of life between the groups. However, participants in the surgery group reported fewer ongoing problems related to their gallstone disease or after surgery than those in the conservative management group. Surgery was, however, more costly than conservative management.

What does this mean?

The C-GALL study has shown that for some patients, a conservative management approach may be a sufficient and less costly way of managing their gallstone symptoms rather than going straight on the waiting list for surgery. More research is needed to identify which patients benefit most from surgery.

Scientific summary

Background

Gallstone disease is one of the most common gastrointestinal disorders in industrialised societies. The prevalence of gallstones in the adult population is estimated to be approximately 10–15%, and around 80% remain asymptomatic. Prevalence increases with age and obesity and is higher in women than in men. At present, cholecystectomy is the default option for people with symptomatic gallstone disease. However, some people, after an initial episode of biliary pain or cholecystitis, do not experience persistent symptoms or complications. There is, therefore, an indication that uncomplicated symptomatic gallstone disease may not always require removal of the gallbladder and could be treated conservatively.

Objectives

To assess the clinical and cost-effectiveness of observation/conservative management compared with laparoscopic cholecystectomy for preventing recurrent symptoms and complications in adults presenting with uncomplicated symptomatic gallstones in secondary care.

Methods

Design

Parallel-group, multicentre patient randomised superiority pragmatic trial with up to 24 months follow-up and embedded qualitative research. Within-trial cost–utility and 10-year Markov model analyses. Development of a core outcome set for uncomplicated symptomatic gallstone disease.

Setting

Secondary care elective settings.

Participants

Adults with symptomatic uncomplicated gallstone disease referred to a secondary care setting.

Exclusion criteria

Unable to consent, medically unfit for surgery, current pregnancy, previous open major upper abdominal surgery, gallstones in the common bile duct or evidence of previous choledocholithiasis, history of acute pancreatitis, evidence of obstructive jaundice, evidence of empyema of the gallbladder with sepsis, suspicion of gallbladder cancer, perforated gallbladder (recent or old perforation) and haemolytic disease.

Primary outcome

Quality of life (QoL) measured by area under the curve (AUC) over 18 months using the Short Form-36 items (SF-36) bodily pain domain.

Secondary outcomes

Otago gallstones Condition-Specific Questionnaire (CSQ).
SF-36 domains (excluding bodily pain).
AUC over 24 months for SF-36 bodily pain.
Persistent symptoms.
Complications.
Need for further treatment.

Sample size

The sample size calculation was based on the primary outcome, AUC up to 18 months using SF-36 bodily pain. A total of 430 participants was needed to detect a 0.33 standard deviation (SD) difference with 90% with alpha at 5% and allowing 10% of participants with complete missing outcome data.

Randomisation

Participants were randomised at a 1 : 1 allocation ratio, using the randomisation application at the trial office at the Centre for Healthcare Randomised Trials (CHaRT). The minimisation variables were recruitment site, gender (male/female) and age (< 35; 35–64; ≥ 65 years) as minimisation covariates to allocate treatment. A random element (20% chance) was incorporated into the minimisation algorithm.

Primary economic outcomes

Incremental cost per quality-adjusted life-year (QALY). QALYs were estimated using participants' responses to the SF-36 questionnaire and were assigned a utility score based on the SF-6D UK tariff.

Interventions: health technology assessed

1. Laparoscopic cholecystectomy.
2. Observation/conservative management.

Results

Clinical effectiveness

Participants were recruited to the C-GALL trial between August 2016 and November 2019. In total, 434 participants were randomised from 20 centres within the UK. There were 2667 patients identified to be potentially eligible for inclusion into the trial, of which 1298 were excluded. Of the 1369 eligible patients, 933 were not randomised, with 910/933 (97.5%) having a preference. The main preference reasons were participants preferred cholecystectomy (538/910, 59.1%), observation/conservative management (167/910, 18.4%) and did not want to be randomised (91/910, 10.0%). By 18 months, 54 (24.9%) in the observation/conservative management group and 146 (67.3%) in the laparoscopic cholecystectomy group received surgery. By 24 months, 64 (29.5%) in the observation/conservative management group and 153 (70.5%) in the laparoscopic cholecystectomy group received surgery with a median time to surgery of 9.0 months [interquartile range (IQR) 5.6–15.0] and 4.7 months (IQR 2.6–7.9), respectively.

At 18 months, the mean SF-36 norm-based bodily pain score was 49.4 (SD 11.7, *n* = 135) in the observation/conservative management group and 50.4 (SD 11.6, *n* = 138) in the laparoscopic cholecystectomy group. For the primary analysis, the mean AUC over 18 months was 46.8 for both groups with no difference with 203 in the observation/conservative management group and 205 in laparoscopic cholecystectomy group providing data: mean difference (MD) −0.0, 95% confidence interval (CI) (−1.7 to 1.7); *p*-value 0.996.

The mean CSQ at 18 months was 21.3 (SD 21.0, *n* = 113) for the observation/conservative management group and 15.8 (SD 19.7, *n* = 101) for the laparoscopic cholecystectomy group and showed evidence of a difference in favour of laparoscopic cholecystectomy: MD 6.6, 95% CI (1.9 to 11.3); *p*-value 0.006. At 24 months, there was evidence of a difference in favour of laparoscopic cholecystectomy (*n* = 205) compared with observation/conservative management group (*n* = 203): MD 9.0, 95% CI (4.1 to 14.0), *p*-value < 0.001. There was a similar pattern for the persistent symptoms score.

At 18 months, 32 (10.1%) participants in the observation/conservative management group had had a complication compared with 44 (25.3%) participants in the laparoscopic cholecystectomy group, with no

evidence of a difference between groups: RR 0.72, 95% CI (0.46 to 1.14); *p*-value 0.17. At 24 months, there were an additional two complications in the laparoscopic cholecystectomy group.

Prespecified sensitivity analyses demonstrated clearly that compliance with treatment allocation, missing data and the potential impact of COVID-19 did not change the findings.

Cost-effectiveness

Cost analysis shows that observation/conservative management group was less costly than the laparoscopic cholecystectomy (MD −£1033). The trial did not demonstrate a significant difference in QALYs between the groups – a mean QALY difference of −0.019 favoured cholecystectomy. The base-case incremental cost-effectiveness ratio (ICER) was found to be high (£55,235), meaning significant potential savings to the NHS with limited QALY loss by following an observation/conservative management approach in the short term. Longer-term modelling suggested that following an observation/conservative management approach might be cost-effective, but there was greater uncertainty due to limited information on subsequent surgeries in the randomised groups, and differences in quality of life (QoL) beyond 24 months could reverse this finding. Sensitivity analysis incorporating longer-term QoL scores reduced the potential saving to just £14,700 per QALY lost. The current decision uncertainty could be reduced with a long-term follow-up of the C-GALL trial participants.

Process evaluation

The embedded process evaluation in C-GALL explored ways to improve recruitment and retention across the trial. A total of 16 sites provided 180 audio recordings of consultations for analysis. Analysis of the transcripts identified four core challenge areas for recruiters: (1) providing a balanced presentation about both treatments; (2) discussing and exploring preferences; (3) discussing uncertainty; and (4) discussing participants who did not receive their treatment allocation (crossovers in treatments). A subset of 38 audio recordings of consultations from four sites were included in the analysis of discussion of retention. Thirty (79%) of these consultations did not include any discussion of trial retention. Interviews with participants (n = 9) to explore challenges in returning postal questionnaires identified six themes influencing retention: unclear expectations of trial participation; personal attributes for questionnaire completion; significance of questionnaire non-return; commitment to returning questionnaires given other priorities; individual preferences for presentation mode and timing of the questionnaires and, internal and external strategies to encourage questionnaire return. The innovative adoption of a behavioural science approach to the process evaluation led to structured changes in written and verbal information across the trial including e-mail feedback, amendments to trial information leaflet and updates to cover letters and newsletters.

Core outcome set

The final core outcome set (COS) for symptomatic uncomplicated gallstone disease included 11 critically important outcomes from both patients and healthcare professionals. These were: QoL; overall health state; overall satisfaction; overall pain; common bile duct injury; biliary leak; haemorrhage; need for endoscopic retrograde cholangiopancreatography; intra-abdominal collections; admission/re-admission for problems; and reoperation.

Comparison with similar randomised trials

The clinical outcomes were similar to those seen in previous randomised trials. The economic outcomes, as they relate to the UK NHS, have not been evaluated in previous randomised trials.

Strengths of the study

C-GALL's strengths included the pragmatic randomised controlled trial (RCT) design and methodological rigour. The benefit of the sample size is reflected in the precision with which outcomes were estimated. This multicentre trial also gives confidence in the generalisability of findings to the NHS. Comparison with routine data suggests that the results were representative of the UK population with gallstones.

This trial was pragmatic where patients in the UK may not always receive the treatment they are offered and waiting lists for surgical treatment exist. We carefully tracked treatment after randomisation and monitored compliance. A major strength included our sensitivity analyses, including compliance analysis, imputation for missing data and potential impact of coronavirus disease 2019 (COVID-19). These analyses did not change our findings.

The cost-effectiveness analysis had several strengths. Firstly, the RCT design allowed the collection of data on resource use and QoL collected prospectively for comparable groups. Secondly, in cost analysis, critical model data was supported by RCT data (e.g. survival analysis; QoL data for the reduction in QoL weight before and after surgery).

We incorporated an embedded process evaluation with qualitative interviews to better understand and reduce recruitment and retention challenges. The process evaluation gave some generalisable findings and recommendations for good practice for ongoing and future trials including how to improve the informed consent and follow-up processes.

A key strength of the COS was the extensive outcome mapping exercise on which the COS was built. The development of a core outcome set for uncomplicated symptomatic gallstone disease will help to ensure that important outcomes to patients and the NHS are collected in the future.

Limitations of the study

An unexpected difficulty was the longer-than-expected time on the waiting list for surgery for those patients who were allocated to cholecystectomy. When designing the trial, it was anticipated that this wait would be, on average, 6 months. Therefore an 18-month follow-up was chosen as the primary outcome follow-up time to reflect a time equivalent to 12 months after surgery. However, during the study, we observed that patients often experienced longer times to surgery, initially due to limited existing NHS resources that resulted in longer waiting lists. To address this, we added a 24-month follow-up time point. Our sensitivity analyses on compliance with the treatment suggested that the waiting list was unlikely to be biasing the study findings. The existence of the waiting list may limit the generalisability to some other countries' jurisdictions. A further limitation was the non-blinding of participants and treating surgeons to allocation.

Implications for health care

Current clinical guidelines recommend laparoscopic cholecystectomy for biliary pain or acute cholecystitis and radiological evidence of gallstones. Hence, surgical management remains the default option for people with symptomatic gallstone disease, one of the most common elective surgical procedures performed in the NHS. In the UK, many people with uncomplicated symptomatic gallstone disease are put on a waiting list and operated on electively after several months. The C-GALL trial demonstrates that in adults presenting with uncomplicated symptomatic gallstones, in a secondary care setting, observation/conservative management may be more effective and cost-effective than laparoscopic cholecystectomy in the short term. The crossover between groups suggests that it remains

key to identify patients that will and will not require surgery, and as healthcare workers often underestimate surgical risks, discussion about conservative management should be part of the clinical discussion.

Implications for research

Costs and benefits will continue to be incurred in both groups beyond 24 months, so future research should focus on (1) long-term follow-up data to establish lifetime cost-effectiveness and (2) identification of the cohort of patients that should be offered surgery.

Conclusion

Overall, our results suggest that in the short term (up to 24 months) observation/conservative management may be a cost-effective use of NHS resources in selected patients, but subsequent surgeries in the randomised groups and differences in QoL beyond 24 months could reverse this finding, and longer-term follow-up is needed to verify the safety and cost-effectiveness of this approach.

Trial registration

This trial is registered as ISRCTN55215960.

Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: 14/192/71) and is published in full in *Health Technology Assessment*; Vol. 28, No. 26. See the NIHR Funding and Awards website for further award information.

Chapter 1 Introduction

Introduction

Gallstone disease (cholelithiasis) is one of the most common gastrointestinal disorders in industrialised societies. The prevalence of gallstones in the adult population is approximately 10–15%.[2–6] Gallstones are more common in women and people over the age of 40 years.

Clinical surveys conducted in Europe, North and South America, and Asia indicate that prevalence rates for gallstone disease range from 5.9% to 25% and tend to increase with age.[7–10] A clinical ultrasound survey conducted in the UK reported prevalence rates of 12% among men and 22% among women over 60 years of age.[9] A multicentre population-based study conducted in Italy has reported an annual incidence of gallstone disease of 0.66% in men and 0.81% in women.[11]

In the UK and North America, the number of surgical procedures for gallstone disease increased steadily between the 1950s and 1990s, reflecting the rise in prevalence of identified gallstone disease and the use of cholecystectomy as the treatment of choice. Rates of surgical procedures stabilised in these countries towards the end of the twentieth century.[6]

The natural course of gallstones is benign, with most people being asymptomatic and with relatively low progression from asymptomatic disease to symptomatic disease.[12] In an Italian population-based study, the overall frequency of symptom development in asymptomatic people was around 20% over a long follow-up period (mean 8.7 years).[12] Similarly, a systematic review published in 2007 reported that the progression of asymptomatic to symptomatic disease ranged from 10% to 25% in studies that followed up patients after their initial diagnosis (up to 15 years of follow-up).[13] The annual risk of developing symptoms has been estimated to be around 2–4%.[12]

Most people with symptomatic uncomplicated gallstone disease probably do not develop complications; the annual rates of developing gallstone-related complications (e.g. acute cholecystitis, acute pancreatitis, acute cholangitis obstructive jaundice) have been reported to be as low as 1–3%.[14–16] The Italian Group for the Epidemiology and Prevention of Cholelithiasis study reported an annual incidence of complications of 0.7% for symptomatic patients.[17]

Mortality from gallstone disease is rare, with typically < 1% of people dying from gallstone-related causes.[12,17,18]

From a patient's perspective, the defining symptom of gallstone disease is severe and lasting (i.e. > 30 minutes) abdominal pain.[19,20] Commonly, general abdominal symptoms intensify over a period and become regular pain attacks (biliary colic) and may require medical attention.

A recent large prospective study conducted in the UK (8909 participants) has shown that 10.8% of people experienced complications 30 days after surgery.[21] Furthermore, a proportion of people (up to 40%) may continue to experience pain and abdominal symptoms after surgery.[22] In particular, persistent pain similar to that experienced preoperatively has been reported in about 20% of people after cholecystectomy[23,24] and de novo pain has been reported in up to 14% of people.[25]

The term 'postcholecystectomy syndrome' is an umbrella term widely used to describe the range of symptoms that occur after cholecystectomy.[26] The term 'persistent postcholecystectomy symptoms' has been suggested as a more accurate description of these symptoms.[27] Symptoms include biliary and non-biliary abdominal pain, dyspepsia, heartburn, nausea, vomiting and jaundice. Persistent diarrhoea or constipation is often reported after cholecystectomy, and flatulence may arise de novo after surgery.[25] There is no consistent pathophysiological explanation for persistent postcholecystectomy symptoms and, in about 5% of people, the reason for constant abdominal pain remains unknown.[28,29]

The rationale for the trial

Current clinical guidelines recommend expectant treatment for asymptomatic gallstones and hence laparoscopic cholecystectomy is considered for biliary pain or acute cholecystitis with radiological evidence of gallstones (i.e. symptomatic gallstones).[30] As per most of the international guidelines, cholecystectomy is the default option for people with symptomatic gallstone disease,[30] and one of the most common and, in terms of total costs, costly elective surgical procedures performed in the NHS in the UK. Some 74,373 cholecystectomies were performed in England in 2019 at an average cost of £3581 per procedure[31] (61,584 of these following elective admissions).[32] These figures indicate that although some patients are operated on in the acute hospital setting, many people with uncomplicated symptomatic gallstone disease are put on a waiting list and operated on electively after several months. Mean waiting time varies according to available resources; in the UK pre-COVID it has been reported to be around 12 months.[33]

Observation and relief of symptoms, delivered mainly in primary care, may be a valid therapeutic option in people presenting with uncomplicated disease, depending on their age, clinical presentation and evolution of symptoms over time. Symptom management includes the prescription of analgesics alongside dietary advice and, when necessary, anti-inflammatory drugs or antibiotics. Moreover, as symptoms of uncomplicated gallstones are usually not urgent, it may be reasonable to consider a conservative option first, which could save a considerable amount of NHS resources.

Early natural history studies, and more recent observational and population-based studies, have suggested that a proportion of people with symptomatic gallstone disease no longer experience biliary pain after the onset of symptoms.[12,17,18,22,34] Larsen *et al.*[22] found that 45% of symptomatic people on watchful waiting were relieved from symptoms during a 1-year observation period. Similarly, Festi *et al.* observed that 58% of people with initially mild symptoms, and 52% of those with more severe symptoms, did not experience further pain episodes during a follow-up period of 10 years, or an increase in disease severity over time.[12] If about half of the people treated conservatively were likely to be symptom-free, up to 30,000 cholecystectomies per year could potentially be avoided with a likely saving for the NHS of around £68 million/year.

A National Institute of Health and Care Research (NIHR) Health Technology Assessment (HTA) found that, on average, cholecystectomy is more costly but more effective than observation/conservative treatment for symptomatic gallstones or cholecystitis.[35] Nevertheless, half of the people treated conservatively were symptom-free and did not require surgery in the long term (14-year follow-up) indicating that there is probably a proportion of patients with uncomplicated symptomatic gallstone disease who could benefit from a conservative approach. The specific results were that participants randomised to observation/conservative treatment were significantly more likely to experience gallstone-related complications [risk ratio (RR) 6.69; 95% confidence interval (CI) 1.57 to 28.51; $p = 0.01$], in particular acute cholecystitis (RR 9.55; 95% CI 1.25 to 73.27; $p = 0.03$), but less likely to undergo surgery (RR 0.50; 95% CI 0.34 to 0.73; $p = 0.0004$) and experience surgery-related complications (RR 0.36; 95% CI 0.16 to 0.81; $p = 0.01$) than those randomised to receive surgery. Fifty-five per cent of people randomised to observation/conservative treatment did not require an operation during the 14-year follow-up period, and 12% of people randomised to cholecystectomy

did not undergo the scheduled surgical operation. These results were subject to major uncertainties in the reported economic model. Even when cholecystectomy occurred after conservative management, a conservative management strategy had between 40% and 60% chance of being cost-effective for alternative values of willingness to pay for an additional quality-adjusted life-year (QALY).

Furthermore, results were strongly influenced by the proportion of individuals initially treated conservatively who subsequently required surgery. Due to the limited evidence available and the current lack of UK NHS data, the C-GALL Research Group highlighted the need for a well-designed trial assessing the effects and safety of observation/conservative treatment compared with cholecystectomy.

Aims and objectives

The *primary aim* of the study is to assess the clinical and cost-effectiveness of observation/conservative management with laparoscopic cholecystectomy for preventing recurrent symptoms and complications in adults presenting with uncomplicated symptomatic gallstones in a secondary care setting.

The *primary patient objective* is to compare observation/conservative management with laparoscopic cholecystectomy in terms of participants' quality of life (QoL) using the Short Form-36 items (SF-36) health survey bodily pain domain at up to 18 months after randomisation.

The *primary economic objective* is to assess the cost-effectiveness of observation/conservative management versus laparoscopic cholecystectomy in terms of the incremental cost per QALY.

The *secondary objectives* are to compare observation/conservative management with laparoscopic cholecystectomy in terms of condition-specific quality of life (CSQ); SF-36 domains (excluding bodily pain domain); complications; need for further treatment; persistent symptoms; healthcare resource use (HCRU); and costs. Secondary outcomes are assessed at 18 and 24 months after randomisation. The bodily pain domain of the SF-36 health survey will also be assessed up to 24 months after randomisation.

The null hypothesis being tested is that there is no difference between observation/conservative management and laparoscopic cholecystectomy. The alternative hypothesis is that laparoscopic cholecystectomy is superior.

Chapter 2 Trial design and methods

The study was prospectively registered on a publicly available website on 27 May 2016 as International Standard Randomised Controlled Trial Number (ISRCTN) 55215960.

Trial design

The study protocol has been published in an Open Access journal.[1] The study protocol and study paperwork will be available on the project webpage at https://www.fundingawards.nihr.ac.uk/award/14/192/71 (accessed January 2024).

C-GALL was a pragmatic, multicentre parallel-group patient randomised superiority trial (with internal pilot phase) to test if the strategy of laparoscopic cholecystectomy is more (cost-) effective than observation/conservative management at 18 months post randomisation. The aim was to recruit 430 adults with symptomatic uncomplicated gallstone disease (biliary pain), who were electively referred to a secondary care setting and considered suitable for cholecystectomy, to assess the clinical and cost-effectiveness of observation/conservative management with laparoscopic cholecystectomy for preventing recurrent symptoms and complications.

The trial design is summarised in *Figure 1*. Patients were recruited to the trial, and all were followed up to at least 24 months post randomisation and every 6 months thereafter to the end of the trial.

The primary patient objective was to compare observation/conservative management with laparoscopic cholecystectomy in terms of participants' QoL using the SF-36 bodily pain domain at up to 18 months after randomisation.

The primary economic objective was to assess the cost-effectiveness of observation/conservative management versus laparoscopic cholecystectomy in terms of the incremental cost per QALY. This information was collected at each follow-up time point (3, 9, 12, 18 months and 6 months thereafter post randomisation till end of trial).

Embedded process evaluation

An embedded process evaluation was incorporated into the study design to identify challenges relating to trial design and/or conduct that could be addressed and modified. The process evaluation component included, where necessary: analysis of participant flow data; audio recording of recruitment consultations with potential trial participants; and semistructured telephone interviews with patients and trial participants [trial consenters, trial non-consenters, those who crossed over trial groups, those who returned questionnaires (returners) and those who did not (non-returners)], and in-depth semistructured telephone interviews with clinical site staff. The trial process evaluation is described in more detail in *Chapter 6*.

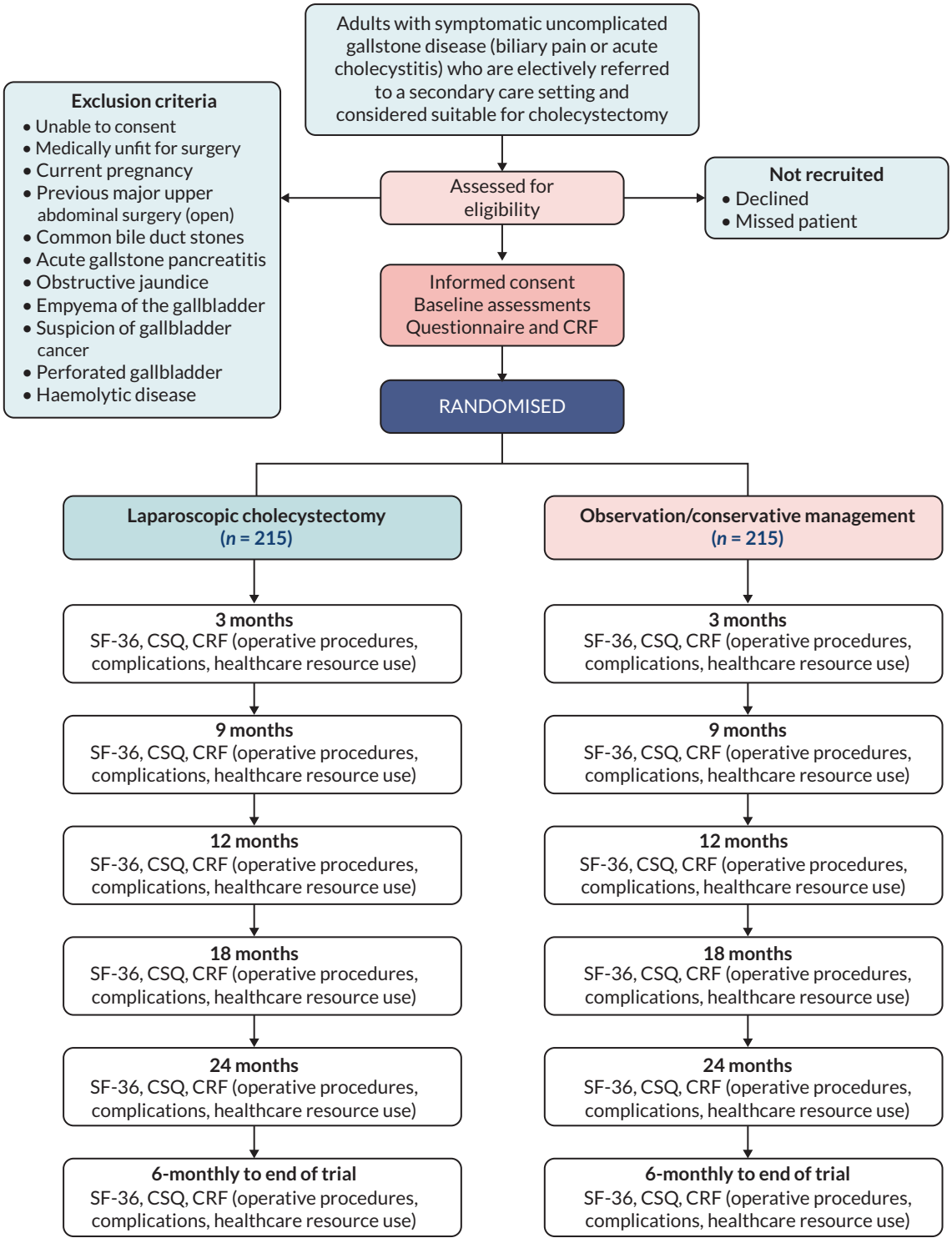

FIGURE 1 C-GALL trial design. CRF, case report form.

Participants

Participants were adults with symptomatic uncomplicated gallstone disease (biliary pain from previous biliary colic or acute cholecystitis) who were electively referred to a secondary care setting and considered suitable for cholecystectomy. Adult patients with diagnosed gallstone disease electively referred to a secondary care setting via general practitioner (GP) referral, accident and emergency (A&E) department or elsewhere, not requiring emergency surgical or endoscopic intervention were approached by the research teams. The following inclusion criteria were used to identify eligible participants.

Inclusion criteria

All adult patients with confirmed symptomatic gallstones electively referred to a secondary care setting for consultation. Clinical diagnosis of gallstone disease was confirmed by imaging. Transabdominal ultrasonography was the standard imaging technique for the diagnosis of gallbladder stones, but diagnosis by any imaging technique was acceptable.

Exclusion criteria

- Unable to consent.
- Medically unfit for surgery.
- Current pregnancy.
- Previous open major upper abdominal surgery.
- Gallstones in common bile duct or evidence of previous choledocholithiasis.
- A history of acute pancreatitis.
- Evidence of obstructive jaundice.
- Evidence of empyema of the gallbladder with sepsis.
- Suspicion of gallbladder cancer.
- Perforated gallbladder (recent or old perforation detected on imaging).
- Haemolytic disease.

Identification

Potential participants were recruited from secondary care hospitals across the UK. Participants were identified by the local research team at these participating centres. Following identification of potential participants, an invitation letter and patient information leaflet (PIL) detailing the trial were sent out, inviting them to attend a hospital outpatient clinic visit, where the trial and their treatment would be discussed. Potential participants not identified prior to a clinic visit, or at sites that were unable to send the PIL in advance, were given the PIL at their outpatient clinic visit. The PIL also highlighted that the clinical consultation might be audio-recorded; in sites who had agreed to do this, participants were asked to consent to do so.

At the hospital outpatient clinic visit, the local research team outlined the trial and asked the patient if they were willing to discuss participation and have their conversation audio-recorded. If the patient did not wish to have their conversation audio-recorded the consultation went ahead as normal, and the patient still had the choice to take part in the trial. For those patients who were happy to discuss the trial, a member of the local research team completed a trial screening form using information from the prospective participant and from the clinical record to document fulfilment of entry criteria. Eligibility criteria were then cross-checked with the patients' clinical records. If the patient was eligible and in provisional agreement, a local research team member met with the patient immediately in the clinic. Eligible participants who expressed an interest in participating had the study explained to them by local research staff and were asked if they had any questions or concerns about participating in the trial. If they agreed to take part, they gave written consent to be randomised.

Recruitment and consent

All staff involved in recruitment and consent had evidence of up-to-date good clinical practice (GCP) training. Written informed consent was sought from those patients interested in participating in the trial. Patients were given sufficient time to accept or decline involvement and were free to leave the study at any time. Patients made the decision to participate during the initial consultation, during a subsequent visit to hospital, or alternatively at home. If the patient decided to take part during the initial consultation, this became the baseline visit. For patients who decided to take part during a subsequent

visit to hospital, this became the baseline visit. If the patient agreed to be contacted at home, they received a telephone call from the local research nurse (RN) to discuss any queries. Patients who decided to participate following telephone counselling could either send their completed documents (consent form and baseline questionnaire) through the post to the local team at their treating hospital or bring it with them if they were returning to hospital for another consultation.

Randomisation/treatment allocation

Eligible participants consenting to the trial were randomised to receive either laparoscopic cholecystectomy or observation/conservative management in a 1 : 1 allocation ratio, using the randomisation application at the trial office at the Centre for Healthcare Randomised Trials (CHaRT). The randomisation application was available via a 24-hour telephone Interactive Voice Response randomisation system or a web-based application. The minimisation algorithm used recruitment site, gender (male/female) and age (< 35; 35–64; ≥ 65 years) as minimisation covariates to allocate treatment. A random element (20% chance) was incorporated into the minimisation algorithm.

For patients who consented to take part in the trial at their initial consultation, randomisation happened during this visit, and they were informed of their allocated treatment group immediately. If the patients were not present in the clinic, they were contacted by the research teams to inform them of the allocated treatment group.

Interventions evaluated

1. Laparoscopic cholecystectomy (surgical management): the current standard surgical procedure for the management of symptomatic gallstone disease. The gallbladder is removed with the stones within it using keyhole techniques (laparoscopy). The procedure is undertaken under a general anaesthetic. It usually involves three to four small incisions in the abdomen, which allow the surgeon to dissect the gallbladder from its attachments and safely divide the key anatomical structures (the cystic duct and artery) that link it to the main bile ducts. The gallbladder is then separated from the under surface of the liver. Usually the gallbladder (containing the stones) is removed within a retrieval bag via one of the small incisions. The operation takes between 45 and 120 minutes, and many patients are admitted for one night, although day-case laparoscopic cholecystectomy is safely undertaken in otherwise fit patients with appropriate social support. All surgical cases are initially started laparoscopically (keyhole) with an intention to remove the gallbladder. Occasionally it might have to be converted to open surgery either to deal with a complication or due to difficulty in progressing safely. Moreover, an alternative procedure may be performed if there is anticipated difficulty in removing gallbladder safely (i.e. drainage of the gallbladder, subtotal cholecystectomy, etc.).
2. Observation/conservative management: in the context of gallstone disease, this involves the prescription of analgesics to relieve the biliary pain, if and when required. Typical therapy includes paracetamol, non-steroidal anti-inflammatory drugs (NSAIDs) (e.g. ibuprofen), narcotic analgesics (e.g. opiates), antispasmodics (e.g. Buscopan), together with generic healthy lifestyle advice. In the longer term, observation/conservative management may involve strategies for symptom control (e.g. analgesia and antispasmodics) alongside the advice to follow a healthy diet and eat regular meals.

Participants who were randomised to the observation/conservative management group of the study were also given a copy of the medical management PIL which provided them with information about what to do if they had a flare-up of their condition and advice on diet.

Blinding of personnel in the study

Baseline data were reported by study participants before randomisation using self-completed questionnaires. Blinding was not possible due to the interventions.

Data collection

The patient-reported outcomes SF-36, CSQ and HCRU were collected at recruitment [baseline (before surgery), except HCRU] and then at 3, 9, 12, 18 and 24 months (post randomisation) and 6 monthly thereafter until the end of trial, December 2021. An additional questionnaire, participant costs, was issued at 18 months. Case report forms (CRFs) were completed after gallstone surgery had taken place (with details of operative procedures), complications, and resource use in hospital. A RN completed a CRF at 24 months for all participants to confirm if they had received any gallbladder-related surgery, and if so the type of procedure and the date it was received.

The schedule for data collection is outlined in *Table 1*.

Baseline

For each participant two CRFs were completed at baseline: participant details and baseline. The participant details CRF recorded name and full contact details, date of birth, gender and ethnicity. The baseline CRF recorded clinical information (including height, weight, information about the gallbladder, confirmation of diabetes and hypertension). The data from the CRFs were entered into the study website.

At baseline, participants completed the baseline questionnaire: SF-36 and CSQ. At the end of the baseline consultation, a reminder card was given to all participants to record information about any surgery they went on to have for their gallstones. Participants were asked to return this to the trial office in a prepaid envelope.

Follow-up

Participants were followed up by questionnaire (issued by post, e-mail or telephone) which collected patient-reported outcomes (SF-36; CSQ; HCRU) at 3, 9, 12, 18 and 24 months (post randomisation) and 6 monthly thereafter until the end of trial. If participants did not respond to the first issue of the questionnaire, a reminder was sent 3 weeks later. If no response was received after 3 weeks, then this was followed up with a telephone call, and if no contact could be made a final questionnaire for this time

TABLE 1 Schedule of data collection

Outcome measure	Baseline	Surgery	3 months	9 months	12 months	18 months	24 months	6-monthly thereafter
SF-36	X		X	X	X	X	X	X
CSQ	X		X	X	X	X	X	X
CRF	X	X	X	X	X	X	X	X
HCRU			X	X	X	X	X	X
Participant costs questionnaire						X		

point was issued. In addition, at 18 months an additional questionnaire, participant costs, was issued to collect information on participant costs in using health services.

Where data were collected by telephone, participants were offered the option of only completing the responses required to capture the primary outcome, safety and surgery data.

A participant newsletter was issued in August 2019 in an attempt to encourage questionnaire response. This included personalised information about their stage in the trial, how many questionnaires were left to complete and, more generally, trial progress.

A RN completed CRFs after any gallstone surgery had taken place, providing details of the operative procedures, complications and resource use in hospital. Costs of the initial intervention procedures were estimated from resource use data recorded on the CRFs coupled with routine unit cost data. Costs associated with subsequent contacts with primary and secondary care (due to symptomatic gallstones) were estimated from patient questionnaires at 3, 9, 12,18 and 24 months (post randomisation) and 6-monthly thereafter until the end of trial and checked at source. QALYs were estimated from patients' responses to the SF-36.

Where completed follow-up questionnaires identified that a participant had received an operation to remove their gallbladder or received a further treatment or surgery to treat their gallstone symptoms, this was followed up with the site RN to complete the appropriate CRFs.

A RN completed a CRF at 24 months for all participants to confirm if they had received any gallbladder-related surgery, and if so the type of procedure and the date it was received. If participants had been found to have had a surgery, the relevant CRFs were then completed.

Data processing

Research nurses at each of the participating centres entered both baseline and CRF data for their participants onto the study website through an online portal. Follow-up questionnaire data, which were sent from the participant directly to the trial office, were entered into the study website by trial office staff.

As part of the trial's monitoring plan, the trial office carried out data accuracy checks on a sample of baseline data entered by each site. No data accuracy checks were carried out on the data entered by trial office staff upon receipt of follow-up questionnaires.

Participant withdrawal

Participants remained in the trial unless they chose to withdraw consent or if they were unable to continue for a clinical reason. All changes in status (with the exception of complete withdrawal of consent) meant the participant was followed up for all trial outcomes wherever possible. All data collected up to the point of complete withdrawal were retained and used in the analysis.

Outcomes

Primary outcome

The *primary patient outcome measure* was QoL. This was measured by area under the curve (AUC) at up to 18 months post randomisation using the SF-36 bodily pain domain (AUC measures at 3, 9, 12 and 18 months).

Secondary outcomes

The secondary outcomes were measured at both 18 and 24 months post randomisation. These were:

- AUC up to 24 months post randomisation for SF-36 bodily pain
- CSQ
- SF-36 domains (excluding bodily pain domain)
- complications (defined as any presurgery, intraoperative or postoperative complications)
- need for further treatment (patient reported)
- persistent symptoms [patient reported and consisted of two sections (pain and dyspepsia) of the CSQ]
- HCRU
- costs.

Safety and breaches

Adverse events

An adverse event (AE) is defined as any untoward medical event affecting a clinical trial participant. Each AE will be considered for severity, causality and expectedness and may be reclassified as a serious event based on prevailing circumstances.

In the C-GALL trial, AEs were anticipated to occur during or after any type of surgery and while in observation/conservative management. In this trial, the following events were expected.

Adverse events during or after laparoscopic cholecystectomy

Intraoperative complications:

- bleeding > 500 millilitres (ml)
- injury to abdominal viscera, including liver tear or laceration
- anaesthetic complications (including hypersensitivity to the general anaesthesia and/or any of the medications or material used)
- injury to the bile duct
- bile leak from the bile duct, hepatic duct or ducts at the base of the liver or bile spillage from the gallbladder
- bile/stone spillage from the gallbladder.

Immediate postoperative complications:

- postoperative bleeding > 500 ml
- injury to the abdominal viscera, including liver tear or laceration
- injury to bile duct
- bowel obstruction
- wound infection
- pain requiring additional analgesia
- bile leak
- thrombosis (deep vein thrombosis/pulmonary embolism)
- urinary retention
- infection (sepsis, septicaemia, abscess)
- retained/missed common bile duct stone.

Late postoperative complications:

- incisional/port site hernia
- chronic wound pain

- infection (sepsis, septicaemia, abscess)
- biliary pain (right upper quadrant pain)
- non-specific abdominal pain
- post cholecystectomy jaundice.

Potential adverse events during observation/conservative management/presurgery

The anticipated risk of developing a potential AE in the conservative management group that might require further surgery or endoscopic treatment was 0.7%/year.[35] The following were expected:

- acute cholecystitis
- empyema/mucocele
- gallbladder perforation
- acute pancreatitis
- common bile duct stone
- obstructive jaundice
- gallstone ileus.

Adverse events that met the criteria for 'serious' were reviewed in order to determine whether or not the event was 'related'. Within C-GALL, 'related' was defined as an event that occurred as a result of a procedure required by the protocol, whether or not it was either (1) the specific intervention allocated at randomisation or (2) it was administered as an additional intervention as part of normal care.

All serious adverse events (SAEs) that were considered to be 'related' were recorded on a trial-specific SAE form. SAEs are described as complications or need for further treatment in *Chapter 3*. SAEs that were not related were not recorded. Deaths were also recorded on this SAE form.

A SAE is any AE that:

- results in death
- is life-threatening (i.e. the subject was at risk of death at the time of the event; it does not refer to an event which hypothetically might have caused death if it were more severe)
- requires hospitalisation or prolongation of existing hospitalisation
- results in persistent or significant disability or incapacity
- is a congenital anomaly or birth defect
- is otherwise considered medically significant by the investigator.

All AEs and SAEs that met the criteria for recording within C-GALL were recorded from the time a participant consented to join the trial until the end of the trial.

Breaches

Sites were asked to report potential breaches of trial protocol or GCP to the trial office. Trial office staff could also report potential breaches. There were two breaches recorded within the study and these are summarised in *Appendix 1*, *Table 36*. Both breaches were assessed by the Chief Investigator and the Sponsor as non-serious.

Ground rules for statistical analysis

The trial analysis followed a statistical analysis plan (SAP), which was agreed by the Project Management Group (PMG) and Trial Steering Committee (TSC) before analysis was started. Apart from the trial statistician, all other authors were blinded to the data when agreeing the SAP. The main analyses were based on the intention-to-treat (ITT) principle (i.e. analysed as randomised irrespective of non-compliance or crossover) and took place after the 24-month follow-up was completed for all

participants. Baseline and follow-up data were summarised using appropriate statistics and graphical summaries. Statistical significance was at the two-sided 5% level with corresponding 95% CIs derived. All analyses were carried out using Stata 16.[36]

Sample size

The primary outcome was AUC measured from SF-36 bodily pain up to 18 months. To detect a 0.33 standard deviation (SD) difference, with 90% power and alpha at 5%, 194 participants per group (388 in total) were required. A 0.33 difference in generic health status is considered clinically relevant and in terms of treatment effect size, in the small to medium ranges as observed in other clinical studies. We allowed for 10% of participants to have completely missing outcome data, with no AUC calculable, inflating the sample size to 430 participants in total.

Primary outcome analysis

The primary outcome, AUC SF-36 bodily pain up to 18 months, was analysed using a mixed effects regression model with adjustment for the minimisation covariates gender (male, female), age (< 35; 35–64; ≥ 65 years) and including centre as a random effect. The AUC for each participant was generated by the trapezium rule using baseline, 3-, 9-, 12- and 18-month time points. Missing SF-36 bodily pain baseline values were imputed using centre-specific baseline mean. Our primary analysis included all participants who had at least one time point up to 30 months post randomisation. For participants with missing data at 18 months, multiple imputation (MI) using Rubin's rule under a missing at random assumption was used to impute their SF-36 bodily pain score at 18 months only. Variables included in the imputation model were SF-36 bodily pain at follow-up time points up to 30 months as well as baseline characteristics. The number of imputations used was the proportion of missing data. If participants were missing SF-36 bodily pain at other time points (apart from 18 months), then the AUC was calculated using the time points available only. Secondary analysis of the primary outcome was performed for participants who had an 18-month score. A sensitivity analysis was performed including all participants who had at least one time point up to 18 months with MI being used for missing 18 months' data.

Secondary outcome analysis

Short Form-36 bodily pain AUC up to 24 months post randomisation was analysed in a similar way to the primary outcome analysis. CSQ, SF-36 (excluding bodily pain) and persistent symptoms were analysed using repeated-measures mixed-effects regression model correcting for baseline score, minimisation covariates gender (male, female), age (< 35; 35–64; ≥ 65 years) and time as a fixed effect. The repeated measures were assessed at 3, 9, 12, 18 and 24 months with treatment effects estimates from time-by-treatment interactions at each time point. This approach uses participant data from all time points and incorporates a random effect for centre and participant. Data missing at baseline were reported as such. For the analysis, missing baseline data were imputed using the centre-specific mean of that variable. Complications and need for further treatment were analysed using a Poisson model adjusting for minimisation covariates gender (male, female), age (< 35; 35–64; ≥ 65 years) and including a random effect for centre using robust error variance.[37] This allows results to efficiently calculate adjusted relative risk. Analysis was performed separately for data up to 18 months and data up to 24 months.

Subgroup analysis

Planned subgroup analyses explored the potential treatment effect moderation of gender (male, female), age (< 35; 35–64; ≥ 65 years) and ethnicity on the primary outcome. For ethnicity, we planned to use the UK census ethnic groupings; however, due to limited numbers in certain categories we categorised ethnicity as white or other than white. The subgroup-by-treatment interaction was assessed by including interaction terms in the models outlined above. We used a stricter level of significance (two-sided 1% significance level) and 99% CIs to reflect the exploratory nature of these analyses.

Compliance analysis

We explored the influence of compliance by undertaking a complier-adjusted causal estimation analysis. Compliance was defined as participants who received their allocated treatment within 24 months. For the laparoscopic cholecystectomy group, participants who received emergency cholecystectomy were defined as non-compliant. For the complier-adjusted causal estimation analysis, an instrumental variable two-stage least squares regression model with compliance instrumented by random allocation was used adjusting for baseline score, minimisation covariates gender (male, female), age (< 35; 35–64; ≥ 65 years) and adjusting for centre using cluster robust variance.

Coronavirus disease 2019

The impact of coronavirus disease 2019 (COVID-19) was assessed by looking at the AUC for the subset of data pre-COVID-19 defined as before 11 March 2020[38] using the same analysis as described in the primary outcome analysis.

Criteria for the termination of the trial

Due to the staggered nature of recruitment and, therefore, the measurement of the primary outcome at 18 months, there were no planned interim analyses for futility or benefit. We proposed one main effectiveness analysis at the end of the trial. During the trial, safety and other data were monitored by reports prepared for the Data Monitoring Committee (DMC).

Differences between the statistical analysis plan and the published protocol

In the published protocol,[1] we did not include ethnicity in the subgroup analysis. This was because when writing the SAP, we highlighted the need to look at the effect of different ethnic groups on the primary outcome. Also, we stated we would look at a per-protocol analysis; however, we decided to do a compliance analysis as it does not exclude participants who did not receive their allocated intervention.

Economic evaluation

Within this study both a 'within-trial' and a model-based economic evaluation were conducted. These are described in detail in *Chapters 4* and *5*.

Ethics approval and monitoring

C-GALL received favourable ethics opinion from North of Scotland Research Ethics Committee (REC) A, on 23 May 2016 (REC reference number 16/NS/0053).

Sponsorship

The University of Aberdeen and NHS Grampian co-sponsored the trial.

Management of the trial

The trial management team (consisting of the Trial Manager, Data Coordinator and the Co-Chief Investigators), based within CHaRT, University of Aberdeen, provided day-to-day support for the recruiting centres. Recruiting centres were led by a local Principal Investigator (PI). The PIs in most cases were supported by RNs, trial co-ordinators or dedicated staff, who were responsible for all aspects of the local organisation, including recruitment of participants, delivery of the interventions and notification of any problems or unexpected developments during the study period. The study was supervised by the PMG, which consisted of representatives from the study office and grant holders.

Oversight of the study

Project Management Group

The PMG was responsible for overseeing the management of the trial. This group consisted of the representatives from the study office, and grant holders which included a grant holder who was the patient and public involvement (PPI) representative and a senior programmer. Members of the PMG are listed in the *Acknowledgements*.

Trial Steering Committee

The TSC was responsible for monitoring and supervising the progress of the C-GALL study. The committee met nine times between August 2016 and September 2021 at agreed intervals. The TSC consisted of independent experts, patient representative, the Co-Chief Investigators and key members of the PMG. Members of the TSC are listed in the *Acknowledgements*.

One of the independent members of the TSC was a patient representative who contributed their individual perspectives of gallstone disease and the perspectives of gallstone disease of the wider community. Our grant holder PPI representative also provided us with perspectives of gallstone disease from the wider community. The TSC reviewed and commented on the study design, protocol and all study documentation, including patient-facing documents that were sent to potential and recruited participants in the C-GALL study. In addition, the PPI partners (grant holder PPI representative and TSC patient representative) also contributed to regular funder progress reports.

Patient and public involvement

The PPI partners were actively involved in discussions of the study results with the TSC and the trial investigators and contributed to the preparation of the plain language summary. They continue to be involved in developing dissemination materials for participants and contribute to academic papers. The PPI partner on the TSC will comment on the participant results letter. At the end of the study, the PPI partners reflected on their input and made suggestions for future research, which is included in the discussion.

Participants who took part in the focus group for the core outcome set (COS), described in *Chapter 7*, have remained involved with the C-GALL study as members of the C-GALL PPI group. The PPI group were actively involved in discussions of the study results with the PPI partners and contributed to the review of the Plain language summary. They continue to be involved in the review of dissemination materials for participants.

Data Monitoring Committee

The DMC was independent of the trial and was responsible for monitoring safety and data integrity. The DMC met nine times between October 2016 and November 2021 at agreed intervals. The trial statistician provided the data and analyses requested by the DMC prior to each meeting. The committee consisted of three independent experts. Members of the DMC are listed in the *Acknowledgements*.

Protocol amendments

There were 10 protocol amendments, and these are summarised in *Appendix 2*, *Table 37*. All were minor clarifications within the protocol. All were reviewed by the sponsor and the study funder before being submitted to, and then approved by, the REC.

Important changes to the methods after trial commencement

Extension to recruitment

Slower than anticipated recruitment and longer than anticipated waiting lists for surgery meant an 18-month recruitment extension was required to achieve full sample size and the opportunity for more participants to receive surgery. The 24-month time point, as an outcome, was added as part of this extension. Due to the staggered recruitment of participants, when the later randomised participants reached 24 months, earlier participants would have reached a longer follow-up time point. Therefore, it was decided to carry on collecting data, as it was deemed this would be useful in the analysis and any long-term follow-up plans.

Core outcome set

A COS for uncomplicated symptomatic gallstones was developed and is described in more detail in *Chapter 7*.

Study Within A Trial

The C-GALL study was involved in the Christmas Card Study Within A Trial (SWAT).[39] Full details of the methods and results can be accessed in the associated publication.[40]

In brief participants receiving postal questionnaires in eight host studies (including C-GALL) were randomised to receive a Christmas card or no Christmas card. The primary outcome of the SWAT was response to the next postal questionnaire that was due. The results of the SWAT showed that sending a Christmas card did not increase response rates compared to not sending a Christmas card.

The C-GALL study was also involved in the STICKER SWAT.[41] Participants receiving postal questionnaires in two studies (including C-GALL) are randomised to have a sticker with the trial logo included on the outgoing envelope or a plain envelope. This SWAT is ongoing and, as such, results are not yet available but will be reported in the future.

Chapter 3 Baseline, trial results and clinical effectiveness

Recruitment to the C-GALL trial

Participants were recruited to the C-GALL trial between August 2016 and November 2019. In total, 436 participants were randomised from 20 centres within the UK (*Table 2*), 218 to each group. The trajectory of recruitment from all centres is shown in *Figure 2*. Initial recruitment was limited to four pilot centres (Aberdeen Royal Infirmary, Nottingham City Hospital, Musgrove Park Hospital, Taunton and Royal Free Hospital, London). Roll-out to other centres began in April 2017. Recruitment was slower than originally projected and an 18-month extension to recruitment was requested and approved. A revised recruitment projection was initiated in March 2018 (see long dashed line in *Figure 2*).

TABLE 2 Recruitment by centre

Centre	Observation/ conservative management, *N* = 218	Laparoscopic cholecystectomy, *N* = 218	Overall, *N* = 436
Aberdeen Royal Infirmary	48 (22.0)	45 (20.6)	93 (21.3)
Nottingham City Hospital	24 (11.0)	24 (11.0)	48 (11.0)
Coventry University Hospital	23 (10.6)	21 (9.6)	44 (10.1)
Musgrove Park Hospital, Taunton	18 (8.3)	16 (7.3)	34 (7.8)
Birmingham Heartlands Hospital	15 (6.9)	15 (6.9)	30 (6.9)
North Tees University Hospital	13 (6.0)	15 (6.9)	28 (6.4)
Queen Elizabeth Hospital Birmingham	13 (6.0)	14 (6.4)	27 (6.2)
Queen Elizabeth University Hospital, Glasgow	11 (5.0)	12 (5.5)	23 (5.3)
University Hospital of Wales, Cardiff	10 (4.6)	11 (5.0)	21 (4.8)
Sandwell Medical Research Unit	9 (4.1)	9 (4.1)	18 (4.1)
University Hospital North Durham	9 (4.1)	6 (2.8)	15 (3.4)
Royal Free Hospital, London	4 (1.8)	10 (4.6)	14 (3.2)
Royal Gwent Hospital, Newport	5 (2.3)	7 (3.2)	12 (2.8)
Warwick Hospital	4 (1.8)	4 (1.8)	8 (1.8)
Yeovil District Hospital	4 (1.8)	4 (1.8)	8 (1.8)
Bedford Hospital NHS Trust	3 (1.4)	1 (0.5)	4 (0.9)
Victoria Hospital, Fife	1 (0.5)	2 (0.9)	3 (0.7)
Royal Liverpool University Hospital	2 (0.9)	1 (0.5)	3 (0.7)
Borders General Hospital	1 (0.5)	1 (0.5)	2 (0.5)
Plymouth Hospitals NHS Trust	1 (0.5)	0 (0.0)	1 (0.2)

Note
Values are *n* (%).

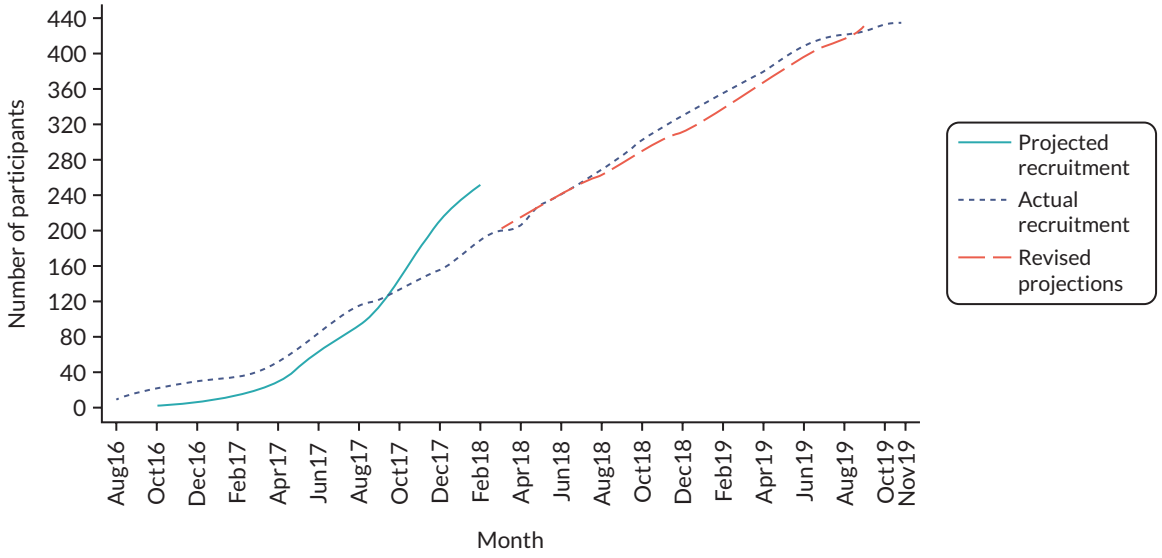

FIGURE 2 Recruitment over time.

Participant flow

The Consolidated Standards of Reporting Trials (CONSORT) diagram for the C-GALL trial is shown in *Figure 3*. There were 2667 patients identified to be potentially eligible for inclusion into the trial, of which 1298 were excluded. The main reasons for patients' exclusion were that they were ineligible (647/1298, 49.9%). Ineligibility was due to unconfirmed symptomatic gallstones (233/647, 36.0%), clinical diagnosis of symptomatic gallstone disease not confirmed by imaging (144/647, 22.3%) and being medically unfit for surgery (105/647, 16.2%). Of the 1369 eligible patients, 933 were not randomised with 910/933 (97.5%) having a preference. Main preference reasons were participants preferred laparoscopic cholecystectomy (538/910, 59.1%), observation/conservative management (167/910, 18.4%) and did not want to be randomised (91/910, 10.0%). Further details on reasons why patients were excluded and not randomised are shown in *Appendix 3*, *Table 38*.

Of the 436 participants randomised, one participant was a post-randomisation exclusion in the observation/conservative management group due to the participant immediately withdrawing from the study due to a previously unstated preference for laparoscopic cholecystectomy. One participant was randomised twice in the laparoscopic cholecystectomy group through error. At 24 months, in the observation/conservative management group, 136 (62.1%) participants responded to participant questionnaires (PQs) and 12 (5.5%) had declined further follow-up. In the laparoscopic cholecystectomy group, 138 (63.6%) participants responded to PQs and 11 (5.1%) had declined further follow-up and one participant had died. *Appendix 3*, *Table 39* shows the response rates for each of the follow-up time points.

Baseline characteristics

The baseline characteristics are shown in *Table 3*. Overall, the randomised groups were well balanced in baseline characteristics. The mean age of participants was approximately 50 years, over 78% were female and over 85% were white. In the observation/conservative management group, 60.4% had a normal gallbladder wall confirmed by transabdominal ultrasonography or another imaging technique compared with 55.3% in the laparoscopic cholecystectomy group. The mean SF-36 norm-based bodily pain score was 44.5 (SD 11.7) in the observation/conservative management group and 43.3 (SD 11.1) in the laparoscopic cholecystectomy group.

Non-responders to questionnaires tended to be younger (mean 46 vs. 52 years), have worse SF-36 bodily pain (mean 41.8 vs. 46.6) and worse disease-specific measures at baseline compared to responders.

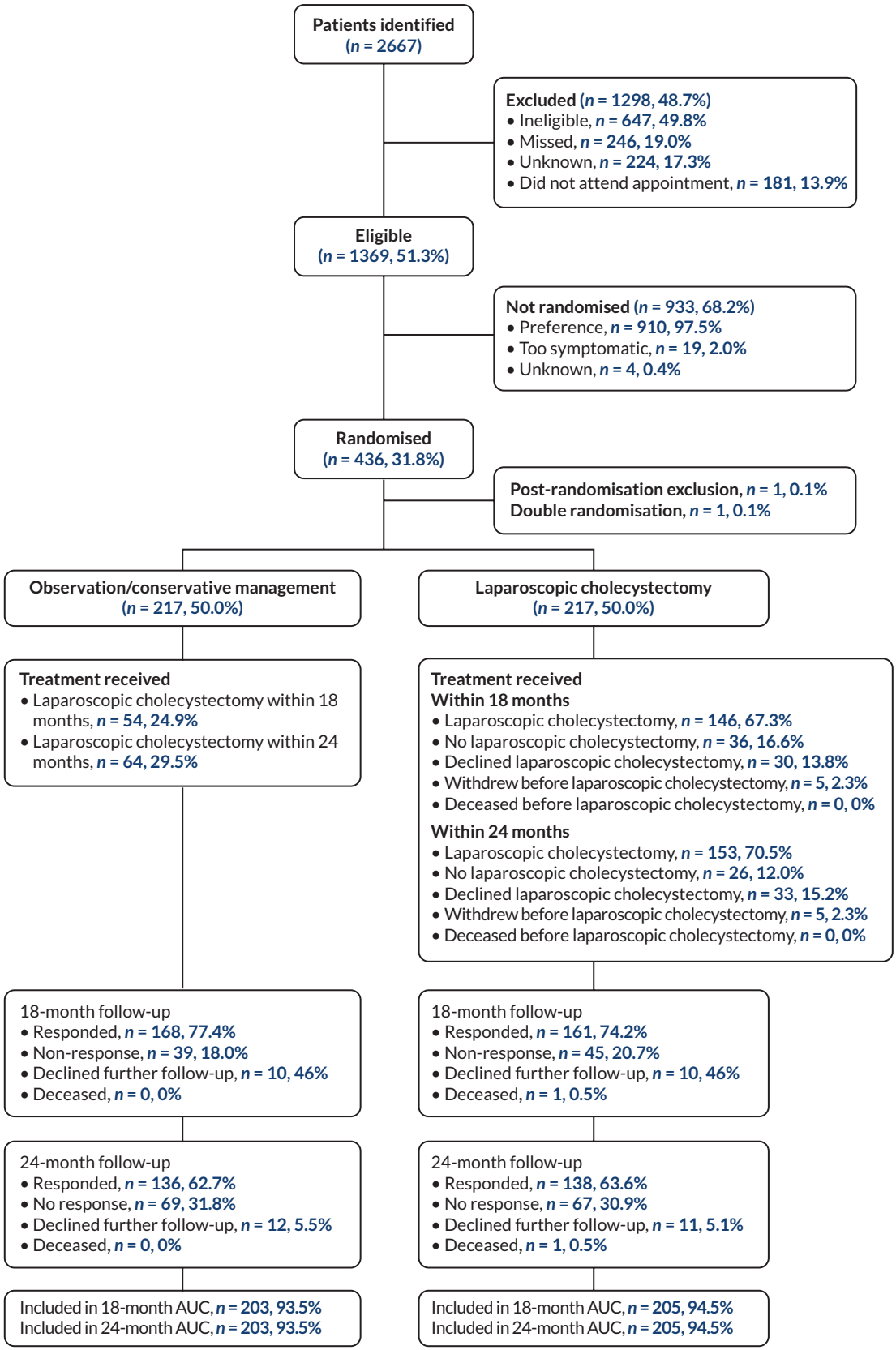

FIGURE 3 Participant flow (CONSORT) diagram.

TABLE 3 Baseline characteristics

Participant characteristics	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Age (years) – mean (SD); *n*	50.4 (15.1); 217	50.5 (15.3); 217
Sex – *n* (%)		
Male	46 (21.2)	47 (21.7)
Female	171 (78.8)	170 (78.3)
Ethnicity – *n* (%)		
White	185 (85.3)	188 (86.6)
Mixed/multiple ethnic groups	2 (0.9)	1 (0.5)
Asian/Asian British	15 (6.9)	15 (6.9)
Black/African/Caribbean/Black British	7 (3.2)	5 (2.3)
Arab	–	2 (0.9)
Other	7 (3.2)	6 (2.8)
Missing	1 (0.5)	–
BMI (kg/m^2) – mean (SD); *n*	32.0 (7.0); 215	31.5 (7.1); 217
Diagnosed with diabetes – *n* (%)		
No	200 (92.2)	203 (93.5)
Type 1	–	2 (0.9)
Type 2	17 (7.8)	12 (5.5)
Gallbladder wall[a] – *n* (%)		
Normal	131 (60.4)	120 (55.3)
Thick	27 (12.4)	30 (13.8)
Not recorded	59 (27.2)	67 (30.9)
Thickness of gallbladder wall if thick[a] (mm) – mean (SD); *n*	5.3 (2.1); 10	5.9 (3.4); 15
Hypertension – *n* (%)		
No	173 (79.7)	182 (83.9)
Yes	43 (19.8)	35 (16.1)
Missing	1 (0.5)	–
SF-36 norm-based scores – mean (SD); *n*		
Bodily pain	44.5 (11.7); 215	43.3 (11.1); 216
Physical functioning	48.2 (10.6); 214	47.3 (10.9); 216
Role physical	47.7 (10.3); 215	46.4 (11.4); 216
General health	45.0 (9.3); 213	43.3 (10.4); 216
Vitality	46.7 (10.0); 213	44.7 (10.9); 216
Social functioning	45.6 (11.7); 213	43.9 (12.5); 216
Role emotional	45.9 (12.4); 215	44.7 (13.3); 216

TABLE 3 Baseline characteristics (*continued*)

Participant characteristics	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Mental health	47.7 (10.4); 213	46.1 (11.1); 216
PCS	46.7 (9.3); 213	45.6 (9.7); 216
MCS	46.4 (11.5); 213	44.72 (12.1); 216
Otago gallstones CSQ – mean (SD); *n*	33.2 (19.9); 210	35.4 (20.6); 211
Persistent symptoms score[b] – mean (SD); *n*	43.0 (20.9); 213	44.6 (22.8); 215

a Confirmed by transabdominal ultrasonography or another imaging technique.
b Derived from two CSQ domains, pain and dyspepsia.

Note
For SF-36 norm-based scores, a higher score indicates better quality of life. For Otago gallstones CSQ, higher score indicating higher symptom burden and therefore poorer quality of life ranging from 0 to 100.

Treatment received

At 18 months, 54 (24.9%) in the observation/conservative management group and 146 (67.3%) in the laparoscopic cholecystectomy group received surgery. Further surgery details up to 18 months are provided in *Appendix 3*, *Table 40*. At 24 months, 64 (29.5%) in the observation/conservative management group, and 153 (70.5%) in the laparoscopic cholecystectomy group received surgery with a median time to surgery of 9.0 months (IQR 5.6–15.0) and 4.7 months (IQR 2.6–7.9), respectively. The majority of the surgical operations were elective in both groups, and over 95% were performed laparoscopically. In the observation/conservative management group, surgery was straightforward for 36/64 (56.3%) compared with 94 (61.4%) in the laparoscopic cholecystectomy group. In the observation/conservative management group, 54/64 (84.4%) did not have a normal gallbladder and of these 44 (81.5%) had chronic cholecystitis. In the laparoscopic cholecystectomy group, 143/153 (93.5%) did not have a normal gallbladder and of these 130 (90.9%) had chronic cholecystitis. Further surgery details at 24 months are shown in *Table 4*.

For those who had not had surgery by 24 months (*Table 5*), 15 (6.9%) were on a waiting list and 7 (3.2%) declined any further follow-up from hospital records in the observation/conservative management group; 13 (6.0%) and 5 (2.3%) in the laparoscopic cholecystectomy group.

Primary outcome

Figure 4 shows the mean of SF-36 norm-based bodily pain score. At 18 months, the mean SF-36 norm-based bodily pain score was 49.4 (SD 11.7) in the observation/conservative management group and 50.4 (SD 11.6) in the laparoscopic cholecystectomy group (*Table 6*). For the primary analysis, the mean AUC over 18 months was 46.8 for both groups with no difference: MD −0.0, 95% CI (−1.7 to 1.7); *p*-value 0.996.

Sensitivity and complete case analyses are shown in *Appendix 3*, *Table 41*.

Compliance analysis

Figure 5 shows the mean of SF-36 norm-based bodily pain score by compliance, where compliance was defined as participants receiving their allocated treatment within 24 months (excluding those two participants in the cholecystectomy group where surgery was done as an emergency case). For the

TABLE 4 Surgery details up to 24 months

Surgery details	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Received surgery		
Yes	64 (29.5)	153 (70.5)
No	153 (70.5)	64 (29.5)
	Received surgery, *N* = 64	**Received surgery, *N* = 153**
Time to surgery (months) – median (IQR); *n*	9.0 (5.6–15.0); 63	4.7 (2.6–7.9); 153
Time between surgery and 24 months follow-up (months) – median (IQR); *n*	15.0 (9.0–18.4); 63	19.3 (16.1–21.4); 153
Length of hospital stay (days) – median (IQR); *n*	1 (0–1); 61	0 (0–1); 150
Operation time (minutes) – median (IQR); *n*	65 (50.0–101.5); 56	61 (50.0–85.0); 139
Elective surgery		
Yes	56 (87.5)	149 (97.4)
No	6 (9.4)	2 (1.3)
Missing	2 (3.1)	2 (1.3)
Procedure type		
Laparoscopic	61 (95.3)	149 (97.4)
Open	1 (1.6)	1 (0.7)
Laparoscopic converted to open	1 (1.6)	1 (0.7)
Missing	1 (1.6)	2 (1.3)
Grade of operating surgeon		
Consultant	37 (57.8)	100 (65.4)
Consultant supervised by another consultant	7 (10.9)	7 (4.6)
Registrar	2 (3.1)	6 (3.9)
Registrar supervised by a consultant	7 (10.9)	19 (12.4)
Specialty (specialty and associate specialist grade) supervised by a consultant	2 (3.1)	–
Senior House Officer supervised by a consultant	1 (1.6)	3 (2.0)
Specialist trainee	–	2 (1.3)
Specialist trainee supervised by a consultant	1 (1.6)	4 (2.6)
Other	–	2 (1.3)
Other supervised by a consultant	2 (3.1)	6 (3.9)
Unknown operating surgeon but supervised by a consultant	–	1 (0.7)
Missing	5 (7.8)	3 (2.0)
Prophylactic antibiotic used in the operation		
Yes	35 (54.7)	81 (52.9)
No	23 (35.9)	64 (41.8)
Missing	6 (9.4)	8 (5.2)

TABLE 4 Surgery details up to 24 months (*continued*)

Surgery details	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Difficulty of surgery[a]		
Straightforward	36 (56.3)	94 (61.4)
Mildly difficult	5 (7.8)	12 (7.8)
Moderately difficult	4 (6.3)	16 (10.5)
Extremely difficult	3 (4.7)	1 (0.7)
Missing	16 (25.0)	30 (19.6)
Admitted to ICU or HDU		
No	58 (90.6)	145 (94.8)
ICU	–	2 (1.3)
HDU	1 (1.6)	–
Missing	5 (7.8)	6 (3.9)
Time in ICU (hours) – median (IQR); *n*	–	30 (24–36); 2
Time if HDU (hours) – value; *n*	47; 1	–
Required additional pain relief predischarge[b]	12 (18.8)	31 (20.3)
Histopathology		
Normal gallbladder		
Yes	6 (9.4)	7 (4.6)
No	54 (84.4)	143 (93.5)
Missing	4 (6.3)	3 (3.0)
Cholecystitis in abnormal gallbladder		
No	4 (7.4)	7 (4.9)
Acute	5 (9.3)	5 (3.5)
Chronic	44 (81.5)	130 (90.9)
Missing	1 (1.9)	1 (0.7)
Incidental biliary cancer		
No	59 (92.2)	149 (97.4)
Yes	–	–
Missing	5 (7.8)	4 (2.6)

a Completed by the operating surgeon.
b Additional over and above standard pain relief following surgery.

Note
Values are *n* (%) unless otherwise stated.

TABLE 5 Participant status at 24 months

Status	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Gallbladder removed		
Yes	64 (29.5)	153 (70.5)
No	153 (70.5)	64 (29.5)
If no, status of participant		
Participant is on surgical waiting list	15 (6.9)	13 (6.0)
Participant not on surgical waiting list	131 (60.4)	13 (6.0)
Withdrew before 24 months	7 (3.2)	5 (2.3)
Declined laparoscopic cholecystectomy	–	33 (15.2)

Note
Values are *n* (%).

compliance analysis of AUC over 18 months, there was no evidence of a difference: MD −0.0, 95% CI (−5.1 to 5.1); *p*-value 0.997 (see *Table 6*).

Subgroup analyses

Figure 6 shows the prespecified subgroup analyses for AUC SF-36 norm-based bodily pain score, gender, age (< 35, 35–64, ≥ 65 years) and ethnicity (white vs. other) at 18 months (primary outcome). Overall, there was no evidence that the treatment effect was moderated by any subgroups. *Appendix 3*, *Figures 18* and *19* show the subgroup analyses for the sensitivity and complete case analyses.

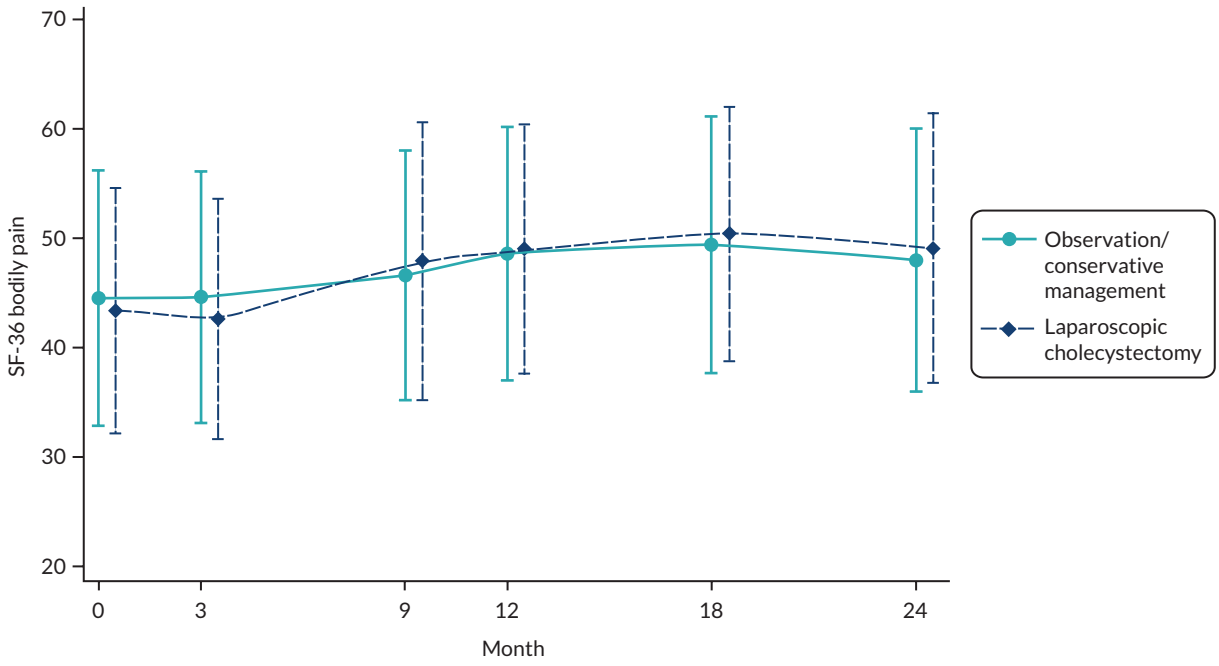

FIGURE 4 Mean SF-36 norm-based bodily pain score up to 24 months. Higher score indicates better quality of life. Bars represent ± SD.

TABLE 6 Primary outcome – AUC SF-36 norm-based bodily pain score over 18 months

Primary outcome	Observation/ conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217	Observation/conservative management, *N* = 217		Laparoscopic cholecystectomy, *N* = 217	
	ITT	ITT	Complied[a]	Not-complied[b]	Complied[c]	Not-complied[d]
Primary analysis						
Summary measures based on available data						
Baseline	44.5 (11.7); 202	43.4 (11.2); 205	46.3 (11.1); 142	40.3 (11.8); 60	43.1 (11.1); 147	44.1 (11.7); 58
3 months	44.6 (11.5); 176	42.6 (11.0); 174	46.4 (10.8); 124	40.1 (12.0); 52	42.0 (11.0); 126	44.2 (10.7); 48
9 months	46.6 (11.4); 144	47.9 (12.7); 160	46.0 (10.8); 103	48.1 (12.9); 41	48.7 (12.5); 119	45.6 (13.1); 41
12 months	48.6 (11.6); 156	49.0 (11.4); 149	47.8 (11.0); 118	50.8 (13.3); 38	50.0 (10.6); 112	45.9 (13.1); 37
18 months	49.4 (11.7); 167	50.4 (11.6); 161	48.9 (11.6); 121	50.6 (11.8); 46	51.5 (11.1); 119	47.3 (12.3); 42
24 months	48.0 (12.0); 135	49.1 (12.3); 138	46.8 (11.9); 102	51.9 (11.7); 33	49.9 (11.5); 100	47.1 (14.1); 38
AUC over 18 months[e]	46.8 (8.8); 203	46.8 (8.7); 205	47.2 (8.6); 143	45.8 (9.0); 60	47.2 (8.2); 147	45.6 (9.8); 58
MD, 95% CI; *p*-value	−0.0	(−1.7 to 1.7); 0.996			−0.0	(−5.1 to 5.1); 0.997

a Received observation/conservative management.
b Received laparoscopic cholecystectomy.
c Received laparoscopic cholecystectomy.
d Received observation/conservative management.
e If 18 months data were missing, then MI was used to calculate an SF-36 bodily pain score at this time point.

Note
Values are mean (SD); *n* unless otherwise stated.
Compliance was defined as participants who received their allocated treatment within 24 months (apart from the laparoscopic cholecystectomy group if surgery was an emergency).

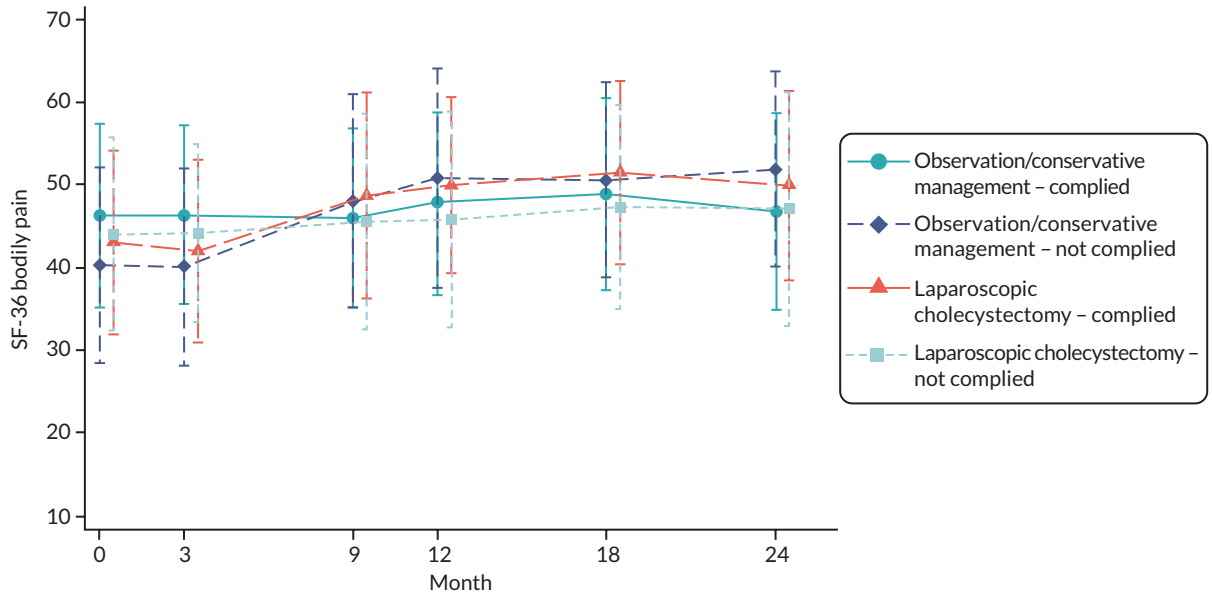

FIGURE 5 Mean SF-36 norm-based bodily pain score up to 24 months by compliance. Higher score indicates better quality of life. Compliance was defined as participants who received their allocated treatment within 24 months (apart from the laparoscopic cholecystectomy group if surgery was an emergency).

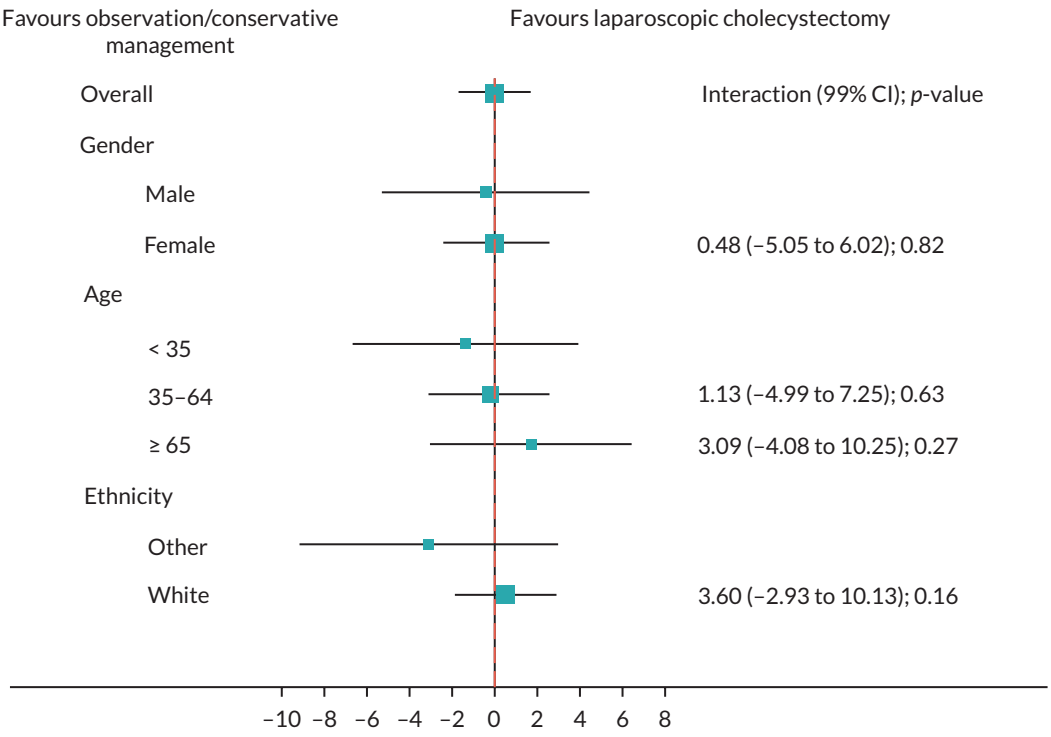

FIGURE 6 Subgroups for observation/conservative management vs. laparoscopic cholecystectomy up to 18 months for AUC SF-36 bodily pain primary analysis. Boxes represent differences in AUC and lines are confidence intervals.

Secondary outcomes

Patient-reported quality of life

The mean AUC SF-36 norm-based bodily pain score over 24 months was 47.2 (SD 8.6) in the observation/conservative management group and 47.3 (SD 8.7) in the laparoscopic cholecystectomy group and with no evidence of a difference for the primary analysis: MD −0.1, 95% CI (−1.8 to 1.7); *p*-value 0.948. Sensitivity and complete case analyses are shown in *Appendix 3*, *Table 42*.

For the SF-36 norm-based scores (apart from bodily pain), some small differences were observed as statistically significant at 18 months, but these disappeared at 24 months (*Table 7*). None of the observed effect sizes were clinically important.

The mean CSQ at 18 months was 21.3 (SD 21.0) for the observation/conservative management group and 15.8 (19.7) for the laparoscopic cholecystectomy group and showed evidence of a difference in favour of laparoscopic cholecystectomy: MD 6.6, 95% CI (1.9 to 11.3); *p*-value 0.006. At 24 months, again there was evidence of a difference in favour of laparoscopic cholecystectomy: MD 9.0, 95% CI (4.1 to 14.0), $p < 0.001$. There was a similar pattern for the persistent symptoms score.

Complications

At 18 months, 32 (10.1%) participants in the observation/conservative management group had had a complication compared with 44 (25.3%) participants in the laparoscopic cholecystectomy group, with no evidence of a difference between groups: RR 0.72, 95% CI (0.46 to 1.14); *p*-value 0.17 (further details are shown in *Appendix 3*, *Table 43*). At 24 months, there were two additional complications in the laparoscopic cholecystectomy group with no evidence of a difference between groups: RR 0.69, 95% CI (0.44 to 1.09); *p*-value 0.11. Details of the complications are shown in *Table 8*. In regard to complications

TABLE 7 Secondary outcome – QoL

Patient-reported secondary outcome	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217	MD	95% CI; *p*-value[a]
SF-36 norm-based score				
Physical functioning				
Baseline	48.2 (10.6); 214	47.3 (10.9); 216		
3 months	47.3 (11.0); 161	46.6 (11.1); 155	0.5	(−1.3 to 2.3)
9 months	46.5 (11.6); 131	47.4 (11.6); 143	−1.4	(−3.3 to 0.4)
12 months	47.7 (10.6); 127	48.2 (11.1); 125	−0.4	(−2.3 to 1.5)
18 months	47.8 (10.4); 120	49.6 (10.0); 114	−2.1	(−4.0 to −0.1); 0.04
24 months	47.7 (10.4); 103	49.1 (10.9); 99	−1.8	(−3.9 to 0.2); 0.08
Role physical				
Baseline	47.7 (10.3); 215	46.4 (11.4); 216		
3 months	46.8 (10.9); 161	44.8 (11.5); 152	2.1	(0.07 to 4.14)
9 months	46.1 (11.5); 130	46.1 (12.5); 138	−0.5	(−2.63 to 1.70)
12 months	46.7 (11.8); 126	48.8 (10.9); 121	−1.6	(−3.80 to 0.67)
18 months	47.7 (10.7); 121	48.5 (11.0); 114	−1.1	(−3.3 to 1.2); 0.36
24 months	46.5 (11.0); 103	47.8 (11.9); 99	−2.0	(−4.5 to 0.4); 0.10
General health				
Baseline	45.0 (9.3); 213	43.3 (10.4); 216		
3 months	43.1 (10.8); 160	42.3 (10.3); 155	0.0	(−1.7 to 1.8)
9 months	43.8 (10.7); 130	44.6 (10.8); 139	−1.5	(−3.3 to 0.4)
12 months	44.5 (10.6); 127	45.2 (10.6); 125	−1.7	(−3.5 to 0.2)
18 months	44.3 (10.9); 121	46.6 (11.0); 111	−2.0	(−3.9 to −0.1); 0.04
24 months	44.9 (10.5); 103	44.8 (11.3); 99	−0.9	(−2.9 to 1.1); 0.40
Vitality				
Baseline	46.7 (10.0); 213	44.7 (10.9); 216		
3 months	44.8 (10.8); 160	44.2 (10.5); 156	−0.6	(−2.5 to 1.3)
9 months	44.9 (10.9); 130	46.3 (11.3); 139	−2.9	(−5.0 to −0.9)
12 months	45.9 (11.5); 127	46.7 (11.2); 125	−2.1	(−4.2 to 0.0)
18 months	45.6 (11.2); 121	48.8 (11.4); 113	−3.9	(−6.0 to −1.7); < 0.001
24 months	46.3 (11.2); 102	47.0 (11.3); 99	−2.2	(−4.5 to 0.0); 0.06
Social functioning				
Baseline	45.6 (11.7); 213	43.9 (12.5); 216		
3 months	43.9 (11.9); 161	42.5 (12.4); 155	0.8	(−1.4 to 3.1)
9 months	44.1 (11.6); 129	46.1 (12.4); 140	−2.4	(−4.8 to 0.0)
12 months	44.5 (13.0); 126	46.2 (12.2); 124	−1.5	(−4.0 to 1.0)

continued

TABLE 7 Secondary outcome – QoL (*continued*)

Patient-reported secondary outcome	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217	MD	95% CI; *p*-value[a]
18 months	46.2 (11.3); 119	47.8 (12.0); 111	−1.2	(−3.8 to 1.3); 0.34
24 months	45.9 (12.5); 101	45.0 (12.5); 97	0.6	(−2.1 to 3.3); 0.68
Role emotional				
Baseline	45.9 (12.4); 215	44.7 (13.3); 216		
3 months	44.5 (12.6); 160	42.7 (12.8); 152	1.8	(−0.5 to 4.1)
9 months	44.2 (12.7); 129	44.6 (13.6); 138	−1.5	(−4.0 to 1.0)
12 months	44.7 (12.8); 124	46.3 (12.4); 121	−1.7	(−4.2 to 0.9)
18 months	44.1 (12.7); 121	46.4 (12.5); 114	−3.2	(−5.8 to −0.6); 0.015
24 months	45.6 (12.0); 103	44.8 (12.9); 99	−0.5	(−3.3 to 2.2); 0.71
Mental health				
Baseline	47.7 (10.4); 213	46.1 (11.1); 216		
3 months	44.9 (11.4); 160	45.1 (11.4); 156	−0.7	(−2.8 to 1.3)
9 months	45.2 (11.4); 130	47.2 (11.2); 139	−2.8	(−4.9 to −0.6)
12 months	46.4 (12.4); 127	47.1 (11.6); 125	−1.6	(−3.8 to 0.6)
18 months	45.9 (11.0); 121	48.1 (11.0); 112	−2.4	(−4.6 to −0.1); 0.040
24 months	46.6 (11.5); 103	45.3 (11.4); 99	−0.0	(-2.4, 2.3); 0.98
Physical component summary – PCS				
Baseline	46.7 (9.3); 213	45.6 (9.7); 216		
3 months	46.3 (10.1); 157	44.8 (10.2); 150	1.4	(−0.4 to 3.2)
9 months	46.4 (10.4); 127	46.9 (11.7); 136	−0.8	(−2.7 to 1.1)
12 months	47.4 (10.8); 123	48.4 (10.5); 119	−0.8	(−2.7 to 1.2)
18 months	47.8 (10.3); 117	49.3 (10.2); 108	−1.2	(−3.2 to 0.8); 0.24
24 months	47.2 (10.9); 100	48.7 (11.4); 97	−1.9	(−4.0 to 0.1); 0.07
Mental component summary – MCS				
Baseline	46.4 (11.5); 213	44.7 (12.1); 216		
3 months	43.9 (12.3); 157	43.2 (12.4); 150	−0.1	(−2.3 to 2.1)
9 months	44.2 (12.4); 127	46.0 (11.7); 136	−2.8	(−5.1 to −0.5)
12 months	45.1 (13.3); 123	46.1 (11.9); 119	−1.8	(−4.2 to 0.6)
18 months	44.9 (12.6); 117	47.8 (11.7); 108	−2.9	(−5.4 to −0.5); 0.020
24 months	45.9 (12.1); 100	44.4 (12.6); 97	0.2	(−2.3 to 2.8); 0.85
CSQ total				
Baseline	33.2 (19.9); 210	35.4 (20.6); 211		
3 months	29.5 (23.4); 148	30.9 (22.6); 147	−0.8	(−4.9 to 3.4)
9 months	26.6 (22.5); 122	23.6 (22.4); 132	4.4	(−0.0 to 8.8)
12 months	21.7 (22.1); 119	18.4 (19.4); 120	4.7	(0.2 to 9.2)

TABLE 7 Secondary outcome – QoL (*continued*)

Patient-reported secondary outcome	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217	MD	95% CI; *p*-value[a]
18 months	21.3 (21.0); 113	15.8 (19.7); 101	6.6	(1.9 to 11.3); 0.006
24 months	20.7 (20.1); 98	14.0 (17.0); 91	9.0	(4.1 to 14.0); < 0.001
Persistent symptoms score[b]				
Baseline	43.0 (20.9); 213	44.6 (22.8); 215		
3 months	32.9 (26.6); 156	34.6 (24.7); 153	−1.4	(−6.4 to 3.5)
9 months	31.1 (26.5); 128	26.7 (26.1); 139	5.5	(0.3 to 10.8)
12 months	23.4 (24.2); 125	20.2 (23.1); 121	4.6	(−0.9 to 10.0)
18 months	23.1 (24.1); 117	17.4 (22.2); 106	6.7	(1.0 to 12.3); 0.02
24 months	23.1 (23.0); 101	15.1 (18.4); 95	10.1	(4.2 to 16.0); 0.001

a *p*-values only reported for 18 and 24 months.
b Derived from two CSQ domains, pain and dyspepsia.

Note
Values are mean (SD), *n*; For SF-36 norm-based scores, a higher score indicates better quality of life; For Otago gallstones CSQ, higher score indicating higher symptom burden and therefore poorer quality of life.

TABLE 8 Secondary outcome – complications up to 24 months

Clinical secondary outcome	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Number of participants	32 (14.7)	46 (21.2)
RR (95% CI); *p*-value	0.69	95% CI (0.44 to 1.09); *p*-value 0.11
Number of complications		
1	18	32
2	8	5
3	4	8
4	2	1
Presurgery complications		
Number of participants	25 (11.5)	11 (5.1)
Number of complications		
1	20	9
2	4	
3	1	1
4		1
Details of presurgery complications		
Cholecystitis	14	8
Biliary colic	8	2

continued

TABLE 8 Secondary outcome – complications up to 24 months (*continued*)

Clinical secondary outcome	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Pancreatitis	2	3
Choledocholithiasis	2	–
Cholecystitis and jaundice	1	–
Choledocholithiasis and pancreatitis	1	–
Cholecystitis, choledocholithiasis and jaundice	–	1
Cholecystitis and pancreatitis	1	–
Bouveret syndrome[a]	1	–
Cholecystitis, choledocholithiasis and pancreatitis	–	1
Jaundice	–	1
Right upper quadrant pain	1	–
Intraoperative complications		
Number of participants	9 (4.1)	24 (11.1)
Number of complications		
1	8	23
2	1	1
Details of intraoperative complications		
Bile/stone spillage from gallbladder	6	16
Injury to abdominal viscera (including liver tear or laceration)	1	5
Bleeding > 500 ml	1	2
Bile leak from the bile duct, hepatic duct or ducts at base of liver	1	1
Injury to bile duct	1	–
Ruptured empyema	–	1
Postoperative complications		
Number of participants	7 (3.2)	16 (7.4)
Number of complications		
1	5	10
2	1	4
3	1	3
Details of postoperative complications		
Bleeding > 500 ml	1	2
Bile leak that required no treatment	2	3
Bowel obstruction requiring no treatment	1	4

TABLE 8 Secondary outcome – complications up to 24 months (*continued*)

Clinical secondary outcome	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Bowel obstruction requiring surgery	–	1
Wound infection	2	2
Intraperitoneal – collection/abscess requiring no treatment	1	4
Intraperitoneal – collection/abscess requiring precutaneous drainage	1	–
Vomiting	–	3
Dizziness and hypotension	1	–
Haematoma	–	1
Missed stone in the bile duct	–	1
Renal failure	–	1
Residual gallbladder inflamed	1	–
Wound dehiscence	–	1
Postsurgery complications within 30 days of discharge		
Number of participants	2 (0.9)	3 (1.4)
Number of complications		
1	2	3
Details of postsurgery complications within 30 days of discharge		
Cholangitis	–	1
Surgical site infection	1	1
Bile leak	–	1
Postcholecystectomy syndrome[b]	1	–
Postsurgery complications after 30 days of discharge		
Number of participants	1 (0.5)	1 (0.5)
Details of postsurgery complications after 30 days of discharge		
Right upper quadrant pain	–	1
Incisional hernia	1	–
Death – cardiovascular event		
Number of participants	–	1 (0.5)

a Bouveret syndrome occurs when a gallstone enters the small bowel via a bilioenteric fistula and is impacted in the duodenum or stomach, causing gastric outlet obstruction.
b Persistence of same symptoms reported by the patient post surgery.

Note
Values are *n* (%) or *n*.

that occurred before laparoscopic cholecystectomy (including in participants who did not have a laparoscopic cholecystectomy), they occurred in 25 (11.5%) participants in the observation/conservative management group and 11 (5.1%) in the laparoscopic cholecystectomy group, with the majority of the complications being either cholecystitis or biliary colic. For intraoperative complications, 9 participants (4.1%) had a complication in the observation/conservative management group and 24 (11.1%) in the laparoscopic cholecystectomy group with the majority being bile/stone spillage from the gallbladder. For postoperative complications, there were 7 (4.1%) and 24 (11.1%) participants, respectively, with bile/stone spillage from gallbladder as the main complication. For postsurgery complications, there were 2 (0.9%) and 3 participants (1.4%), respectively, within 30 days of discharge, and 1 participant (0.5%) after 30 days in both groups. There was one cardiovascular-related death in the laparoscopic cholecystectomy group (myocardial infarction). By treatment received at 18 months, there were 8/234 participants (3.4%) who had a complication in the observation/conservative management group and 68/200 (34.0%) in the laparoscopic cholecystectomy group. At 24 months, there were 7/217 (3.2%) and 71/217 (32.7%) participants who had a complication, respectively (see *Appendix 3*, *Tables 44* and *45* for further details).

Need for further treatment

At 18 months, there were 9/202 (4.5%) participants in the observation/conservative management group who had had a further treatment and 12/203 (5.9%) in the laparoscopic cholecystectomy group with no evidence of a difference: RR 0.75, 95% CI (0.31 to 1.78); *p*-value 0.509 (further details are shown in *Appendix 3*, *Table 46*). At 24 months, there were 10/202 (5.0%) participants in the observation/conservative management group who had further treatment and 16/203 (7.9%) in the laparoscopic cholecystectomy group with no evidence of a difference: RR 0.62, 95% CI (0.28 to 1.38); *p*-value 0.242 (*Table 9*). The main treatments were additional pain relief, antibiotics and endoscopic retrograde cholangiopancreatography (ERCP).

TABLE 9 Secondary outcome – further treatment up to 24 months

Clinical secondary outcome	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Number of participants	10/201 (5.0)	16/203 (7.9)
RR (95% CI); *p*-value	0.62	95% CI (0.28 to 1.38); *p*-value 0.242
Number of treatments		
1	8	10
2	2	4
3	–	1
4	–	1
Details of the further treatment		
Pain relief	3	12
Antibiotics	2	5
ERCP	3	4
Antisickness	1	–
Bloating	1	–
Urinary catheter for retention	–	2
Bowel	–	1

TABLE 9 Secondary outcome – further treatment up to 24 months (*continued*)

Clinical secondary outcome	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Colostomy	1	–
Blood transfusion	–	1
Laparotomy washout and haemostasis	–	1
Fluids	–	1
Pancreatitis management	–	1
Unknown	1	–

Note
Values are *n* (%) or *n*.

Appointments with health professionals, medications prescribed and further investigations

Details of appointments with health professionals, medications prescribed and further investigations related to their gallstones by 18 months are shown in *Appendix 3*, *Table 47*. By 24 months (*Table 10*), 108/202 (53.5%) of participants in the observation/conservative management group and 127/203 (62.6%) in the laparoscopic cholecystectomy group required appointments with healthcare professionals with the majority visiting their GP, or an NHS hospital A&E or referred to outpatient department. Further medication for their gallstones was prescribed to 67/202 (33.2%) participants in the observation/conservative management group and 59/203 (29.1%), respectively, and further investigation was required in 11/202 (5.4%) participants in the observation/conservative management group and 10/203 (4.9%) in the laparoscopic cholecystectomy group.

TABLE 10 Professional appointments, medication prescribed and further investigation up to 24 months

Details	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Professional appointments		
Number of participants who required appointments	108/202 (53.5)	127/203 (62.6)
Details of the appointments		
GP	73	91
District nurse	3	8
Practice nurse	15	34
NHS hospital outpatients	54	64
NHS hospital A&E	39	33
Private healthcare	2	3
Appointment with other care provider	8	1
Consultant	5	–
Physiotherapist	2	–
Paramedics	1	–

continued

TABLE 10 Professional appointments, medication prescribed and further investigation up to 24 months (*continued*)

Details	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Medications prescribed		
Number of participants who were prescribed and taken medication	67/202 (33.2)	59/203 (29.1)
Details of medication[a]		
Pain relief	100	69
Paracetamol	36	23
Dihydrocodeine	33	24
Ibuprofen	21	18
Tramadol	7	3
Other pain medication	3	1
Antispasmodic	25	20
Antisickness	–	1
Antibiotic	–	3
Antireflux	20	22
Antidiarrhoea	–	1
Other	3	7
Bile salts	1	5
Iron	2	2
Further investigation		
Number of participants	11/202 (5.4)	10/203 (4.9)
Number of investigations		
1	9	9
2	2	1
Details		
MRI	4	2
Ultrasound scan	3	5
CT	1	1
Endoscopy	2	1
SeHCAT test	1	2
Unknown	2	–

a Participants can take more than one medication.

Note
SeHCAT 23-seleno-25-homotaurocholic acid, selenium homocholic acid taurine or tauroselcholic acid test; values are *n* (%) or *n*; denominator is for those that responded to a follow-up questionnaire.

Impact of coronavirus disease 2019

All randomised participants reached the 3-month follow-up time point before the COVID-19 pandemic began (defined as 11 March 2020 by the WHO).[38] In addition, 97/167 (58.1%) in the observation/conservative management group and 102/161 (63.4%) in the laparoscopic cholecystectomy group had completed their 18 months of follow-up before this date (see *Appendix 3*, *Table 48*). The AUC difference for the SF-36 norm-based bodily pain score by 18 months was MD −0.1, 95% CI (−1.8 to 1.6); *p*-value 0.94 in the subset of data pre-COVID-19. There was no evidence that this differed from the intention-to-treat (ITT) estimate.

In summary, the C-GALL group delivered the trial in the UK NHS setting, across 20 secondary care sites, and randomised 434 participants. We found no evidence of differences in bodily pain (primary outcome), QoL, complications or in need for further treatment between the policies at up to 24 months follow-up. There was statistically significant evidence that gallbladder-specific measures of quality of life improved in the cholecystectomy randomised group at 24 months.

Chapter 4 Economic evaluation: within-trial analysis

Introduction

This chapter reports on the within-trial economic evaluation of laparoscopic cholecystectomy compared with observation/conservative management for the treatment of adults presenting uncomplicated symptomatic gallstones in a secondary care setting conducted during the C-GALL trial.[1] The analysis considers a 24-month time horizon. A longer time horizon is explored with a Markov model extrapolation allowing for further relevant costs and consequences associated with gallstones, laparoscopic cholecystectomy and observation/conservative management in *Chapter 5*. Health service resources are scarce and healthcare technologies should be adopted or maintained as part of the NHS healthcare package only if they can be demonstrated to provide good value for money. Economic evaluation can aid decision-making on the adoption, or not, of new healthcare technologies by providing information on the relative efficiency of the technologies under consideration.

Objectives of the economic evaluation

The primary economic objective of the within-trial cost-effectiveness analysis was to estimate the difference in cost and QALYs between observation/conservative management and laparoscopic cholecystectomy at out to 24 months post randomisation. The secondary economic objective addressed in *Chapter 5* is to model the longer-term cost-effectiveness of observation/conservative management versus laparoscopic cholecystectomy.

Methods

This economic evaluation followed established methods[42,43] and reporting standards[44] with a prespecified protocol and health economics analysis plan. The UK NHS healthcare system perspective was adopted for all the economic analyses (see *Chapters 4* and *5*).

Study design and participants

The C-GALL trial protocol[1] and *Chapter 2* provide details of the trial design. The within-trial economic analysis follows the ITT principle and is based on the same participants randomised and analysed for the main trial analysis.

Cost and outcome assessment

Case report forms and PQs were used to assess costs and outcomes. A RN completed a CRF at the time of surgery providing details of the operative procedures, perioperative complications and resource use in hospital [e.g. transfer to high dependency unit (HDU) or intensive care unit (ICU)]. Costs of the initial intervention procedures were estimated from resource use data recorded on the CRFs coupled with routine unit cost data.[31] Costs associated with subsequent contacts with primary and secondary care (due to symptomatic gallstones or post surgery) were estimated from PQs at 3, 9, 12, 18 and 24 months post randomisation. Health-related QoL weights were estimated from participants' responses to the SF-36 questionnaires at 3, 9, 12, 18 and 24 months post randomisation, and were used to estimate QALYs.[45,46]

Assessment of health service costs

The aim of the economic evaluation is to inform the efficient allocation of the NHS budget. Therefore, an NHS perspective was adopted for the analysis and NHS costs were estimated for health service use by the trial participants in both secondary and primary care. Costs are expressed in Great British pounds and reported using the 2019–20 price year.

Cost of the primary interventions

Cholecystectomy

The cost of the initial surgical intervention was estimated from resource use data recorded in the C-GALL Surgery CRF for each participant. The CRF recorded the type of procedure carried out (i.e. laparoscopic cholecystectomy, open cholecystectomy, laparoscopic converted to open cholecystectomy), operation time (i.e. incision to skin closure), grade of surgeon (and if the main surgeon was supervised by a consultant), intraoperative complications and anaesthetic complications. Further information such as dates for admission and discharge, and destination post operation (i.e. type of ward, HDU or ICU) was provided by the C-GALL postsurgery CRF.

In order to capture patient variation, the primary costing approach assigned costs to individual components of resource use (micro-costing). Laparoscopic cholecystectomy is one of the most common NHS operational procedures with more than 66,000 procedures completed in England in 2020–1.[47] The C-GALL Surgery CRF asked only for time from *incision to skin closure*. Upon discussion within the PMG and clinical opinion, time in the anaesthetic room and recovery room were estimated to be similar to those observed in the Hysterectomy or Endometrial AbLation Trial for Heavy menstrual bleeding (HEALTH)[48] which compared laparoscopic supracervical hysterectomy and second-generation endometrial ablation for the treatment of heavy menstrual bleeding. Time in the anaesthetic room was costed using the cost per hour (incorporating overheads) for a consultant anaesthetist and an anaesthetic nurse (*Table 11*). The unit cost of the recorded grade of surgeon and a consultant anaesthetist was applied to the time in theatre (*incision to skin closure* time). Nursing staff were costed as the requirement for general day surgery: one anaesthetic nurse, a scrub nurse and two further theatre nurses.

In addition, a published unit cost was applied for time in theatre to reflect the average cost of other staff, supplies and consumables, and allocated capital charges and overheads.[49] While this detailed unit cost of theatre time is only available for Scottish hospitals, the average cost per theatre hour in general hospitals in Scotland (£1072 including medical and nursing staff) is comparable with a published estimate for England (£1200 per hour).[50] Therefore, the Scottish estimate for the average cost for time in theatre for gastroenterology (£550 per hour, excluding medical and nursing staff) was applied to the time in theatre for the base-case analysis (see *Table 11*).

The unit cost of a band 6 nurse (inclusive of overheads) assuming one-to-one care was used to cost the time in recovery after surgery. Time on the ward following recovery was costed using an estimate of the cost per excess bed-day following cholecystectomy. Cost for excess bed-days were last published by NHS England in 2018.[51] Therefore, these unit costs were applied after being adjusted for inflation using the NHS cost inflation index.[52]

The National Schedule of NHS costs provides an alternative source for costing the initial cholecystectomy procedure.[31] For this, each patient record was mapped to the appropriate Healthcare Resource Group (HRG). The core HRG code for cholecystectomy is GA10K (laparoscopic cholecystectomy, 19 years and over, with complexity and comorbidity score 0). The C-GALL Surgery CRF contained information about the cholecystectomy being conducted as an elective or emergency procedure. Thus, the NHS cost for cholecystectomy was applied as either a day-case (patient discharged same day) or an elective inpatient admission (stay ≥ 1 day), non-elective short stay (up to 1 overnight stay), or non-elective long stay (more than 1 overnight stay).

TABLE 11 Unit costs (NHS perspective)

Resource	How measured	Source of measurement	Unit cost	Source of valuation
Time in anaesthetic room	Time in minutes	Assumption based on clinical opinion and HEALTH trial	£169 per hour	Band 6 nurse (£50) + consultant anaesthetist (£119); Unit Costs of Health and Social Care, 2020[48,54]
Time in theatre	Time in hours	C-GALL Surgery CRF		
Surgeon time			Consultant (£114 per hour); associate specialist (£117); registrar (£50 per hour); nurse consultant (£69); nurse (floor) (£60)	Unit Costs of Health and Social Care, 2020[52]
Anaesthetist time			Consultant (£119 per hour); anaesthetic nurse (£50)	Unit Costs of Health and Social Care, 2020[52]
Theatre costs (excluding medical and nursing staff)			£550 per hour	Table R140, ISD 2019[49]
Procedure consumables			£247.33	*Source*: Aberdeen Royal Infirmary General Surgery Service (Mr Jamie McAllister, personal communication)
Perioperative complication costs	See methods	C-GALL postsurgery CRF	Various based on recorded reasons and procedures	NHS Reference costs 2019–20[31]
Re-admissions	See methods	C-GALL hospital admission CRFs; C-GALL PQ	Various based on recorded reasons and procedures	NHS Reference costs 2019–20[31]
Outpatient appointments		C-GALL PQs		
Non-admitted face-to-face attendance, first			£145 per attendance	Service Code 301 Gastroenterology. NHS Reference costs 2019–20[31]
A&E visit			£162 per attendance	NHS Reference costs 2019–20[31]
Primary care contacts	See methods	C-GALL PQs		Unit Costs of Health and Social Care, 2020[52]
GP visits			£39.23 per visit	
GP home visits			£39.23 per visit	
GP phone/video consultation			£15.32 per consultation	
Medications	See methods	PQs	Various	British National Formulary[53]

Observation/conservative management

Observation/conservative management involves the prescription of analgesics to relieve the biliary pain (paracetamol, NSAIDs, e.g. ibuprofen etc.; narcotic analgesics, e.g. opiates; antispasmodics, e.g. Buscopan), together with generic lifestyle advice.[2] Furthermore, active monitoring is recommended for individuals without symptoms. This means that participants could let their GP know if any changes in symptoms occurred. Data on medications and GP contacts were collected through PQs. The British National Formulary (BNF)[53] was used to value medications. GP contacts were costed utilising the cost per contact (i.e. visit, telephone consultation or online consultation) reported by Curtis and Burns.[52] The cost for a GP-led telephone triage was used to cost telephone consultations with the GP.

Costs of hospital admissions, perioperative complications and hospital re-admissions

The costs of clinical management of any procedure due to perioperative complications as well as any hospital admissions were based on NHS reference costs.[31] The information on the type of complication experienced and details of the procedures undertaken were obtained from the C-GALL CRFs and associated SAE forms. These included events such as biliary colic, cholecystitis, gallstone pancreatitis and choledocholithiasis (see *Table 8*, *Chapter 3*). Participants who reported no hospital admissions or had no relevant CRF completed were assumed to have zero cost due to hospital admissions.

Costs of hospital outpatient and primary care healthcare utilisation

Primary and secondary outpatient contacts related to gallstone disease incurred over the 24-month follow-up period were obtained from the PQs. Medications prescribed for any ongoing problems related to the patient's gallstone condition were also recorded in these questionnaires. All the primary care contacts were costed using the unit cost of Health and Social Care,[52] and outpatient visits were costed using the NHS reference cost for a general surgery outpatient visit.[31] The number of visits for each patient was multiplied by the appropriate unit cost. The list of medications and quantities primarily prescribed in primary care were costed using prices recorded in the BNF.[53]

Outcome measures

Quality-adjusted life-years were defined as the measure of effectiveness for the cost–utility analysis. QALYs were estimated using participants' responses to the SF-36 questionnaire completed at baseline, 3, 9, 12, 18 and 24 months post randomisation. The SF-36 questionnaire was selected as the recall time for this instrument is the last 4 weeks. We believe this recall time has a better chance to reflect participants' QoL variations due to acute events associated with gallstones compared to other instruments such as the EuroQol-5 Dimensions (EQ-5D) with a focus on 'your health today'.[55] Participant responses were assigned a utility score based on the Short Form-6 Dimensions (SF-6D) UK tariff.[45,46] The appropriate algorithm for the SF-36 version 2 was recreated in Stata[56] to obtain the health state utility weights. QALYs were estimated using the AUC approach, assuming linear change in utility between the observed follow-up time points.

Statistical analysis of trial economic data

Aggregating costs and effects

Costs and QoL data were summed up for each participant for the 24-month follow-up period and total cost and QALYs per participant were obtained. Mirroring the statistical analysis, the principles of ITT analysis were followed to compare cost and QALYs between groups.

Missing data

Reliance on complete case data for cost-effectiveness analysis can introduce bias unless the data are missing completely at random. The total estimated cost is the sum of numerous components over the observed 24-month follow-up period of the trial. Besides, QALYs can only be computed where

participants have responded to the relevant QoL questionnaires at every follow-up point. Therefore, economic evaluations based on participant-level trial data are likely to face problems with missing data. MI[43,56] was implemented as part of the primary analysis, using chained equations with predicted mean matching (*k*th-nearest-neighbour = 5). Implemented in Stata,[56] the method starts with the variable with fewer missing values. Predictive mean matching imputes an observed value from another individual whose predicted value is close to the predicted value of the individual with the missing observation. Twenty imputed data sets were generated with plausible fitted values assigned for missing cost and utility elements. The imputation was conducted at the cost variable level (e.g. inpatient care, outpatient visits and primary care contacts) and SF-6D utility score level. The imputation model included all the variables in the analysis model (age, gender and treatment group allocation) and auxiliary variables that may help to explain missingness (trial centre, indicator for having surgery and type of procedure). Rubin's rules were used to pool estimates across the MI data sets.[57]

Incremental cost-effectiveness analysis

The economic analysis estimated the incremental cost and QALYs between observation/conservative management group and laparoscopic cholecystectomy group up to 24 months post randomisation. General linear regression models adjusted for minimisation factors (centre, age and gender) and baseline SF-6D score were used. Family function was based on a modified Park test, while a modified Hosmer and Lemeshow, Pearson correlation and Pregibon link tests were used to select the link function[43] (see *Appendix 4*, *Tables 49* and *50*, for details). The test results suggested a Gaussian family and identity link as the best model for estimating cost and QALY differences. Adjusted mean values by treatment allocation and the incremental difference between the groups were obtained using the methods of recycled predictions.[43] The incremental cost-effectiveness ratio (ICER) for observation/conservative management group versus laparoscopic cholecystectomy group was calculated as the difference in mean costs divided by the difference in mean QALYs. Moreover, the uncertainty surrounding the joint incremental costs and effects was characterised using non-parametric bootstrapping with 1000 iterations with the MI process (k = 5 and 20 simulated data sets) nested within the bootstrapping process.[58] Results from the bootstrap iterations were used to obtain 95% credible intervals (95% CrI) for the incremental cost and incremental QALYs. While the micro-costing approach makes better use of the trial data, the HRG-based costing might better reflect the opportunity costs for the UK NHS. Therefore, the HRG-based costing was defined as the base-case analysis and all sensitivity analyses were conducted with the HRG-based costing.

Sensitivity analyses

Sensitivity analyses were conducted by predefined subgroups and explored the cost-effectiveness of observation/conservative management versus laparoscopic cholecystectomy according to gender (male/female), age (< 35; 35–64; ≥ 65 years) and ethnicity (white; non-white). Treatment allocation by subgroup interaction terms was defined in the regression models for this analysis. A further analysis was conducted using micro costing for the index surgical intervention.

Results

Resource use and costs

Resource use and cost from the NHS perspective by treatment allocation are presented in *Table 12*. Cholecystectomies in the observation/conservative management group required slightly longer time in theatre (mean 83 minutes) than those in the laparoscopic cholecystectomy group (mean 72 minutes) and longer hospital stays (1.4 days vs. 0.6 days). HRG-based costing was used as the primary costing method to cost cholecystectomy episodes with the cost of the episode also reported in *Table 12*. The costs of time in the anaesthetic room (16 minutes) and recovery room (71 minutes) were added to the costs of the time in theatre and hospital stay for the micro-costing of the cholecystectomy episode. One hundred and ninety-five participants provided complete data on time in theatre and length of hospital stay. The average cost per cholecystectomy episode was £839 and £1912 for the observation/

TABLE 12 Health service resource use and costs by treatment allocation

Variable	No. of observations	Observation/conservative management *N* = 217	Laparoscopic cholecystectomy *N* = 217
Resource use			
Cholecystectomies; *n* (%)	434	64 (29)	153 (71)
Time in theatre (minutes); mean (SD)	195	83.2 (49)	71.6 (42)
Length of stay (days); mean [SD; median; (IQR)]	215	1.4 [3.4; 1; (0–1)]	0.63 [1.3; 0; (0–1)]
Further follow-up			
Further hospital admission; *n* (%)	434	28 (12.9)	21 (9.7)
Length of stay (days); mean [SD; median; (IQR)]	49	2.9 [3.1; 1.3; (0.3–6)]	3.1 [5.1; 1; (0–2)]
Outpatient visits; mean [SD; median; (IQR)]	188	0.6 [1.5; 0; (0–1)]	0.6 [1.3; 0; (0–1)]
Accident and emergency visits; mean [SD; median; (IQR)]	191	0.5 [1.3; 0; (0–0)]	0.2 [0.6; 0; (0–0)]
GP contacts			
Face-to-face visits; mean [SD; median; (IQR)]	151	1 [1.9; 0; (0–1)]	1.6 [3.1; 1; (0–2)]
Phone consultations; mean [SD; median; (IQR)]	142	0.6 [1.6; 0; (0–0)]	0.3 [0.8; 0; (0–0)]
Home visits; mean [SD; median; (IQR)]	137	0 [0.1; 0; (0–0)]	0 [0.2; 0; (0–0)]
Practice nurse; mean [SD; median; (IQR)]	154	0.1 [0.5; 0; (0–0)]	0.3 [0.8; 0; (0–0)]
Medication prescribed; *n* (%)	425	26 (12.2)	30 (14.2)
Costs (£)			
Initial surgical episode cost			
Micro-costing-based estimates; mean (SD)	401	839 (1599)	1912 (1702)
HRG-based costing; mean (SD)	434	1014 (1757)	2190 (1536)
Re-admission costs for further treatment; mean (SD)	434	242 (1398)	121 (603)
Outpatient costs; mean (SD)	188	89 (223)	87 (193)
Accident and emergency visits; mean (SD)	191	76 (205)	36 (92)
Primary care costs; mean (SD)			
Face-to-face visits; mean (SD)	151	39 (74)	61 (121)
Phone consultations; mean (SD)	142	9 (24)	4 (12)
Home visits; mean (SD)	137	0.59 (4.8)	1.11 (9.3)
Practice nurse; mean (SD)	154	1.49 (6.3)	4.61 (10.7)
Medication costs; mean (SD)	425	4 (18)	7 (24)
Total NHS cost; mean (SD)	118	885 (1954)	2184 (1486)

conservative management and laparoscopic cholecystectomy groups, respectively. NHS reference costs were used as the alternative costing method. The mean cost of cholecystectomies is substantially lower in the observation/conservative management group (£1014; *N* = 217) compared to the laparoscopic cholecystectomy group (£2190; *N* = 217), due to the lower percentage undergoing treatment.

Table 12 also reports follow-up resource use up to 24 months post randomisation. Participants in the observation/conservative management group spent more time in hospital in the follow-up period. Reasons for hospital admissions included biliary colic, cholecystitis, gallstone pancreatitis and choledocholithiasis. The average costs of these events were higher in the observation/conservative management group (£242 vs. £121). There were no differences in the mean number of outpatient visits between groups, but there was a higher mean number of visits to A&E department in the observation/conservative management group. While GP face-to-face and practice nurse contacts were higher in the surgical group, GP phone consultations were lower. However, none of these differences appear statistically significant by visual inspection of the SD for each group.

The overall costs for the initial and follow-up healthcare result in a total average NHS cost of £885 for observation/conservative management and £2184 for cholecystectomy, giving an unadjusted difference of £1299.

Utility scores and quality-adjusted life-years

Mean health state utility scores based on the SF-6D algorithm[45,46] by treatment allocation are reported in *Table 13*. QoL data were collected at baseline, and at 3, 9, 12, 18 and 24 months post randomisation. At baseline, there was a small non-significant difference in the mean utility score favouring observation/conservative management. Mean utility score is also slightly higher for this group at 3 months. However, the mean score for laparoscopic cholecystectomy group increases from month 3 to 9 surpassing the corresponding score for the observation/conservative management group. This probably reflects participants going through, and recovering from, surgery in the laparoscopic cholecystectomy group as the median time to surgery was 4.7 months (compared with 9 months for observation/conservative management group). Both study groups show an increase in the mean utility score from 9 months to 18 months and a fall mean utility score value at 24 months. A small non-significant mean QALY difference favouring laparoscopic cholecystectomy emerges from the calculation of QALYs based on 30% of participants reporting complete data.

TABLE 13 SF-6D health state utility scores by treatment allocation

Variable	No. of observations	Observation/conservative management *N* = 217	Laparoscopic cholecystectomy *N* = 217
SF-6D score			
Baseline; mean (SD)	424	0.701 (0.13)	0.682 (0.13)
3 months post randomisation; mean (SD)	294	0.685 (0.13)	0.667 (0.13)
9 months post randomisation; mean (SD)	255	0.678 (0.14)	0.704 (0.15)
12 months post randomisation; mean (SD)	236	0.695 (0.14)	0.717 (0.14)
18 months post randomisation; mean (SD)	256	0.733 (0.15)	0.750 (0.15)
24 months post randomisation; mean (SD)	226	0.716 (0.14)	0.710 (0.16)
QALYs; mean (SD)	128	1.4 (0.21)	1.45 (0.23)

Cost–utility analysis results

The base-case cost–utility analysis results based on the MI data set and adjusted by minimisation factors and baseline SF-6D utility score are reported in *Table 14*.

The adjusted mean costs per participant were £2510 and £1477 for the laparoscopic cholecystectomy and observation/conservative management groups, respectively, resulting in an adjusted cost difference of −£1033 (95% CrI: −1413 to −632). The mean adjusted QALYs per participant were 1.413 and 1.395 for the laparoscopic cholecystectomy and observation/conservative management groups, respectively, producing an adjusted QALY difference of −0.019 (95% CrI: −0.06 to 0.02) for the 24-month follow-up period. Therefore, ITT with laparoscopic cholecystectomy as the first option results in significantly higher mean costs and non-significantly higher-mean QALYs compared with observation/conservative management. The ICER between observation/conservative management and laparoscopic cholecystectomy was £55,235. Therefore, moving from the standard practice of laparoscopic cholecystectomy to observation/conservative management would result, on average, in lower costs and QALYs with a saving of £55,235 per QALY forgone.

The results from the 1000 bootstrapped iterations of the regression analysis nested MI are presented in the cost-effectiveness plane (*Figure 7*). Incremental costs and QALYs for observation/conservative management versus laparoscopic cholecystectomy are reported in this scatterplot. All the iterations show negative mean cost differences, meaning that observation/conservative management is less costly than laparoscopic cholecystectomy. Although 16% of the iterations resulted in higher QALYs for observation/conservative management, the MD in QALYs (−0.019) favoured laparoscopic cholecystectomy. The red dashed line in *Figure 7* represents the £20,000 cost-effectiveness threshold.

TABLE 14 Base-case analysis results – MI

Intervention	Total cost (£)	Incremental cost (£)[a] (95% CrI)	Total QALYs	Incremental QALY[a] (95% CrI)	ICER
Laparoscopic cholecystectomy	2510		1.413		
Observation/conservative management	1477	−1033 (−1413 to −632)	1.395	−0.019 (−0.06 to 0.02)	55,235

a Differences adjusted by study minimisation variables and baseline SF-6D score.

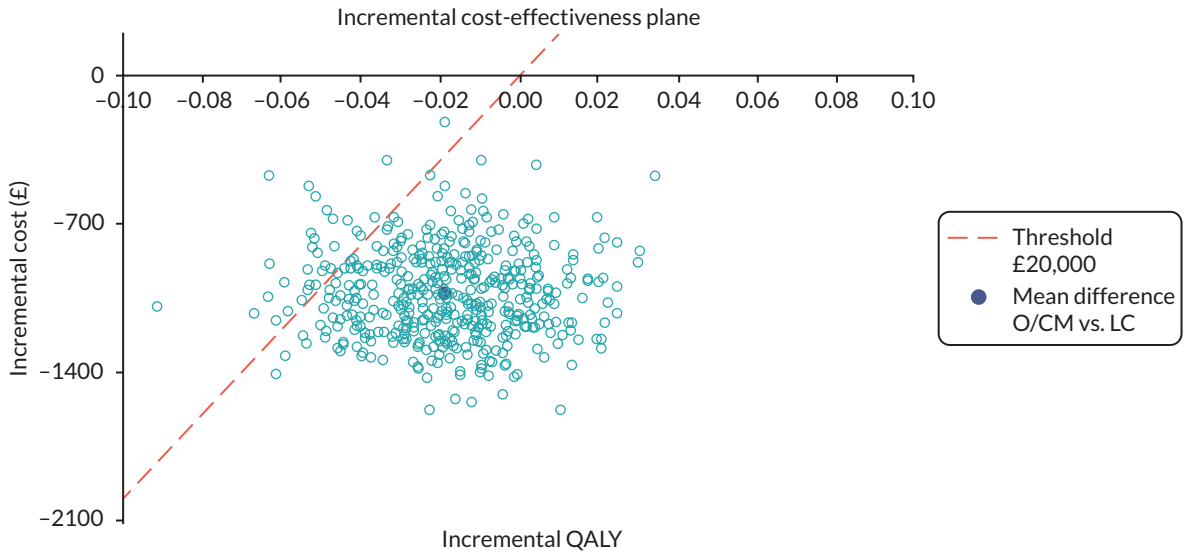

FIGURE 7 Trial-based incremental cost-effectiveness scatterplot for observation/conservative management vs. laparoscopic cholecystectomy (base case; imputed data – 1000 bootstrap iterations).

This threshold separates iterations that result in either trial group being cost-effective. Iterations sitting to the right (left) and below (above) the red dashed line indicate that observation/conservative management (laparoscopic cholecystectomy) is cost-effective. Just 7% of the iterations sit above and to the left of the £20,000 threshold[59] dashed line; that is just 7% of the iterations show QALY differences large enough for laparoscopic cholecystectomy to be cost-effective (see *Figure 7*).

The cost-effectiveness acceptability curves (CEACs) for laparoscopic cholecystectomy and observation/conservative management are reported in *Figure 8*. The CEACs show the probability of observation/conservative management or laparoscopic cholecystectomy being cost-effective at alternative cost-effectiveness threshold values. The CEAC for observation/conservative management shows a 94% probability of being cost-effective at £20,000 threshold value. This probability decreases at higher threshold levels reflecting the higher mean cost and QALYs resulting from laparoscopic cholecystectomy, but conservative management has the higher probability of cost-effectiveness up to a cost-effectiveness threshold of £55,000.

Alternative costing method and subgroup analyses

Results for the alternative costing methodology for the index laparoscopic cholecystectomy and the subgroup analyses are summarised in *Table 15*. The analysis using micro costing to value cholecystectomies generates a slightly smaller cost difference between the study groups (−£930 compared to −£1033) resulting in a lower ICER (£49,747). As for the base-case analysis, this ICER is above the usual cost-effectiveness threshold used in the UK.[59]

The subgroup analyses for gender and age give similar results to those observed for the base case. Although observation/conservative management shows higher mean QALYs for males and for participants over 65 years old, and the subgroup analysis for non-white favours laparoscopic cholecystectomy with an ICER of £15,627, the results for these analyses should be interpreted with caution as they are due to small numbers. It should be noted that all *p*-values for the interaction terms between the subgroup indicator and treatment effect variables were > 0.05 (see *Appendix 4*, *Table 51*); that is, there were no statistically significant differences in the incremental cost or QALYs between the subgroup.

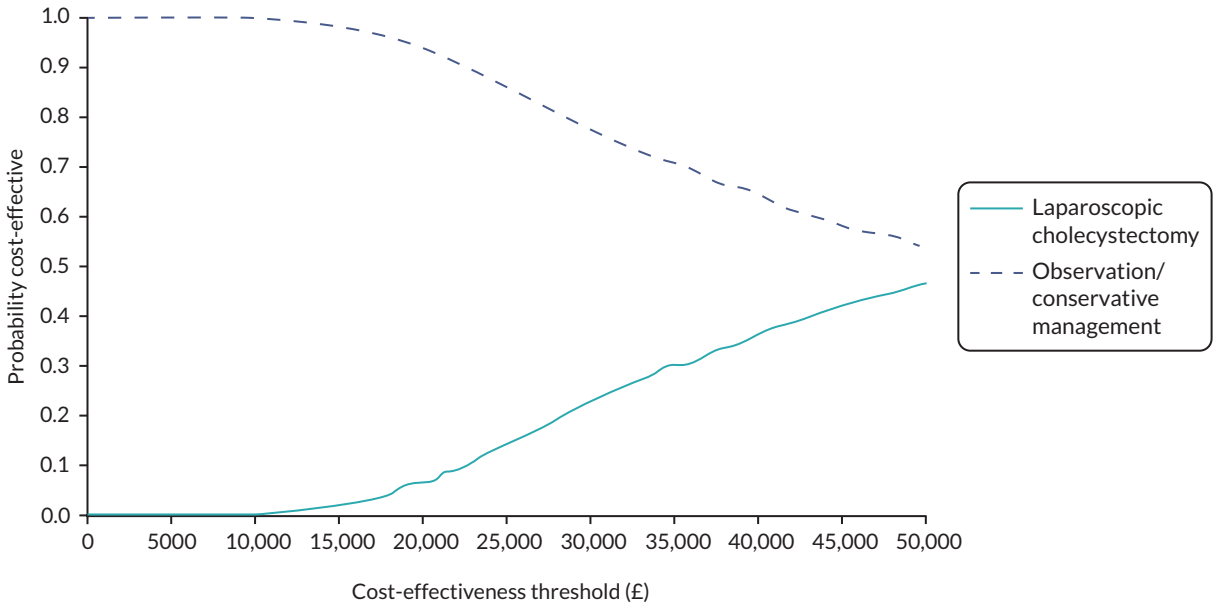

FIGURE 8 Cost-effectiveness acceptability curves for observation/conservative management and laparoscopic cholecystectomy groups (trial-based analysis; base case; imputed data – 1000 bootstrap iterations).

TABLE 15 Trial-based incremental cost-effectiveness analysis by alternative costing methods and predefined subgroups

Intervention	Total cost (£)	Δ cost (£) (95% CrI)[a]	Total QALYs	Δ QALY (95% CrI)[a]	ICER[b]	Probability cost-effective		
						£13,000	£20,000	£30,000
Base-case analysis (full cohort; HRG-based costing)								
Laparoscopic cholecystectomy	2510		1.413			0.011	0.065	0.229
Observation/conservative management	1477	−1033 (−1413 to −632)	1.395	−0.019 (−0.06 to 0.02)	55,235	0.989	0.935	0.771
Laparoscopic cholecystectomy (micro-costing)								
Laparoscopic cholecystectomy	2322		1.413			0.018	0.097	0.271
Observation/conservative management	1391	−930 (−1344 to −539)	1.395	−0.019 (−0.06 to 0.02)	49,747	0.982	0.903	0.729
Gender, females								
Laparoscopic cholecystectomy	2506		1.411			0.036	0.154	0.344
Observation/conservative management	1600	−906 (−1328 to −449)	1.391	−0.020 (−0.07 to 0.02)	45,752	0.964	0.846	0.656
Gender, males								
Laparoscopic cholecystectomy	2527		1.423			0.018	0.049	0.142
Observation/conservative management	1014	−1514 (−2226 to −639)	1.410	−0.014 (−0.07 to 0.06)	111,310	0.982	0.951	0.858
Age 65 years and over								
Laparoscopic cholecystectomy	2750		1.317			0.011	0.094	0.281
Observation/conservative management	1873	−878 (−1915 to 491)	1.322	0.005 (−0.06 to 0.07)	Dominant[c]	0.989	0.906	0.719
Age < 65 years								
Laparoscopic cholecystectomy	2442		1.440			0.137	0.186	0.258
Observation/conservative management	1365	−1077 (−1456 to −731)	1.415	−0.025 (−0.07 to 0.01)	42,563	0.863	0.814	0.742

TABLE 15 Trial-based incremental cost-effectiveness analysis by alternative costing methods and predefined subgroups (*continued*)

Intervention	Total cost (£)	Δ cost (£) (95% CrI)[a]	Total QALYs	Δ QALY (95% CrI)[a]	ICER[b]	Probability cost-effective		
						£13,000	£20,000	£30,000
Ethnicity; white								
Laparoscopic cholecystectomy	2524		1.410			0.008	0.046	0.155
Observation/conservative management	1505	−1018 (−1443 to −558)	1.399	−0.011 (−0.05 to 0.03)	93,408	0.992	0.954	0.845
Ethnicity; non-white								
Laparoscopic cholecystectomy	2417		1.439			0.341	0.571	0.749
Observation/conservative management	1341	−1075 (−1845 to −193)	1.370	−0.069 (−0.14 to 0.01)	15,627	0.659	0.429	0.251

a Δ, incremental; ICER: incremental cost-effectiveness ratio.
b £ per QALY spent.
c Dominant means that, on average, observation/conservative management is less costly and produce more expected QALYs than laparoscopic cholecystectomy.

Note
Differences adjusted by study minimisation variables.

Summary and discussion

The results for the within-trial cost–utility analysis reported in this chapter indicate that ITT with observation/conservative management was, over a 24-month follow-up period, less costly than cholecystectomy (MD −£1033). This cost difference was driven by the higher number of cholecystectomy procedures undertaken in the laparoscopic cholecystectomy group as there was no difference in the cost of other events triggering hospital admissions between the groups. A non-significant QALY difference of −0.019 was observed. This is consistent with the C-GALL trial main statistical analysis that showed no difference in primary outcome of SF-36 bodily pain AUC up to 24 months. The significant cost difference, together with a small non-significant QALY difference favouring cholecystectomy (cholecystectomy is more effective by 0.019 QALYs), resulted in an ICER of £55,235. That is, moving from the standard practice of ITT with laparoscopic cholecystectomy to observation/conservative management would result, on average, in lower costs and QALYs with a saving of £55,235 per QALY lost. The probabilistic analysis showed observation/conservative management having a high probability of being cost-effective (97%) and the sensitivity analyses showed these results to be robust to the alternative costing approach using micro-costing for cholecystectomies. Prespecified subgroup analyses (by gender, age or ethnicity) showed that these findings were generally consistent across subgroups.

The main strength of this within-trial analysis relates to the randomised controlled trial (RCT) study design. The resource use and health-related QoL data used were collected prospectively for individual participants as part of the C-GALL trial. Unbiased and accurate estimation of costs and QALY differences are possible due to the randomised treatment allocation. The pragmatic nature of the C-GALL trial together with the ITT principles facilitates the generalisability of the findings to the patient population treated routinely in the UK NHS. A further strength is that the economic analysis was conducted according to a prespecified and agreed health economics analysis plan.

The pragmatic RCT design is a strength of the analysis. The imposed restrictions on elective procedures such as cholecystectomy might have de facto implemented observation/conservative management on both trial groups for some participants for a limited period of time. A second important limitation relates to the short follow-up and the natural history of gallstone disease. Cholecystectomy might be indicated in the future for those participants in the observation/conservative management group as well as those still waiting for surgery in the cholecystectomy groups. Schmidt *et al.*[60] reported on 14-year follow-up for a RCT that randomised 137 participants to observation or cholecystectomy. The authors reported that median time to cholecystectomy for those participants in the observation group was 28 months. The cost of cholecystectomy was the driver of the cost difference. Participants crossing over from observation/conservative management to the laparoscopic cholecystectomy group may result in improved cost-effectiveness for laparoscopic cholecystectomy over a longer time horizon. Therefore, there is a clear case for a longer-term follow-up of the C-GALL trial. Until then, an extrapolation exercise beyond the trial follow-up is conducted in *Chapter 5* using a Markov model over a 10-year time horizon.

Chapter 5 Economic modelling

Introduction

The aim of this chapter is to report on the model-based extrapolation to extend the economic analysis horizon beyond the 24-month trial follow-up, up to 10 years. While a clear answer for the cost-effectiveness of observation/conservative management versus laparoscopic cholecystectomy was obtained from the within-trial analysis, it was anticipated that further participants would need to undergo surgery for their gallstone disease after 24 months[12,60] and that this could have an impact on the cost-effectiveness results. Moreover, current method guidelines for the conduct of economic evaluations[61] recommend using a time horizon long enough to capture all relevant differences in costs and consequences between the strategies being compared. Therefore, a simple Markov model was developed to estimate longer-term economic differences beyond the 24-month trial follow-up.

Methods

Model structure

The Markov model was constructed in TreeAge Pro software (TreeAge Software, Inc., Williamstown, MA) and was informed by reviewing decision models identified in the literature.[35] A similar structure was used for both trial groups and is illustrated in *Figure 9*. A cohort of individuals referred to secondary care and with confirmed symptomatic gallstones enter the model in the '*No surgery*' health state and are assigned to either laparoscopic *cholecystectomy* or observation/conservative management. Individuals receive treatment as observed in the C-GALL trial on an ITT basis. As such, the model considers the waiting time which occurred in the trial follow-up for the first 24 months, with survival analysis used to extrapolate time from randomisation to surgical procedures. AEs (and costs) related to gallstone disease such as biliary colic or cholecystitis are experienced by a proportion of individuals in the '*No surgery*' health state. Individuals undergoing surgery (cholecystectomy) accrue the cost of the surgical episode and move to the tunnel state '*Recovery from surgery*' where a reduction in quality of life associated with the surgical intervention is accounted for. Markov tunnel states are a series of temporary states, lasting only one cycle, that must be visited in a fixed sequence.[62] On exit from the tunnel Markov state, individuals either have their symptoms resolved or not, moving to the corresponding Markov state (i.e.

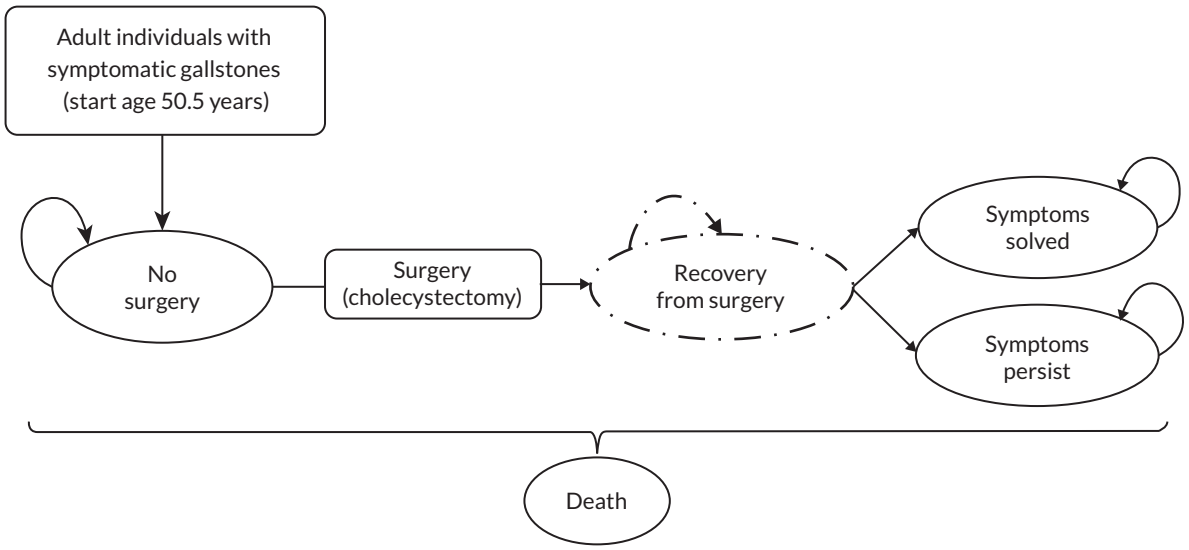

FIGURE 9 Simplified schematic for the C-GALL Markov model.

'*Symptoms solved*' or '*Symptoms persist*'). Finally, Markov models need at least one state that cannot be left, named absorbing states, for the Markov process to end. In the C-GALL economic evaluation model, the absorbing state '*Death*' is included, which can be entered from any other state based on age-specific mortality rates reported in the UK life tables.[63]

The state occupancy and estimated payoffs are updated on a constant monthly Markov cycle.

Population

The model analysis was conducted for a cohort of individuals with characteristics matching those of the participants in the C-GALL trial cohort (baseline mean age and gender proportion 50.5 years and 71% women, respectively). The ITT principle was followed to estimate input parameters for the economic model; therefore, the model reflects the fact that a number of participants in the C-GALL trial did not receive their intended treatment.

Time horizon and discounting

Based on the available literature,[60] and upon discussion within the PMG, a time horizon of 10 years was adopted for this analysis. This decision was based on the expectation that most of the individuals needing surgery would go through this procedure in the first few years after the end of their 24-month C-GALL trial follow-up. Survival analyses of time to surgery (see *Clinical input parameters*) also show that the hazard transition to surgery declines through time with the Kaplan–Meier (KM) curves levelling off before the 10-year mark. The impact of adopting a medium-term time horizon of 5 years was assessed in a sensitivity analysis. The model starts at baseline/randomisation time with costs and QALYs accrued beyond year 1 discounted at an annual discount rate of 3.5%.[61]

Clinical input parameters

Surgery

The crucial clinical parameter in the model is the time-dependent risk of undergoing cholecystectomy. All participants in the C-GALL trial were followed for a minimum of 24 months. In addition, PQs were sent every 6 months for those completing the trial follow-up at 24 months. Therefore, information about surgical procedures beyond 24 months was available for some participants for up to 48 months post randomisation. *Figure 10* shows the KM curve for cholecystectomies for all participants in the C-GALL trial.

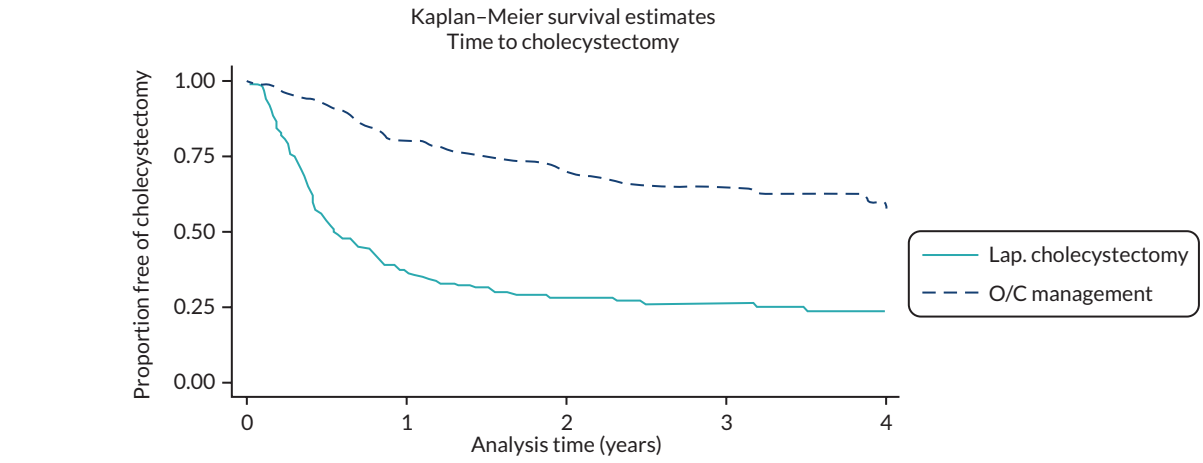

FIGURE 10 Kaplan–Meier plot of time in years for cholecystectomy by treatment allocation. Lap. cholecystectomy, laparoscopic cholecystectomy; O/C management, observation/conservative management.

Parametric survival functions (i.e. exponential, Weibull, Gompertz, log-logistic and generalised gamma) were fitted to 48-month data. However, visual inspection made clear that all curves provided a poor visual fit to observed data, underestimating the earlier proportion of surgeries and overestimating the proportion of surgeries at 48 months (and possibly after). An alternative piecewise approach was then taken, with the first 9 and 12 months of data dropped from the analysis, and parametric curves fitted to the remaining tails of the distributions. The statistical fit of the fitted curves, assessed using the Bayesian and Akaike information criteria, are reported in *Appendix 4*, *Table 52*. The estimated survival curves and the percentages of participants expected to have surgery at 10 years were discussed within the C-GALL PMG (*Table 16*). Furthermore, Schmidt *et al.*[60] stated that almost no operations took place beyond 5 years, when reporting on the 14-year follow-up of a study that randomised 137 participants to observation or cholecystectomy. Given the proportion of C-GALL participants known to have had surgery at 24 months, coupled with those declining surgery in the laparoscopic cholecystectomy group, and based on the advice of the clinical experts, it was agreed that the proportion of participants having cholecystectomy at 10 years would be around 50% and 80% for the observation/conservative management and laparoscopic cholecystectomy groups, respectively. Therefore, it was agreed to use the generalised gamma function fitted to the 9-month time-to-surgery data for the base-case analysis and the Gompertz and log-logistic functions (also fitted to the 9-month data) for sensitivity analyses.

Monthly transition probabilities were obtained directly from the KM curves for the first 24 months of the model to mirror the survival observed in the C-GALL trial, and from the fitted survival curves thereafter.

Adverse events related to gallstone disease

Individuals not going to surgery can experience AEs related to gallstone disease such as biliary colic or cholelithiasis. C-GALL trial data were used to estimate the rate and mean cost for 'presurgical' events. Presurgical events were defined as events needing a hospital admission (e.g. A&E admission) with an onset date earlier than the date of surgery or end of follow-up, whichever happened first. The rate of

TABLE 16 Number of participants estimated to undergo surgery by treatment allocation and alternative survival function

	Observation/conservative management, *n* (%)		Laparoscopic cholecystectomy, *n* (%)	
	With surgery	Without surgery	With surgery	Without surgery
C-GALL data at 24 months	64 (29.5)	153 (70.5)	153 (70.5)	64 (29.5)
Survival functions fitted to post 9-month time-to-surgery data				
Exponential	163 (75.2)	54 (24.8)	211 (97.1)	6 (2.9)
Weibull	127 (58.3)	90 (41.7)	196 (90.2)	21 (9.8)
Gompertz	88 (40.7)	129 (59.3)	172 (79.2)	45 (20.8)
Log-logistic	121 (55.6)	96 (44.4)	188 (86.8)	29 (13.2)
Generalised gamma	111 (51)	106 (49)	180 (82.8)	37 (17.2)
Survival functions fitted to post 12-month time-to-surgery data				
Exponential	165 (75.9)	52 (24.1)	All	None
Weibull	142 (65.7)	75 (34.3)	207 (95.5)	10 (4.5)
Gompertz	105 (48.2)	112 (51.8)	185 (85)	32 (15)
Log-logistic	133 (61.4)	84 (38.6)	199 (91.7)	18 (8.3)
Generalised gamma	123 (56.8)	94 (43.2)	195 (89.9)	22 (10.1)

presurgical events for the model was estimated using time-to-event data, calculating the rate of events per patient year. This was transformed into a monthly rate of 0.0068 events per patient month.

Health state utilities

The health state utilities applied in the model were estimated from SF-36 data collected in the C-GALL trial for the first 24 months (*Table 17*) and retrieved from the literature thereafter.[64] Based on completed trial data, mean SF-6D scores and SDs by treatment allocation were used for the first 24 months of the model with beta distributions associated with these data for the probabilistic sensitivity analysis (PSA). Van den Berg (2012)[64] derived SF-6D population norms for the UK using a sample of 22,166 respondents to a national representative sample of British citizens were used. These population norms were adjusted according to age and gender in the C-GALL trial. Following the methods proposed by Ara and Brazier,[65] a multiplier was calculated as the ratio between the mean SF-6D score at 24 months and the corresponding age-specific population norm SF-6D score (weighted by the C-GALL gender proportions). The 95% CI reported in Van den Berg 2012[64] was used to estimate standard errors and build up beta distributions to reflect the parameter uncertainty associated with the mean utility scores (see *Table 17*).

Table 17 also reports the utility decrement related to the need for cholecystectomy surgery estimated using data from the C-GALL trial. SF-6D utility scores for the time points before and after the date of surgery were selected. A paired *t*-test was used to test the difference between SF-6D mean scores before and after surgery. The statistically significant mean score difference was then used to calculate the percentage reduction from the postsurgical utility score. The utility decrement was applied to (half of) the mean time between the utility scores before and after surgery obtained from C-GALL data. Applying the utility decrement to half of the mean time assumes a linear extrapolation between the lower (before surgery) and higher (after surgery) utility scores.

TABLE 17 SF-6D health state utility scores applied in the economic model

Month	Observation/conservative management, mean (SE)	Cholecystectomy, mean (SE)	Distributional form	Source
SF-6D score				
Baseline	0.701 (0.13)	0.682 (0.13)	Beta	C-GALL trial
3 months	0.685 (0.13)	0.667 (0.13)	Beta	C-GALL trial
9 months	0.678 (0.14)	0.704 (0.15)	Beta	C-GALL trial
12 months	0.695 (0.14)	0.717 (0.14)	Beta	C-GALL trial
18 months	0.733 (0.15)	0.756 (0.15)	Beta	C-GALL trial
24 months	0.716 (0.14)	0.716 (0.16)	Beta	C-GALL trial
C-GALL multiplier		0.915		
	Males, mean (SE)	**Females, mean (SE)**	**C-GALL pop norm**[a]	
50–54 years	0.794 (0.006)	0.785 (0.004)	0.716	Based on Van den Berg 2012[64]
55–59 years	0.803 (0.005)	0.779 (0.004)	0.704	Based on Van den Berg 2012[64]
60–64 years	0.782 (0.006)	0.775 (0.004)	0.706	Based on Van den Berg 2012[64]

a Twenty-one per cent males plus 79% females scores and multiplied by the C-GALL multiplier. Utility scores used after *recovery from surgery* Markov tunnel state.

Health service resource use and costs

Costs for the cholecystectomy episode for the model were obtained from NHS reference costs[31] (*Table 18*). Unit costs for elective laparoscopic cholecystectomy (HRG code GA10K) were assumed for all cholecystectomies in the model. Unit costs for day-case and non-elective long-stay (emergency) cholecystectomies were used to define alternative scenarios (see *Table 18*). The emergency cholecystectomy scenario is assumed to incorporate perioperative and postsurgery AEs associated with the index cholecystectomy procedure.

Additional health service costs' input data for the model were informed by the analysis of C-GALL trial data. A total of 55 AEs associated with gallstone disease such as biliary colic or abdominal pain needing hospital admission were counted for participants who were either waiting for surgery or for whom surgery was not yet being considered. Mean cost and SD for these events were used to attach a cost to the rate of presurgical events that was applied to those participants staying in the *No surgery* Markov health state (see *Adverse events related to gallstone disease*).

As no differences in the number of outpatient or GP contacts were observed by study group in the C-GALL trial, a simplifying assumption was made and no cost for these contacts was incorporated in the model.

Finally, gamma distributions were used for all unit costs to explore parameter uncertainty by conducting PSA.

Model validation

A number of steps were taken in order to secure the internal validity of the model. Testing was conducted throughout the model implementation such as verification of the model formulae using Microsoft Excel® (Microsoft Corporation, Redmond, WA, USA) (white-box testing). The final model results were checked when varying the model input parameter values (e.g. defining all utility scores equal to one and zero discount rate to check whether total QALYs equal total life years) to assess the consistency of the model performance given specific variation in the model input values (black-box testing).

The validity of the projected estimates of cholecystectomies was assessed by plotting the model Markov traces with the Kaplan–Maier curves for the C-GALL trial (*Figure 11*). Visual inspection shows the Markov traces follow closely the C-GALL Kaplan–Maier curves. In addition, the rate of cholecystectomies at 10 years extracted from the model base case are 48% and 79% for observation/conservative management and cholecystectomy groups, respectively. This is in line with the expected rates discussed in the PMG (i.e. 50% and 80%, respectively).

Finally, within-trial analysis and model results were compared. For this, the model was run for a 24-month time horizon using the unit cost for the day-case cholecystectomy as this unit cost was

TABLE 18 Unit costs for cholecystectomy episode and presurgical events

	Unit cost (£) mean (SE)[a]	Source
Laparoscopic cholecystectomy		
Elective	3579 (358)	HRG GA10K; NHS Reference cost
Day-case	2693 (269)	HRG GA10K; NHS Reference cost
Non-elective long stay	4883 (488)	HRG GA10K; NHS Reference cost
Presurgical event	853 (677)	C-GALL trial

a Assumption.

Note
Unit cost expressed in Great British Pounds, 2019–20 price year.

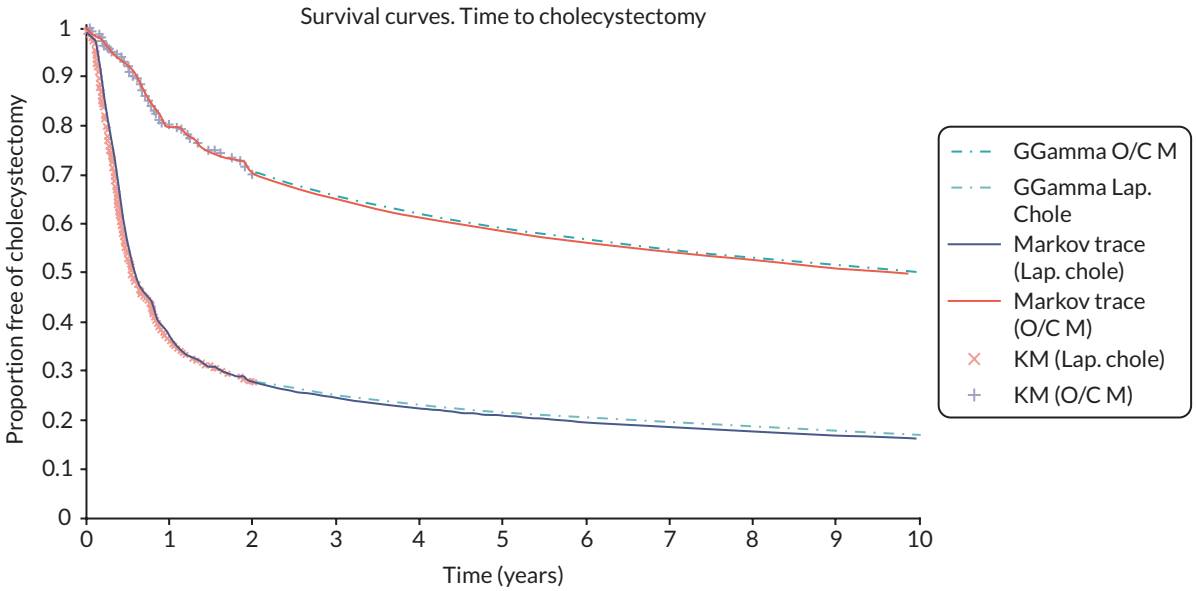

FIGURE 11 Kaplan–Meier and model Markov trace plot of time for cholecystectomy by treatment allocation. GGamma, generalised gamma; KM, Kaplan–Meier; Lap. Chole, laparoscopic cholecystectomy; O/C M, observation/conservative management.

similar to the cost of the cholecystectomy episode resulting from the within-trial analysis. The cost-effectiveness results for this comparison are reported in *Appendix 4*, *Table 53*. While the model expected costs and expected QALYs up to 24 months are lower than the corresponding values for the within-trial analysis, the difference in costs and QALYs between the two analyses produce consistent but substantially different ICERs (i.e. £55,235 and £88,930 for the within-trial and model-based analyses, respectively). The difference in costs reflects the fewer cost categories incorporated in the modelling analysis. The difference in the incremental QALYs is the result of the alternative approach used to obtain the AUC (linear interpolation for the within-trial analysis vs. simple extrapolation for the model) and the fact that neither cost or QALYs are adjusted for minimisation factors in the model-based analysis.

Model analysis

The model was run deterministically (using expected costs and outcomes) and probabilistically (using second-order Monte Carlo simulations) to characterise the joint uncertainty in the incremental costs and QALYs between observation/conservative management and cholecystectomy, arising from the model input parameter values. Probability distributions were assigned to model input mean parameter values reflecting uncertainty due to sampling variation. Beta distributions were defined for probabilities (risks) and utilities, and gamma for costs. Mean and standard errors, used to define these distributions using TreeAge® (TreeAge Software, Inc., Williamstown, MA, USA) software, are provided in *Tables 17* and *18*. Probabilistic analyses were conducted by running 10,000 random draws from the allocated, producing 10,000 estimates of incremental costs and QALYs. The point estimate was calculated by averaging across the 10,000 estimates for costs and QALYs. Results tables also show the probability of observation/conservative management and laparoscopic cholecystectomy being cost-effective at £13,000, £20,000 and £30,000 per QALY thresholds.

Further deterministic analysis was conducted to assess the sensitivity of the model results to changes in key input parameters and structural assumptions. The following scenario and one-way sensitivity analyses were conducted:

1. Gompertz distribution for the survival analysis for the laparoscopic cholecystectomy and observation/conservative management groups. This results in a lower proportion of individuals going to cholecystectomy compared with the base case that used a generalised gamma survival model.

2. Log-logistic distribution for the survival analysis for the laparoscopic cholecystectomy and observation/conservative management groups. This shows a higher proportion of individuals going to cholecystectomy compared with the base case that used a generalised gamma survival model.
3. Gompertz distribution for laparoscopic cholecystectomy group and log-logistic for the observation/conservative management group. This scenario assumes a relatively higher proportion of individuals going to surgery in the observation/conservative management group and a lower proportion in the laparoscopic cholecystectomy group compared with the base case.
4. Day-case cholecystectomy cost (£2693) assumed for all cholecystectomies.
5. Emergency cholecystectomy cost (£4897) assumed for all cholecystectomies.
6. Day-case (£2693) and emergency (£4897) cholecystectomy cost assumed for laparoscopic cholecystectomy and observation/conservative management groups, respectively.
7. One-way sensitivity analysis: additional cost for surgery in the observation/conservative management group reflecting the possible higher costs associated with unexpected cholecystectomies in the observation/conservative management group, that are not necessarily associated with emergency procedures.
8. One-way sensitivity analysis. The utility reduction due to the need for cholecystectomy after 24 months is varied from 0 to 20% (base case, 5.5%).
9. One-way sensitivity analysis. Additional cost for surgery in the observation/conservative management group and scenario 3 for survival analysis.
10. Using C-GALL within-trial analysis results into the model for the first 24 months as the base case thereafter.

Although the C-GALL trial follow-up was 24 months, participants continued to be sent questionnaires every 6 months at 24 months. SF-36 data were collected and SF-6D scores obtained and analysed. A mixed-effects regression model for repeat measures with adjustment for the minimisation covariates (gender, age) and including centre as a random effect was used to obtain the SF-6D utility score difference between trial groups by data collection time point. The results of this analysis are reported in *Appendix 4*, *Table 54*. Consistent with the trial's main statistical analysis, non-significant utility score differences between randomised groups were observed for up to 24 months. However, statistically significant differences are seen for 30 and 36 months post randomisation. Therefore, a final sensitivity analysis (scenario 11) was conducted using the utility data reported in *Appendix 4*, *Table 54* over the initial 48 months of the 10-year time horizon.

Results

Base-case analysis

The results for the base-case cost–utility analysis are reported in *Table 19*. As anticipated, the expected costs and QALYs are higher for the 10-year time horizon compared to the 24-month follow-up for the within-trial analysis. The higher costs are explained by the higher proportion of individuals going to surgery and the longer time period considered. Differences in costs and QALYs decrease compared with

TABLE 19 Base-case model results

Strategy	Cost (£)	Δ cost (£)	QALYs	Δ QALYs	ICER (£/QALY)	Probability cost-effective		
						£13,000	£20,000	£30,000
Laparoscopic cholecystectomy	3020		5.910			0.000	0.016	0.109
Observation/conservative management	2016	−1003	5.894	−0.013	78,063	1.000	0.984	0.891

Δ, increment.

the within-trial base-case analysis (see *Table 14*). However, the ICER rises, indicating that the expected saving per QALY loss is increased.

The expected costs and QALY differences from the PSA are reported in *Figure 12*. Cost and QALY differences are measured in the *y*- and *x*-axis, respectively. All PSA results show observation/conservative management is less costly than laparoscopic cholecystectomy. Only 23% of the results show observation/conservative management is more effective; however, the QALY difference in favour of laparoscopic cholecystectomy is not large enough for this strategy to be considered cost-effective in the base case. Moreover, the red dotted line represents the £20,000 cost-effectiveness threshold. Just 1.6% of the PSA iterations cross over the left of this line; that is, the probability of laparoscopic cholecystectomy being cost-effective at a £20,000 threshold is just 1.6% for the base-case analysis (see *Table 19*).

The CEACs showing the probability of observation/conservative management or laparoscopic cholecystectomy being cost-effective at alternative cost-effectiveness threshold values are reported in *Figure 13*. The CEAC for observation/conservative management shows a 98% probability of being cost-effective at £20,000 threshold value. This probability decreases at higher threshold levels reflecting the higher mean cost and QALYs resulting from laparoscopic cholecystectomy, but observation/conservative management has the higher probability of cost-effectiveness up to a cost-effectiveness threshold of £78,000 (data not shown).

Scenario analyses

In *Table 20* the results of the first six scenario analyses are presented. Assuming alternative distributions such as Gompertz or log-logistic for the survival analysis, or lower (day-case) or higher (emergency procedures) unit cost for all cholecystectomies, have an impact on the ICER. However, ICERs are well above the usual cost-effectiveness threshold used in the UK for decision-making (i.e. £20,000).[66] Moreover, the probability of observation/conservative management being considered cost-effective remains very high for scenarios 1 to 5. Assuming all cholecystectomies being conducted as emergencies for the observation/conservative management and as elective procedures for the laparoscopic cholecystectomy group (scenario 6) reverse the result seen in scenarios 1 to 5. Laparoscopic cholecystectomy becomes less costly and more effective (dominating) than observation/conservative

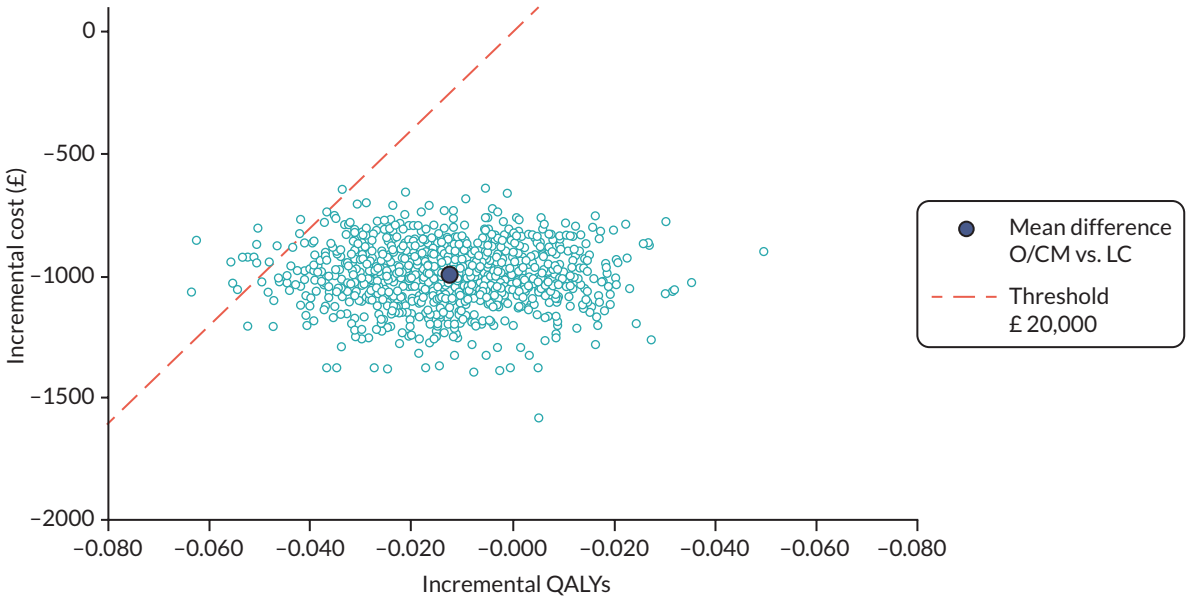

FIGURE 12 Model-based incremental cost-effectiveness scatterplot for observation/conservative management vs. laparoscopic cholecystectomy (base case; probabilistic sensitivity analysis – 1000 iterations).

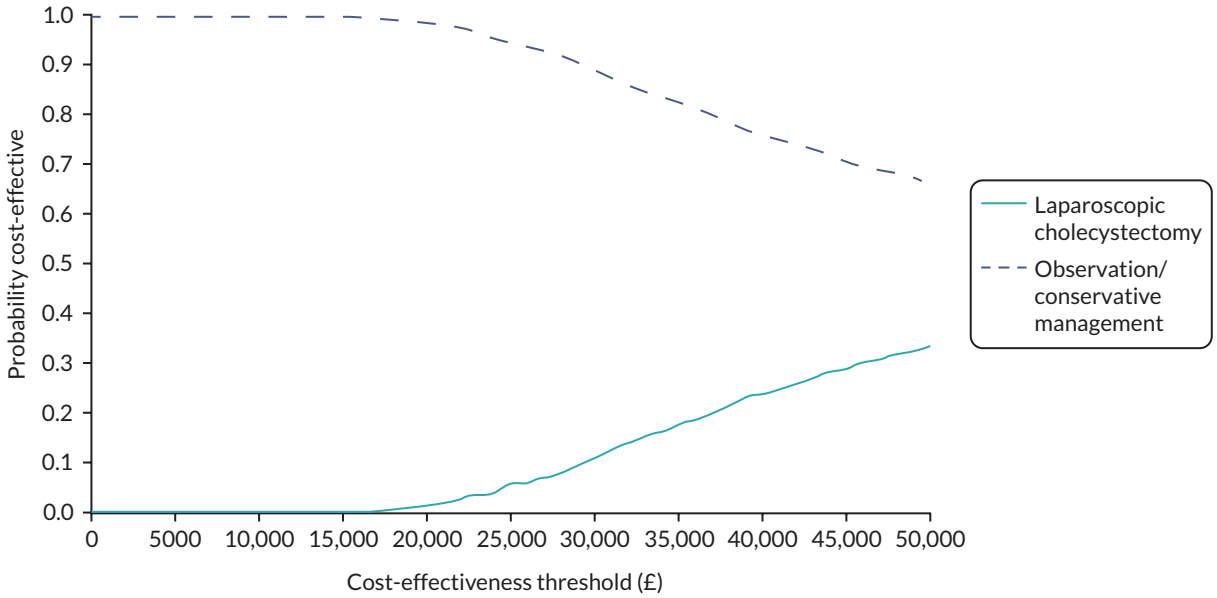

FIGURE 13 Cost-effectiveness acceptability curves for observation/conservative management and laparoscopic cholecystectomy groups (model-based analysis; base case; probabilistic sensitivity analysis – 1000 iterations).

TABLE 20 Selected scenario analyses

Strategy	Cost (£)	Δ cost (£)	QALYs	Δ QALYs	ICER (£/QALY)	Probability cost-effective		
						£13,000	£20,000	£30,000
Base case								
Laparoscopic cholecystectomy	3020		5.907			0.00	0.02	0.11
Observation/ conservative management	2016	−1003	5.894	−0.013	78,063	1.00	0.98	0.89
Scenario 1: Gompertz distribution for laparoscopic cholecystectomy and observation/conservative management								
Laparoscopic cholecystectomy	2871		5.907			0.00	0.00	0.04
Observation/ conservative management	1638	−1233	5.895	−0.012	101,160	1.00	1.00	0.96
Scenario 2: Log-logistic distribution for laparoscopic cholecystectomy and observation/conservative management								
Laparoscopic cholecystectomy	3135		5.906			0.00	0.03	0.13
Observation/ conservative management	2175	−960	5.893	−0.013	74,012	1.00	0.98	0.87
Scenario 3: Gompertz distribution for laparoscopic cholecystectomy and log-logistic for observation/conservative management								
Laparoscopic cholecystectomy	2871		5.907			0.00	0.00	0.04
Observation/ conservative management	2175	−697	5.893	−0.014	50,685	1.00	1.00	0.96

continued

TABLE 20 Selected scenario analyses (*continued*)

Strategy	Cost (£)	Δ cost (£)	QALYs	Δ QALYs	ICER (£/QALY)	Probability cost-effective		
						£13,000	£20,000	£30,000
Scenario 4: Day-case surgery cost assumed for all cholecystectomies								
Laparoscopic cholecystectomy	2311		5.907			0.01	0.09	0.25
Observation/ conservative management	1608	−703	5.894	−0.013	54,680	0.99	0.91	0.75
Scenario 5: Emergency cholecystectomy cost assumed for all cholecystectomies								
Laparoscopic cholecystectomy	4075		5.907			0.00	0.00	0.01
Observation/ conservative management	2624	−1451	5.894	−0.013	112,864	1.00	1.00	0.99
Scenario 6: Day-case and emergency cholecystectomy cost assumed for laparoscopic cholecystectomy and observation/conservative management groups, respectively								
Laparoscopic cholecystectomy	2311		5.907			0.91	0.90	0.89
Observation/ conservative management	2624	313	5.894	−0.013	−24,364	0.09	0.10	0.11

Δ, increment.

management with a high probability of being cost-effective. This result was expected as scenario 6 is an extreme case scenario.

Table 21 shows the results for the one-way sensitivity analysis where alternative differences in the unit cost for the surgical episode between the trial groups were defined. Differences over £1700 are needed for a drop in the expected savings per QALY lost for the £20,000 threshold. As a reference, the difference in the NHS reference cost for laparoscopic cholecystectomy episode between a day-case and an elective procedure with an overnight stay and a non-elective procedure with long stay are £886 and £2205, respectively. That is to say, a large proportion of individuals in the observation/conservative management group will need to undergo emergency cholecystectomies for the laparoscopic cholecystectomy strategy to be cost-effective.

The results for the combined one-way analysis are reported in *Table 21* and those for scenario 3 are reported in *Table 22*. Scenario 3 assumed a Gompertz survival model for the laparoscopic cholecystectomy group and log-logistic for the observation/conservative management group (see *Table 20*, scenario 3). Differences in the unit cost for the cholecystectomy episode over £800 would reduce the expected saving per QALY lost below £20,000.

Table 23 shows the results for the one-way sensitivity analysis where a higher reduction in QoL (utility decrement) due to cholecystectomy after 24 months was assumed. All ICERs remained above the usual cost-effectiveness threshold used in the UK, meaning observation/conservative management is still cost-effective for the utility decrements assumed.

TABLE 21 One-way sensitivity analysis; additional cost for surgery in the observation/conservative management group

Value (£)	Intervention	Total cost (£)	Δ cost (£)	Total QALYs	Δ QALY	ICER
0	Laparoscopic cholecystectomy	3020		5.907		
(Base case)	Observation/conservative management	2016	−1003	5.894	−0.013	78,063
500	Laparoscopic cholecystectomy	3020		5.907		
	Observation/conservative management	2247	−773	5.894	−0.013	60,138
1000	Laparoscopic cholecystectomy	3020		5.907		
	Observation/conservative management	2477	−543	5.894	−0.013	42,213
1500	Laparoscopic cholecystectomy	3020		5.907		
	Observation/conservative management	2708	−312	5.894	−0.013	24,287
1700	Laparoscopic cholecystectomy	3020		5.907		
	Observation/conservative management	2800	−220	5.894	−0.013	17,117
2000	Laparoscopic cholecystectomy	3020		5.907		
	Observation/conservative management	2938	−82	5.894	−0.013	6362
2100	Laparoscopic cholecystectomy	3020		5.907		
	Observation/conservative management	2984	−36	5.894	−0.013	2777
2200	Laparoscopic cholecystectomy	3020		5.907		
	Observation/conservative management	3030	10	5.894	−0.013	−808

Δ, increment.

The results for the last two scenarios are reported in *Table 24*. For scenario 10, the results from the C-GALL within-trial base-case analysis in *Chapter 4* were incorporated in the model. The cost difference between the strategies reduces and the QALY difference increases slightly, resulting in a sharp reduction in the ICER compared with the base-case analysis (see *Table 19*). However, the ICER is still above the £20,000 cost-effectiveness threshold and the probability of observation/conservative management being cost-effective remains high.

In scenario 11, the adjusted SF-6D utility scores from the analysis of the C-GALL trial data were used for the initial 48 months of the 10-year model time horizon. The SF-6D scores show a significant difference in QoL at 30 and 36 months post randomisation between the trial groups, and non-significant utility scores favouring laparoscopic cholecystectomy for all but three and 48 months time points. As anticipated, there are no changes in costs or cost difference with respect to the base-case analysis as only input utility scores were varied. Expected QALYs are higher for the laparoscopic cholecystectomy group and slightly lower for the observation/conservative group, dropping the ICER below the £20,000 threshold. Laparoscopic cholecystectomy has 83% probability of being cost-effective at £20,000 cost-effectiveness threshold for this final scenario.

Discussion

The results from the 10-year extrapolation modelling exercise reported in this chapter suggest that laparoscopic cholecystectomy is more costly than observation/conservative management. Only under extreme assumptions, such as assuming all surgeries conducted as emergency procedures in the observation/conservative group and as day-cases for the laparoscopic cholecystectomy group, was the

TABLE 22 One-way sensitivity analysis; additional cost for surgery in the observation/conservative management group and scenario 3 for survival analysis[a]

Value (£)	Intervention	Total cost (£)	Δ cost (£)	Total QALYs	Δ QALY	ICER
0	Laparoscopic cholecystectomy	2871		5.907		
(as scenario 3)	Observation/conservative management	2175	−697	5.893	−0.014	50,685
500	Laparoscopic cholecystectomy	2871		5.907		
	Observation/conservative management	2429	−442	5.893	−0.014	32,169
600	Laparoscopic cholecystectomy	2871		5.907		
	Observation/conservative management	2480	−391	5.893	−0.014	28,466
700	Laparoscopic cholecystectomy	2871		5.907		
	Observation/conservative management	2531	−340	5.893	−0.014	24,763
800	Laparoscopic cholecystectomy	2871		5.907		
	Observation/conservative management	2582	−289	5.893	−0.014	21,060
900	Laparoscopic cholecystectomy	2871		5.907		
	Observation/conservative management	2633	−239	5.893	−0.014	17,357
1000	Laparoscopic cholecystectomy	2871		5.907		
	Observation/conservative management	2683	−188	5.893	5.893	13,653
1500	Laparoscopic cholecystectomy	2871		5.907		
	Observation/conservative management	2938	67	5.893	−0.014	−4862
2000	Laparoscopic cholecystectomy	2871		5.907		
	Observation/conservative management	3192	321	5.893	−0.014	−23,378

a Gompertz distribution for laparoscopic cholecystectomy and log-logistic for observation/conservative management
Δ, increment.

cost difference favouring observation/conservative management reversed. However, higher resource used should be related to individuals presenting with more severe clinical cases. Clinical cases at presentation may be more severe in the observation/conservative management group once individuals have waited several months without surgery. Moreover, a small but consistent QALY difference favouring laparoscopic cholecystectomy was observed for the base case and all sensitivity analyses. These cost and QALY differences translated into an ICER of £78,063 for the base case, meaning that important savings could be obtained for a relatively small QALY loss. A final scenario analysis used SF-6D-adjusted utility scores up to 48 months post randomisation. This produced a QALY difference large enough to reverse the base-case results, reducing the ICER to £14,698.

A strength of the modelling exercise is that it is supported by randomised data collected in a prospective, large multicentre and pragmatic RCT. Survival analysis was used to extrapolate time from randomisation to surgical procedures and time to event and cost data were used to estimate rate and costs associated with presurgical AEs. In addition, the reduction in QoL for individuals undergoing cholecystectomy was estimated using SF-6D scores collected in the trial covering a period of 6 months before and after cholecystectomy.

There are a number of limitations in our analysis. The estimated survival curve data were directly incorporated into the decision modelling software in a deterministic way and therefore there was no parameter uncertainty associated with the survival analysis in the PSA. As such, the probabilistic analysis

TABLE 23 One-way sensitivity analysis; utility reduction when going through surgery

Value (%)	Intervention	Total cost (£)	Δ cost (£)	Total QALYs	Δ QALY	ICER
20.0	Laparoscopic cholecystectomy	3020		5.904		
	Observation/conservative management	2016	−1003	5.890	−0.015	67,418
15.0	Laparoscopic cholecystectomy	3020		5.905		
	Observation/conservative management	2016	−1003	5.891	−0.014	70,745
10.0	Laparoscopic cholecystectomy	3020		5.906		
	Observation/conservative management	2016	−1003	5.892	−0.013	74,417
5.5	Laparoscopic cholecystectomy	3020		5.907		
(Base case)	Observation/conservative management	2016	−1003	5.894	−0.013	78,063
5.0	Laparoscopic cholecystectomy	3020		5.907		
	Observation/conservative management	2016	−1003	5.894	−0.013	78,491
0.0	Laparoscopic cholecystectomy	3020		5.908		
	Observation/conservative management	2016	−1003	5.895	−0.012	83,037

Δ, increment.

TABLE 24 Further scenarios using C-GALL data

Strategy	Cost (£)	Δ cost (£)	QALYs	Δ QALYs	ICER (£/QALY)	Probability cost-effective		
						£13,000	£20,000	£30,000
Scenario 10: C-GALL within-trial analysis results for the first 24 months								
Laparoscopic cholecystectomy	2919		5.926			0.15	0.28	0.44
Observation/conservative management	2335	584	5.907	0.019	30,000	0.85	0.72	0.56
Scenario 11: C-GALL adjusted utilities up to 48 months								
Laparoscopic cholecystectomy	3020		5.914			0.37	0.83	0.98
Observation/conservative management	2016	1003	5.846	0.068	14,698	0.63	0.17	0.03

Δ, increment.

might underestimate the overall parameter uncertainty in the model. Nevertheless, alternative survival functions were defined to cover the range of extreme but plausible scenarios with the base-case results being robust to the different survival distributions used. Furthermore, there was no difference in the utility score attached to the *Symptoms resolved* and *Symptoms persist* Markov health states after surgery. As the proportion of cholecystectomies are higher in the cholecystectomy group, a reduced quality of life due to symptoms after surgery would make cholecystectomy less cost-effective and therefore our assumtion is conservative. However, the long-term quality-of-life data are needed to assess if this assumption is valid beyond the 24-month trial follow-up.

In summary, the cost–utility analysis based on the C-GALL trial participant data suggests that observational/conservative management might be cost-effective at 24-month follow-up. However, the extrapolation over 10 years introduced uncertainty over this short-term result and could in some scenarios reverse the findings. The current decision uncertainty could be reduced with a long-term follow-up of the C-GALL trial participants.

Chapter 6 Embedded process evaluation

Introduction

The C-GALL trial evaluated very different interventions for the treatment of gallstone disease: laparoscopic cholecystectomy or observation/conservative management. Based on evidence from other surgical versus non-surgical intervention trials, it was anticipated that the trial would face a number of challenges, particularly around informed consent and recruitment, from both the perspective of patients and recruiting clinicians.[69–71]

An embedded process evaluation was incorporated at the design stage of the trial to identify challenges relating to trial design and or conduct that could be addressed and modified. The process evaluation component was designed to be responsive to the 'needs' of the trial, allowing flexibility to address core problems in relation to recruitment and/or retention in a timely manner. The overall structure and methods of data collection within the process evaluation were guided by the QuinteT Recruitment Intervention (QRI) and lent heavily on participant flow data, audio recordings of trial discussions and interviews with trial participants.[72] However, a significant change to the traditional QRI method was the application of a behavioural science approach to understand problems of trial recruitment and retention. Clinical trials depend on behaviours: they rely on people (patients, clinicians, trial staff) performing actions (such as receiving or delivering a trial intervention, attending a clinic, returning a questionnaire or approaching eligible participants) that they would not do otherwise. Emerging evidence suggests behavioural science has the potential to add value with regard to improving the conduct of trials.[73]

Within this process evaluation we applied the theoretical domains framework (TDF) as a method to help inform data collection and analysis. The TDF is an established framework that integrates 33 theories of behaviour into 14 domains that inhibit or enable behaviour (knowledge, skills, social/professional role and identity, beliefs about capabilities, beliefs about consequences, optimism, reinforcement, intentions, goals, memory/attention/decision processes, environmental context and resources, social influences, emotion and behavioural regulation).[74] A handful of projects interested in trial recruitment and retention have now successfully applied the TDF as a tool to detail the challenges, from the perspectives of both patients and healthcare professionals, through a behavioural lens.[75–78] The TDF was chosen as the framework of choice for components of this process evaluation, as it also lends itself to evidence-based methods for intervention development through mapping onto behaviour change techniques (BCTs) through established methods.[79] BCTs are defined as the smallest active ingredient of an intervention such as feedback on behaviour or action planning, and they can be used alone or in combination with other BCTs.[80] This process evaluation applied these methods from the science of behaviour change to address problems of recruitment and retention in trial. We developed and implemented behaviour change interventions to target the challenges identified during the behavioural investigation.

The process evaluation team worked with the Co-Chief Investigators, PMG and site staff to identify and address challenges to trial recruitment and retention across two sequential phases:

- Phase 1 focused on identifying and understanding the 'problem' for trial recruitment and/or retention using multiple methods and data sources.
- Phase 2 sought to develop solutions to the 'problem' identified in Phase 1 through the development and delivery of interventions, with the wider trial team, to target the conduct challenges.

Methods

The process evaluation was approved through the main trial application from NHS North of Scotland Research Ethics Committee (16/NS/0053). Informed consent was obtained from all participants.

Phase 1: identifying and understanding challenges for trial recruitment and retention

Phase 1 focused on the collection and analysis of data from three main components: (1) participant flow data; (2) audio recordings of trial discussions; and (3) interviews with trial participants. Each of these components is described below with regard to data collection and analysis.

Component 1 – Participant flow data

Data collection

Data on the number of participants screened, eligible, approached and randomised were taken from screening logs and the trial website for all recruiting centres. Additional data were collected on: eligibility (including whether they met the protocol inclusion/exclusion criteria, or any reason they were deemed ineligible); whether eligible participants were approached about the trial (if not, why); and finally, whether participants were randomised and if not, why not, and which treatment they chose. Reasons for participants' declining participation, specifically patient or surgeon preference for surgery or observation/conservative management were elicited. Finally, the number of participants who had requested or had received the treatment to which they had not been randomised (i.e. potential and actual crossovers) were also downloaded from the trial website.

Data analysis

Alongside primary qualitative data, an in-depth analysis of participant flow at each recruiting site was conducted. Analysis of patient recruitment pathways using the Screened, Eligible, Approached, Randomised (SEAR) framework was applied to identify and assess areas of complexity and protocol compliance.[81] Simple counting of data collected in SEAR logs can provide useful information about the complexity of the recruitment process; differences between centres or over time can give indications of difficulties that can be investigated further. These data were compared across study sites to illustrate any variation between centres and identify areas of good practice that can be shared. Data were used to guide decisions about prioritisation of feedback by comparing activity across sites to identify core problem areas, for example, sites that had much lower (proportionally) approached to randomised rates, or high crossover rates, etc.

Component 2 – Audio recordings of trial discussions

Sampling and recruitment

All staff at the trial sites involved in discussing the C-GALL trial with potential participants were asked to routinely record consultations in which the C-GALL trial was discussed. To facilitate C-GALL-qualitative (C-GALL-QUAL) audio recording of recruitment discussions, a specific PIL was given to participants, at the same time but before any discussion of the trial was initiated, to explain the purpose of, and the request to, audio-record their consultation which involved discussion of the trial (herein referred to as recruitment consultations). Patients were not obliged to participate in audio recording of recruitment consultations and their decision did not affect their invitation to take part in the C-GALL trial. Following

receipt of the audio recording PIL, and after having had the opportunity to ask any questions, willing participants were asked to verbally confirm their consent to the recruitment consultations being audio-recorded. Recruitment consultations were recorded after an initial greeting and introduction to the consultation. If a participant consented, the recording would continue and there would be a record of consent. If a participant declined, the audio recording was stopped, and the file was deleted. With regard to staff consent, a nominated individual at each site (usually the PI) distributed staff information sheets about the recording process and obtained a one-off written consent from all staff involved in audio recording that covered all subsequent recordings captured throughout the study period.

For the analysis of discussions relating to trial retention, sites included in the sampling had to have trial participants who had reached the 9-month follow-up time point to be eligible for inclusion. From the seven eligible sites, we selected sites based on a positive (i.e. high postal questionnaire response rates) and negative (i.e. low postal questionnaire response rates) deviant approach to provide a variety of discussions from sites with varying retention patterns. We aimed to analyse the 10 most recent (assuming that analysing most recent practices would be required if intending to implement a change based on findings) consultations where available.

Data collection

Efforts were made to audio-record recruitment consultations across sites throughout the duration of the trial (September 2016–November 2018) for those participants who consented to the audio recording. Sites were provided with devices for recording conversations and were asked to upload audio files to a secure trial webpage following data capture. The audio files were then deleted from the devices. All conversations within the recordings related to C-GALL trial (where recruiters explain the design and details of the C-GALL trial, and patients decide whether or not to take part), as determined by the researcher, were transcribed for the purpose of analysis that is, targeted transcription. At least 10 consultations per site were required before site-specific analysis was conducted.

Data analysis

All recordings (from both recruitment consultations and interviews) were transcribed verbatim by a professional transcription service, anonymised and labelled with a unique identifier, to ensure confidentiality. Data were initially analysed deductively against a coding framework informed by existing research using audio recordings to improve recruitment processes in RCTs.[72] Coding categories included how the trial was introduced, whether the trial or treatment was introduced first, balance of options/risks, patients' preference, discussion of randomisation, discussion of uncertainty, discussion of crossover and discussion relevant for retention such as in relation to questionnaires or withdrawal. Initial coding was conducted by one researcher with 25% checked by another member of the team and any discrepancies discussed with an arbiter. In addition, analysis took the form of constant comparison alongside case study methods, both within and across sites, to determine problem areas or identify aspects of good practice.

Various information sources were drawn upon when deciding which BCTs to incorporate into the e-mail feedback delivered to sites. Evidence from existing studies that had explored healthcare professionals' actions reporting challenges to recruitment (in surgical trials) was reviewed to consider how challenges should be mapped against the BCT taxonomy.[82,83] Feasibility, with regard to what could be delivered in the mode we had available (i.e. e-mail), was a consideration, alongside input from Co-Chief Investigators. Finally, we requested guidance from the first two sites to receive the feedback regarding positive versus negative framing of messages.

In addition to the more general discussions about trial participation, we were also interested in how trial retention was discussed during consultations. Therefore, in addition to the content analysis described above, qualitative data from the audio recordings were also transformed into quantitative data using the quanti-qualitative appointment timing approach (Q-QAT)[84] to quantify time spent on discussions of retention during the trial consultation.

Component 3 – Interviews with trial participants during the trial

Sampling and recruitment

A purposive sample of trial participants who had not returned at least one questionnaire at any time point during the study, and therefore had discontinued their follow-up at least once, were invited to interview. An invitation letter and a qualitative interview PIL were distributed to all potential participants with a reply slip and a prepaid envelope. A researcher then contacted participants who responded to confirm they were willing to be contacted to discuss the qualitative interview study further and book a mutually convenient time for a telephone interview. All invitation packs were distributed from the C-GALL trial office to maintain the confidentiality of potential participants. Two attempts were made to engage with potential participants. A total of 279 invitations were issued.

Patients approached for the C-GALL trial (both consenters and non-consenters) were also invited to interviews to explore outcomes of relevance to patients to inform the development of a COS. This work is reported in *Chapter 7*.

Data collection

The topic guide was informed by the TDF and refined by the research team. The topic guide was iteratively updated to ensure robustness. Broadly, the topic guide explored participants' reasons for not returning questionnaires and the behavioural barriers and facilitators to returning their questionnaires. Telephone interviews were conducted by two members of the research team between January and November 2019, with interviews lasting between 25 minutes and 1 hour 28 minutes (median = 36 minutes). Verbal informed consent was sought from all participants. All interviews were audio-recorded and transcribed verbatim. Sufficiency of our sample size was judged against five key aspects of information power: whether the study aim was broad or narrow (with a focused aim requiring a smaller sample); dense or sparse sample specificity (where dense specificity requires a smaller sample); application or not of established theory (theoretical perspectives requiring smaller sample); quality of the dialogue (with strong clear communication requiring less); and finally, whether case or cross-case analysis (with cross-case requiring more participants).[85]

Data analysis

A TDF coding guide specifying each of the domains, relevant constructs and example quotes relevant to said domains, was developed to facilitate coding. Three interview transcripts were coded independently using the draft coding guide prior to comparing the coding results, with discrepancies resolved through discussion. One researcher coded all participants' interview responses into the relevant theoretical domains. After coding data into theoretical domains, belief statements were generated which represented similar underlying beliefs held by participants within each domain.[79] We used the recommended procedure for TDF analysis to identify the relevant domains that were most likely to influence the behaviour:[79] (1) the frequency of belief statements across all domains; (2) the presence and prevalence of conflicting beliefs; and (3) evidence of strong beliefs that influence the behaviour. All three criteria were considered concurrently to judge the relevance of each domain. We developed overarching themes that described the content of related belief statements and domains to effectively summarise the findings. Overarching themes were initially generated by one member of the research team. Themes and belief statements were refined by two other members to ensure that they summarised the data accurately.

Following identification of the TDF domains relevant for PQ return, existing patient-facing documents relevant for retention were analysed for inclusion of BCTs. Existing questionnaire cover letters and newsletters were coded against the BCT Taxonomy to identify pre-existing BCTs and compared to templates of previously developed BCT-informed cover letters to promote questionnaire response.[86] These BCTs were then compared against the relevant TDF domains identified in the interviews to ensure all relevant domains were being targeted by pre-existing BCTs. If a TDF domain was not covered, relevant BCTs were added into the text of the letter/newsletter, also, if a relevant domain required enhancing the BCTs were dosed up by inclusion of additional relevant techniques. This process was conducted by one researcher and checked by two others with disagreements resolved through discussion.

Phase 2: development and delivery of interventions to target recruitment and retention challenges

The findings from Phase 1 were developed into strategies to target the recruitment and retention challenges identified. As per the QRI method for generating trial-focused solutions, data generated in Phase 1 was discussed with the PMG and Co-Chief Investigators to directly inform plans.[72] Strategies developed as a direct result of the analysis of the participant flow and audio-recorded consultations were at a site level, with (depending on the data collected) opportunities for generic and/or site-personalised feedback. Feedback had originally been planned as face-to-face meetings (which did happen for the one site); however, due to logistical challenges this feedback was moved to e-mail. Face-to-face generic feedback did occur at regional meetings for investigators. Other mechanisms for feedback included a 'top tips' recruitment sheet, which was also shared via e-mail. No identifiers of individuals or clinical centres were included in feedback. Findings from the interviews informed the development of strategies to target return of trial questionnaires. These strategies included adaptations to the cover letters accompanying the questionnaire and updating newsletters. Data from the audio recording analysis of discussions of retention also directly informed the trial-specific PIL provided to potential participants when making a decision about participation.

Findings

Phase 1: identifying and understanding challenges for trial recruitment and retention

Recruitment

Twenty-one of the twenty-two C-GALL sites were included in the participant flow-data analysis. One site was not included in this analysis due to inactivity. Of these twenty-one, three had very low recruitment numbers due to them becoming active recruiting centres late in the trial, and one site was closed down due to poor recruitment numbers. As such, these four sites were not progressed for analysis. All further data presented are based on 17 sites. Diagnostic data from the participant flow revealed that for most sites (n = 13; 76%) the main challenge for recruitment existed in converting approached patients to randomisations. The biggest barrier to randomisation was patient preference, largely for surgery (*Table 25*). A total of 16 sites provided 180 audio recordings for analysis (median number of recordings per site = 10). A further five sites did not audio-record consultations, due to sites not agreeing or forgetting to record. Consultations included consultant surgeons, RNs and patients who were sometimes accompanied by family members (partner or child). Analysis of the transcripts identified four core challenge areas for recruiters discussing the trial:

- providing a balanced presentation about both treatments
- discussing and exploring preferences
- discussing uncertainty
- discussing crossovers.

Examples of excerpts from the audio consultations are provided in *Table 25*.

Retention

Audio recordings of trial discussions

A subset of 38 audio recordings from 4 sites were included in the analysis of discussion of retention with 3 sites providing 10 recordings each and the fourth site providing 8. The findings from this aspect of the process evaluation have been published in full elsewhere.[68] In brief, most of these consultations (n = 30/38; 79%) did not include any discussion of trial retention. Where retention was discussed, the median proportion of consultation time in which retention was discussed was 3.8%, with the longest discussion of retention lasting 89 seconds and shortest lasting 20 seconds. Presence of discussion of retention also varied by site, with one of the four sites never discussing retention in any of the

TABLE 25 Summary data by site regarding participant flow, examples quotes, targets for feedback and level of feedback

Site ID	'Problem' identified from participant flow data	Audio	Example quotes[a]	Targets for feedback	Feedback
1	Crossovers	Y	Patient: *'It is the surgery I am thinking I'd rather have'* Site staff: *'We can see what randomisation says and then we can take it from there'*	Positive: site is the second top recruiting site for the whole of the C-GALL trial Negative: site is also the site with the highest rate of patients crossing over	Site specific
2	Approached to randomised patients who are declining trial participation by expressing a preference for surgery	Y	Not available at time of feedback	Positive: are discussing the trial with 97% of eligible patients Negative: has the second lowest proportion of approached patients being randomised	Generic
3	Approached to randomised patients who are declining trial participation by expressing a preference for surgery	Y	Site staff: *'The trial's only there if you're interested. If you feel that you have a preference, then that's absolutely fine'* Patient: *'I think for my circumstances I think I would opt to just have it taken out and hopefully it would deal with everything'.* Site staff: *'Okay'*	Positive: approaching 80% of eligible patients to talk to them about C-GALL Negative: only 20% of trial discussions go on to lead to randomised participants Why: majority have a preference for surgery	Site specific
4	Crossovers	Y	*'Let's do the randomisation right? And see what you are allocated. You might be allocated surgery and that's happy days. If you are not allocated surgery, you can tell us immediately to swap you over'* *'So obviously let's see … what you are randomised into. If you are randomised to surgery, that's the answer. If you are not randomised to surgery, and then we can have a discussion, and then you are allowed to cross over at any time'* *'You are not stuck. You can immediately within the same minute say, "no I want surgery" and we'll swap you over'* *'You can cross over to the surgical arm without any problems, or even giving an explanation. This is a completely flexible trial'* *'If you are randomised to no surgery and you want surgery, you can cross over. You are in the driving seat; we are just trying to observe and learn from your disease process'*	Self-identified the problem of crossovers	Site specific

TABLE 25 Summary data by site regarding participant flow, examples quotes, targets for feedback and level of feedback (*continued*)

Site ID	'Problem' identified from participant flow data	Audio	Example quotes[a]	Targets for feedback	Feedback
5	Approached to randomised patients who are declining trial participation by expressing a preference for surgery	Y	Site staff: *'What's your initial thoughts on that?'* Patient: *'Yeah, that's okay'*	Positive: approaching 97% of eligible patients to talk to them about C-GALL	Generic (No audio uploaded at time of feedback. Quote provided as example content of site discussions retrieved post feedback)
		Y	Site staff: *'Do you want to go for surgery? Are you wanting to see how things go?'* Patient: *'Just happy'* Site staff: *'Are you interested in the study or what do you think?'* Patient: *'Just happy to get it gone'* Site staff: *'You just want an operation?'* Patient: *'Yeah, just get it gone'*	Negative: site had only 22% of participants go on to be randomised Why: 63% of those who decline to take part have a preference for surgery	
6	Approached to randomised patients who are declining trial participation by expressing a preference for surgery	N		Positive: upwards of 85% of eligible patients talked about C-GALL Negative: only 19% of participants go on to be randomised Why: 61% of those who decline to take part had a preference for surgery	Generic
7	Eligible to approached and crossovers	Y	*'And if you're in the study what we do is we allocate you to one of those two arms in the study, and one is that we do the gallbladder operation, and the other arm is we do what's called watchful waiting, and either way we keep a very close eye on you and see what happens. And if we need to take your gallbladder out during the study, if you're allocated to watchful waiting for instance and then you suddenly came in with a dreadful attack of gallbladder problems then obviously, we do what's needed clinically to take your gallbladder out. But maybe from the circumstances you don't get any further attacks, so that's the reason it's difficult to advise you on what we do'*	Positive: randomising 40% of patients approached about the C-GALL trial (average) Negative: only 71% of participants being told about the C-GALL trial Negative: site has the third largest number of patients intending to crossover	Generic (No audio uploaded at time of feedback. Quote provided as example content of site discussions retrieved post feedback)

continued

TABLE 25 Summary data by site regarding participant flow, examples quotes, targets for feedback and level of feedback (*continued*)

Site ID	'Problem' identified from participant flow data	Audio	Example quotes[a]	Targets for feedback	Feedback
8	Approached to randomised patients who are declining trial participation by expressing a preference for surgery	N		Positive: upwards of 80% of eligible patients talked about C-GALL Negative: only 25% of participants go on to be randomised Why: 42% of those who decline to take part had a preference for surgery	Generic
9	Approached to randomised patients who are declining trial participation by expressing a preference for surgery	Y	*'I think that having the operation is the best thing to do'* *'if you don't ever want the pain again you could have the operation'* *'you may go for surgery and be no better off, I suspect you'll be better off'*	Positive: discussing the trial with 100% of eligible patients Negative: only 29% of trial discussions lead to randomised participants Why: 73% of those who decline to take part had a preference for surgery. Surgery is consistently mentioned to be the best option	Site specific
10	Approached to randomised patients who are declining trial participation by expressing a preference for surgery	Y	*'We know that obviously if we do the operation, it would be done by keyhole, 95% of the time. Good chance of helping your symptoms, but we do know that after having had your gallbladder out, up to a third of patients still have ongoing pain… Obviously if you go down the line of being observed, and things become worse, then we would see you again to look at taking your gallbladder out, should that be required. Obviously if you got put down the line of having your gallbladder out, if you changed your mind, then that's absolutely fine. You are under no obligation. Would you be happy to be part of the study?'*	Positive: upwards of 87% of eligible patients Negative: 43% of trial discussions lead to randomised participants Why: 49% of those who decline to take part had a preference for surgery	Generic (No audio uploaded at time of feedback. Quote provided as example content of site discussions retrieved post feedback)
11	Approached to randomised patients who are declining trial participation by expressing a preference for surgery	Y	*'We find the gallbladder – this is all under general anaesthetic, so you are asleep – find the gallbladder, free it up, and we take the gallbladder with the gallstone out, because you can't leave any of that behind because you are at the risk of forming more. That's the best way to stop any further problems (…) There's a risk of damage to upper abdominal structures – bowel, blood vessels, liver, in that area (…) if you go onto the non-operative arm, you'll take painkillers if you need it, or antibiotics, or whatever, and then that will be part of your questionnaire when it gets sent out'*	Positive: approaching 93% Negative: 47% of trial discussions lead to randomised participants Why: 87% patient preference for surgery	Site specific

TABLE 25 Summary data by site regarding participant flow, examples quotes, targets for feedback and level of feedback (*continued*)

Site ID	'Problem' identified from participant flow data	Audio	Example quotes[a]	Targets for feedback	Feedback
12	Approached to randomised patients who are declining trial participation by expressing a preference for surgery	Y	Site staff: *'So, you'd rather have it removed?'* Patient: *'I'd rather have it out, yeah'* Site staff: *'Yeah, okay. Well, fair enough'* Patient: *'Is there no other way for the gallbladder if you have stones? Is it just surgery or that's it?'* Site staff: *'Yes'*	Positive: 90% approached Negative: 43% led to randomisations Why: 42% patient preference for surgery	Site specific
13	Approached to randomised patients who are declining trial participation by expressing a preference for surgery	Y	Not available at time of feedback	Positive: 97% of eligible approached Negative: only 40% of participants approached were then randomised Why: 90% of patient preference for surgery	Generic
14	Eligible to approached	Y	Not available at time of feedback	Positive: has highest percentage of participants approached randomised Negative: lowest amount initially approached, high percentage of patient preference for surgery	Generic
15	Approached to randomised patients who are declining trial participation by expressing a preference for surgery	Y	*'The attitude has always been, we do an operation we remove your gallbladder, and you get better when we look back on people, that 20 years ago that's what we did, we can help 80% of them did, but up to 20% of people did not feel better (…) We can consider an operation … that is option 1, option 2, we do know that increasingly that some people get better by losing weight, having healthier diet (…) we are also looking at doing research'*	Positive: 53% of eligible patients approached Negative: 12% of those trial discussions lead to randomised participants Why: 52% of those who decline to take part had a preference for surgery	Site specific
16	Approached to randomised	Y	*'So, if you go down the surgical route, then obviously the advantages – you get rid of the stones. Not only are you taking the gallstones away, we take the gallbladder as well, so they don't come back, and it gets rid of the source of the pain for good'*	Negative: only 39% of participants approached were randomised	Generic (No audio uploaded at time of feedback. Quote provided as example)

continued

TABLE 25 Summary data by site regarding participant flow, examples quotes, targets for feedback and level of feedback (*continued*)

Site ID	'Problem' identified from participant flow data	Audio	Example quotes[a]	Targets for feedback	Feedback
17	No immediate issues	Y	Not available at time of feedback	One of the highest percentage of participants approached to randomised	Generic
18	Small overall numbers and no immediate issues	N			Generic
19	Small overall numbers and no immediate issues	Y	Not available at time of feedback		Generic
20	Small overall numbers and no Rimmediate issues	N			Generic
21	Site was in process of closing	N			Generic

a All quotes are from site staff unless otherwise specified.

consultations included in the analysis. Where consultations did discuss retention, some contained inaccuracies (*n* = 1, 9%), some failed to detail the frequency of follow-up questionnaires (*n* = 1, 9%) and the majority focused on the participant's right to withdraw (*n* = 6, 54%). In some instances, the patient initiated the conversation around retention, with respect to frequency of follow-up. Examples of excerpts from the consultations relevant to trial retention are presented in *Table 26*.

Interviews with trial participants during the trial

Nine interviews were conducted with trial participants who had not returned at least one questionnaire during their participation in the C-GALL trial. The median age of these participants was 53 years old (range: 35–75 years) and the majority (*n* = 7/9; 78%) identified as female. This was comparable to the entire population of C-GALL trial non-responders at the point of data analysis (February 2020; 79% female with a median age of 51 years). The length of interviews ranged from 25 minutes to 1 hour 28 minutes (median = 36 minutes).

Domains and associated belief statements were categorised into overarching themes to summarise the main messages arising from the data (*Figure 14*). The specific beliefs, the TDF domains they are associated with, and the sample quotes are presented in *Table 27*. The relationship between

TABLE 26 Example quotes relevant to trial retention

Theme	Quote
Timing and purpose of questionnaires	*'So, with the quality of life, I mentioned there was a baseline – and with those forms you'll get sent one 3 months, 9, 12, and 18 …'. 'They will be very similar to the baseline ones if you want to go into the study. And the study centre will send those out' RN Site B*
	Surgeon: *'Umm I mean you can always say uhh uhh that ah regardless of which treatment group you do – the study allocates to you, for example if it says observation then uh you will uh will require observation in about 3 months' time. Is that right?' (Doctor asking RN)*
	RN: *'Umm yeah, we just send you the questionnaire by email or by phone – so you will just answer them every couple of months. That's all we will do. There is nothing where you physically have to come or need to be examined or anything like that' Site C*
	'Emm what we do is that we send you a questionnaire in the post. 3, 6, 9, 12, and 18 months. So, it's every 3 months you get questions in the post. Kind of day-to-day thing about how your pain is and things and such'.
	'It does not involve another visit back for us. Emm, If you do get randomised to surgery – emm, it's a kind of a standard way to check for that. And then we also send you the questionnaires in the post as well' Surgeon Site D
Participant's right to withdraw	*'Like Mr X said, it is a randomised controlled study – so if somebody agrees to go into the study there is some baseline information that we do when they consent. So there is a consent process that I'd go through. Umm a consent form that you fill in. And you can withdraw your consent at any time. So if you fill in your consent from today, decided you wanted to do it – and then went home, thought about it – and thought, actually no this isn't right for me; you can withdraw at any time. Umm and again with the follow ups – so the follow up questionnaires you can always decide that when life gets busy and get in the way – if you decide actually, oh I've not got time to do these questionnaires, or too much other things going on – again you can choose to withdraw from the follow up as well. So if you decide to go into the study it is not set in stone. You can withdraw at any time alright?' RN Site B*
	'With the follow-up, it's the same for both arms for the questionnaires. I will show you a copy so you know what is expected of you. This is the baseline. But it is the same for all of them. There is no massive essays. It's a tick box – and I don't know if you want to have a look through that. Umm it's all about the general health, following activities of daily living, discussing pain …' RN Site B
	'That's pretty good, perfect! And the questionnaire don't forget the questionnaires. In a way that is kind of confirming your ongoing consent' Surgeon Site B

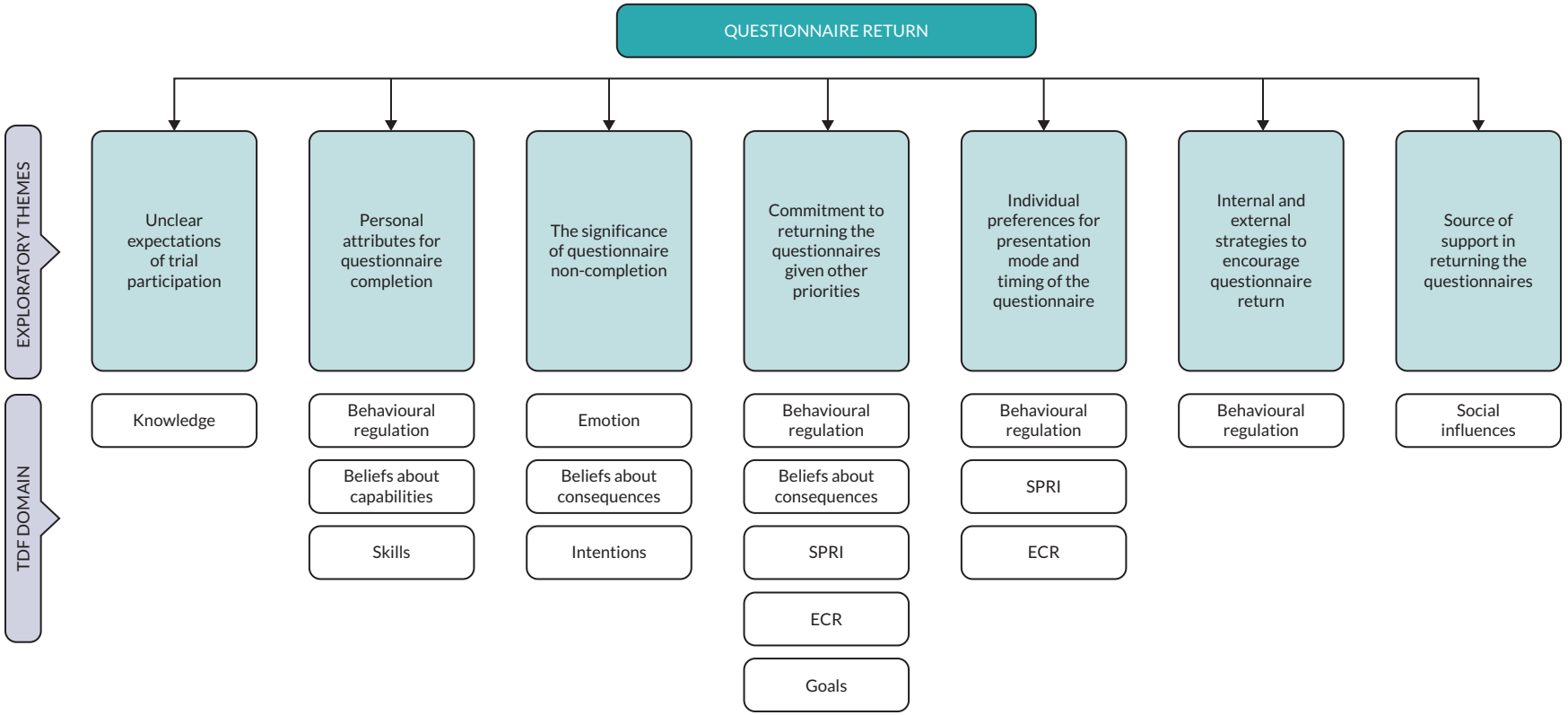

FIGURE 14 Overarching themes and associated domains. ECR, environmental context and resources; SPRI, social professional role and identity; TDF, theoretical domains framework.

TABLE 27 Table of findings with theoretical domain framework frequencies and belief statements

Domain (frequency = participants who mention domain/number of total participants)	Specific belief	Sample quote
Environmental context and resources (9/9)	The format, layout and postal administration of the questionnaire was (not) convenient for me	*'That's it, yeah. It comes with an envelope to just pop it straight back in the post. You don't have to put a stamp on or write the envelope, so I don't see any issues with it'* Participant 1 *'I much preferred it when I could give oral answers to a simplified version in which she said, "These are just the main questions that we need to know". That didn't take nearly as long, and it was quite focused and straightforward'* Participant 2
	It was (not) easy to schedule completion and return of the questionnaires with other priorities in my life	*'Well, as I say, I put it aside and then never got to it. So, clearly life got in the way for me. Yes, yes'* Participant 2 *'No, not at all. It would never be a problem to sit down for ten minutes and answer the questions'* Participant 1
	Completing the questionnaire was (not) enjoyable and/or straightforward	R: *'And any struggles to do with any parts of the study at all so far?'* *'No. Actually, in the questionnaire I asked about my health and this, that and the other, it was quite enlightening to thought about things'* Participant 7
		'Tick box things. I think, probably … Sometimes, I ticked a box and I'm thinking, "Well, I'm ticking it, but maybe I need to explain a bit more". Probably I'm ticking it because I was not on one side … . You know what I mean?' Participant 5 *'There was a lot of, "Hmm, how I do I answer this one? I'm not quite sure"'* Participant 9
Social influences (9/9)	I did not feel pressurised to complete and return the questionnaires	R: *'And did anyone support you to return the questionnaire?'* P: *'No, I mean, my partner's here but he doesn't really take any notice of my post anyway. No, I've got no pressures from anyone. Nobody encourages me either. At first, my sister half encouraged me to take part because I was about 90% sure I was going to do it, and then after talking to her and a couple of others, just things that were said I just said, Yeah, I'll do it'* (also coded under belief statement, support from family members)
	I (did not) receive support from family members to complete and return the questionnaires	*'Once I've filled it in, my husband supports me in terms of he gets it to you, as in he posts it for me because I don't go out to post it when I've got work or whatever. But that's the only support, (inaudible) support, my husband is understanding. He understands the pain and he helps me when I'm suffering and stuff, but no, I don't have any other support, not really'* Participant 6
		R: *'And does anyone support you to return the questionnaire?'* P: *'No, like I say, I just do it straightaway'* R: *'And it's all off your own back, you're able to post it yourself and such'* P: *'Yeah'* Participant 8

continued

TABLE 27 Table of findings with theoretical domain framework frequencies and belief statements (*continued*)

Domain (frequency = participants who mention domain/number of total participants)	Specific belief	Sample quote
Beliefs about consequences (9/9)	Completing and returning the questionnaires will help people with gallstones in the future	P: *'Just to help other people further down the line, really. If the questionnaire helps the study, it can obviously benefit other people further down the line'* Participant 8
	The questions were not relevant to my circumstances	*'I think filling it in would have been medical management, and I've not had any medical management. It was a bit like, yeah, I've not had anything to help with it so I was just filling it in in terms of the first time'* Participant 8. *'Well, there were a few. I mean, I started off well and then because nothing had happened … I hadn't had any pain or anything and then I'd put them aside and then somebody would send another one and say, "You haven't completed it"'* Participant 2
	It's important to fill in the questionnaires for the research team	*'No, like I said, just the information for yourselves to go towards the study and get the information that you need, really. That's the only thing that pushed me to do it if that makes sense'* Participant 8
Reinforcement (7/9)	(No) incentives or rewards would have encouraged me to complete and return the questionnaires	*'I suppose so. Obviously if it's going to benefit and you're going to get something back, it would be nice …'* Participant 8 *'Yeah, I just want to do it because I want to do it. I don't really want anybody to try and bribe me to do it or encourage me to do it due to any kind of incentive, no. In some ways, actually that would put me off'* Participant 1
Behavioural regulation (8/9)	It would (not) have been better to receive the questionnaire in a different format	*'Yeah, I suppose that's possible. Over the phone would be fine; face to face wouldn't be ideal because I don't live that close to [study centre]. Over the phone would be all right, I wouldn't mind that, but in some ways that would be more difficult to fit into your day as well. Filling it out on a piece of paper, you can do that when you've got a few minutes to get it done and from my point of view, I'd probably find that easier to fit in'* Participant 1 *'Initially. I mean, even if they were sort of three, six months apart, it would have been something. I find it easier to talk to a person probably than have to fill in these forms'* Participant 5
	I (do not) make a plan to help me complete and return the questionnaires	*'I would actually think, "Well, I've got to get this done" and do it and then send it off. Saying that, yes, probably … You're right, I would probably say, "Oh, I'm going to sit down this evening, fill this questionnaire in and then get it sent off." So, yes, in that way, there was a plan'* Participant 5 *'The questionnaire thing for me is a barrier because I'll think, "Oh, I'll put it to the side, I'll do it when I've got time," and I never get the time. I've got my second one to fill and it's been there for months. Really bad!'* Participant 6
	The time point of receiving the questionnaires influences questionnaire completion	*'I suppose if the questions had been really complicated or difficult to answer, maybe. Or if you just can't remember what exactly has happened over the time. I suppose the time between questionnaires, if it's every three months then that's not too bad, but if it was longer than that then I'd definitely have trouble remembering exactly what symptoms I'd had and how many times and all the fine details'* Participant 1 *'We're all busy and I think I'd have been more engaged with it, had I had a lot of symptoms and I was really desperately wanting to get sorted out, which I actually feel a little bit more now'* Participant 2

TABLE 27 Table of findings with theoretical domain framework frequencies and belief statements (*continued*)

Domain (frequency = participants who mention domain/number of total participants)	Specific belief	Sample quote
Beliefs about capabilities (9/9)	I was (not) confident that I could return the questionnaire(s)	R: *'How confident are you that you could return a questionnaire when requested?'* P: *'100%, I know I can'* Participant 1 *'As I say, it was the paperwork just kind of threw me and … I should have been more on top and organised with that, but'* Participant 2
Intentions (9/9)	I am (not) committed to completing and returning the questionnaires	P: *'Well, there's not a lot … It's commitment to answering the actual survey questions that took … not the survey, the questionnaires that come in. You've got to commit to that. I think you have to, because otherwise there's no point you doing it'* Participant 5 *'There's something about these forms that you just look at them … like a tax form … You just look at it and think, "Oh, I'll do it later." You know'* Participant 2
Knowledge (9/9)	The activities and tasks involved in the trial did (not) meet my expectations	*'It's the same as most questionnaires like that where you've got, I think, to give a level of how much something's inconvenienced you. Yeah, just what you'd expect from a questionnaire. Sorry, I'm not very helpful on this one!'* Participant 1 *'Well, actually the paperwork that comes through, some of the questions that we're asked were sort of non-sp-…'* (inaudible) non-specific Participant 5
	I was (un)aware of when the trial/my participation ended	*'I think at the appointment initially, they said that the trial had been extended. They didn't give me a timeframe for how long it would go on for'* Participant 1 *'Yeah, I know it's every six months that they wanted me to do a questionnaire … two is coming to mind, two plus'* Participant 6
Skills (9/9)	Literacy and communication are key skills required to complete the questionnaires	*'But actually, the actual skill involved is trying to be … I mean, hopefully, you're literate enough to be able to do it …'* Participant 5
	You need a good memory to complete the questionnaire accurately	*'It's hard, it's very hard. "During the past four weeks, to what extent has your physical health and emotional … how have they interfered with your personal life or your home life?" Things like that. You forget, do you know what I mean? You kind of forget in that four weeks what's happened. They should ask me about it there and then, so this questionnaire, "How are you feeling today?" or whatever, or last week, or "When was your pain bad?" But the last four weeks, I can't remember how it interfered with my neighbours, with my family. It is a valid question, it's just me, I'm a busy person, I don't have …'* Participant 6

continued

TABLE 27 Table of findings with theoretical domain framework frequencies and belief statements (*continued*)

Domain (frequency = participants who mention domain/number of total participants)	Specific belief	Sample quote
Emotion (9/9)	Completing and returning the questionnaire(s) (does not) trigger unpleasant emotions such as boredom	R: *'Yes, I understand. This may be a bit of an odd question but bear with me. What sorts of feelings come to mind when you think of returning a questionnaire?'* P: *'Just slightly bored'* Participant 2 R: *'What sort of feelings come to your mind when you think of returning a questionnaire? Any sort of feelings you experience?'* P: *'Not really, I think I'm quite matter-of-fact about it all'* Participant 7
	Completing the questionnaire brings me satisfaction and a sense of responsibility	*'Well, you've completed a job, you know that it's important, it helps in clinical trials, it's helping finding out stuff. That's the upside of doing it, and you feel satisfaction, that you're making a difference'* Participant 4
Social professional role and identity (9/9)	My personality and social role influence my impression of the questionnaire in terms of content and format	*'My (inaudible) thing is, honestly, I'm a social person. I like to speak to somebody, have a consultation and you will get your answers'* Participant 6 *'I'm not one for paperwork, so … it's not something I look forward to'* Participant 3

TDF domains and the overarching themes is presented in *Figure 14*. A total of seven themes were identified. These were:

1. Unclear expectations of trial participation: included beliefs about activities and tasks in the trial not meeting expectations, highlighting participants were (un)aware of when their participation in the trial ended, the relevance of the questionnaire to their circumstances (i.e. when not experiencing pain), and confusion over receiving the same questionnaire on multiple occasions.
2. Personal attributes for questionnaire completion: the need to remember levels of pain experienced was mentioned as a skill required to complete the questionnaire, as was the need to concentrate which was sometimes reported as a barrier to completion and return.
3. Significance of questionnaire non-return: participants reported feeling guilty for not returning questionnaires largely linked to them feeling they were letting the trial team down. Others felt satisfaction by completing and returning the questionnaires and saw the activity as their duty as a trial participant and that their contribution would help research.
4. Commitment to returning questionnaires given other priorities: some participants reported they did not feel committed to completing and returning the questionnaire as it was not a priority for them, reporting they were too busy. In contrast, some participants recognised the negative consequences of not returning the questionnaire (such as wasting the researchers' time) and so felt committed to completing.
5. Individual preferences for presentation mode and timing of the questionnaires: there was variability among participants regarding preferences for questionnaire format with some preferring postal and others stating they would have preferred to complete it electronically or to speak with someone.
6. Internal and external strategies to encourage questionnaire return: a range of strategies were reported that participants had used to help them complete questionnaire such as making a plan to complete and return, keeping a note of symptoms and completing the questionnaire as soon as they received it. There were mixed views about whether the use of incentives to encourage response would have been appropriate, but most participants suggested that receiving a prompt or reminder would have been/were helpful.

Sources of support in returning the questionnaires: participants reported that in some cases they received support from family members to return the questionnaires, but for the majority this was done independently. The trial office was noted as providing some level of support for all, which included reminder or completion of the questionnaire over the phone. Participants did not feel pressurised to complete and return the questionnaire, with some stating this was a barrier to completion.

Theoretical domains framework domains, associated belief statement and associated example quotes are presented in *Table 27*.

Phase 2: development and delivery of interventions to target recruitment and retention challenges

Strategies to improve trial recruitment and retention were developed, through discussion with Co-Chief Investigators and the PMG, in response to the data collected during the Phase 1 investigation. Feedback to trial sites was provided on a rolling basis, targeting sites in response to problems identified through the participants' flow diagnostic data.

Site investigator meeting

A site investigators meeting was held in Birmingham in November 2018. This meeting was attended by 17 site representatives (which included consultant surgeons and RNs) from across 10 sites. Findings from the analysis of retention discussion in the audio consultations and participant interviews were shared and discussed. Meeting attendees were encouraged to reflect on the findings and asked to consider what they could do to help improve retention, and specifically, what changes they could make during trial recruitment to raise the profile of what was expected in terms of trial retention.

E-mail feedback on recruitment activity: site specific and generic

Guided by the diagnostic participant flow data, sites were provided with e-mail feedback on their current recruitment activity in relation to other centres (SEAR). The e-mail started with a thank you and general reminder of the aim of the trial and then a positive message about an aspect of their recruitment activity, for example, approaching all eligible patients. The content of the e-mail then moved on to highlight the target behaviour that we wanted them to change (e.g. randomising more of those approached by exploring preference), and data were shared on their performance in relation to other sites (see *Figure 15*, for example text). In addition to the broad areas of content, we also included BCTs within the e-mail text. A list of potential BCTs was proposed and discussed among the immediate process evaluation team regarding relevance and feasibility of inclusion in e-mail feedback to sites to target trial recruitment. Twelve BCTs were deemed relevant and feasible for delivery and incorporated into the site specific and generic feedback (*Table 28*).

Thank you very much for your efforts on the C-GALL trial. I am really grateful for your ongoing support to achieve our recruitment targets. As you will know, achieving our recruitment targets is crucial to being able to answer the research question posed by the C-GALL trial – what is the best treatment approach for patients with uncomplicated symptomatic gallstones?

I would like to acknowledge the fantastic recruitment figures at Site X which reflects all your hard work. Site X is discussing the trial with *upwards of XX% of eligible patients*.

Although this is extremely positive, I also need to share with you that Site X is also the site that *has one of the lowest percentages of participants that go on to be randomised*. As you can see below, only XX% of trial discussions lead to randomised participants, the average across sites is XX% with some sites achieving upwards of XX%. So, although you are doing really well speaking to so many patients about the trial, this, unfortunately is not translating into a high proportion being randomised.

I see that around *XX% of those who decline to take part had a preference for surgery*. This is particularly high compared to the average of XX% across sites. During the consultation, offering a balanced explanation of both treatments and exploring patient preferences may ensure both treatments are given equal consideration. Specific suggestions of how to discuss these elements can be found in the attached materials.

The BCT-enhanced content information was included in the body of the e-mail. In addition to this main text, two attachments were provided. The first contained feedback on the content of the audio

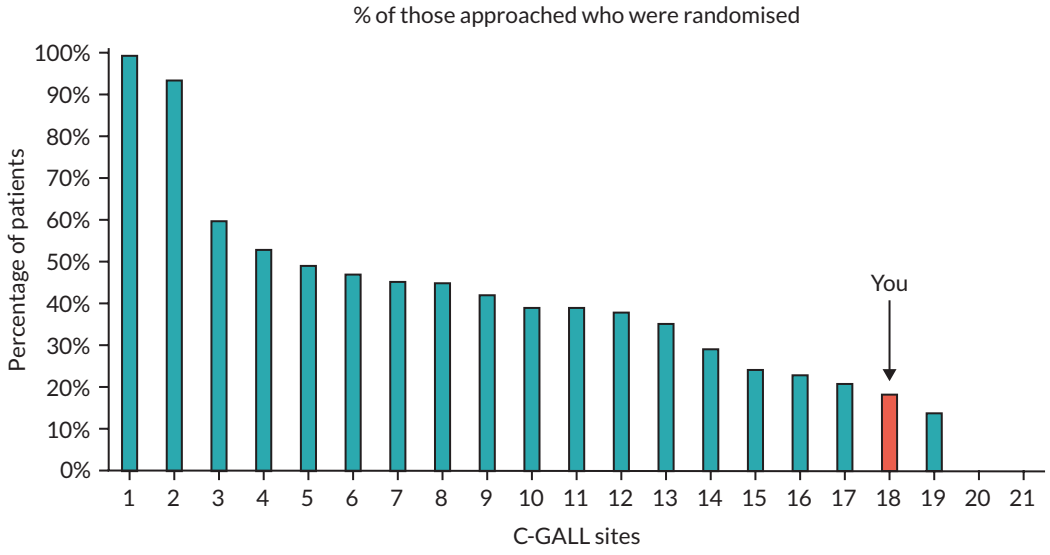

FIGURE 15 Example e-mail content for site feedback (e-mail sent on behalf of the Clinical Co-Chief Investigator).

TABLE 28 Behaviour change techniques identified as relevant for inclusion in site feedback

BCT	Example text
1.2 Problem-solving	There are a number of generic tips on ways to discuss recruitment with potential participants in the attached materials. You could set aside some time to look at these and reflect back on the previous few recruitment conversations you've had – do you think there is anything you could improve upon?
1.4 Action planning	During the consultation, offering a balanced explanation of both treatments and exploring patient preferences may ensure both treatments are given equal consideration. Specific suggestions on how to discuss these elements can be found in the attached materials.
2.2 Feedback on behaviour	Site X is discussing the trial with *upwards of XX% of eligible patients*.
2.7 Feedback on outcomes of behaviour	Although this is extremely positive, I also need to share with you that Site X is also the site that *has one of the lowest percentages of participants that go on to be randomised*. As you can see below, only XX% of trial discussions lead to randomised participants.
3.2 Social support (practical)	Additionally, the trials team would like to offer your site the opportunity to receive further individualised feedback and training on recruitment. Please contact [research assistant] if you would like to arrange a date.
5.1 Information about health consequences	As you will know, achieving our recruitment targets is crucial to being able to answer the research question posed by the C-GALL trial – what is the best treatment approach for patients with uncomplicated symptomatic gallstones?
5.3 Info about social and environmental consequences	Crossovers, especially those in one direction, are problematic for a few reasons. Firstly, they suggest that participants have a strong preference for surgery and are unlikely to have been in equipoise when they agreed to participate. Secondly, a high number of crossovers will make the trial results unreliable and make it impossible for C-GALL to answer the research questions.
5.5 Anticipated regret	I need to minimise the numbers of crossovers so that all the hard work we are doing to recruit participants is not wasted.
6.2 Social comparison	The average across sites is XX% with some sites achieving upwards of XX%.
9.1 Credible source	Letter sent on behalf of Co-Chief Investigators.
10.4 Social reward	Thank you very much for your efforts on the C-GALL trial. I am really grateful for your ongoing support to achieve our recruitment targets.
12.2 Restructuring the social environment	Why not get together with the rest of the research team in Site X and agree a plan about how to discuss C-GALL in the future.

recordings with site-specific examples (where available) alongside examples of what had worked well at other sites (*Table 29*). The second attachment included a one-page 'top tips' sheet which aimed to address the core challenges identified in the consultations (*Box 1*). The sheet focused on tips regarding how to structure the conversation and explain trial-specific processes (e.g. randomisation), building on the examples of good practice provided in the first document they received. It also provided questions for them to ask participants to help explore treatment preferences. All sites were provided with this 'top tips' sheet along with either site specific or generic feedback on consultations.

E-mail feedback was provided to all 21 sites, with generic and site-specific feedback dependent on presence of audio recordings. Sixteen of the sites were responsible for recruiting 400 of the 421 total patients (95%) recruited to the C-GALL trial. This suggests that the hypothesised incremental improvements in recruitment from receiving feedback will have had the opportunity to be realised across these high recruiting sites.

Amendments to the trial patient information leaflet

The PMG discussed the findings from the participant interviews and audio recordings and considered these in the context of the existing content of the trial PIL. It was agreed that the existing PIL could

TABLE 29 Feedback from consultations – attachment to e-mail

	Examples from your site	Examples of what works well at other sites
Balanced presentation about both treatments	*'So, some people get infection of the gallbladder, which you may have already had, which can become more problematic. The risk of that is around about 10%/15% per year, okay. Other risks is something called pancreatitis, which the risk of that is about 1% to 2% a year, and that can be quite, potentially quite serious, and that is why historically if you like we've taken out a lot of gallbladders because people are worried that they don't want that to happen to individuals. But that said, the surgery itself poses risks, and some people still find they get pain and symptoms and things after the operation as well … If you don't have an operation, you are at risk of having an infection related to your gallbladder, which is probably what you had before'* *P: 'Yeah, okay'* *R: 'You can get another … there is other risks in terms of what we call pancreatitis. The risk of that is lower than the risk of getting an infection, but it is potentially quite a serious problem … it can need quite long hospital admission, and it's even potentially life threatening, all right … The risk of that is relatively low, about 1% to 2% of people with gallstones will have that problem at some point in a year you know, but whether that will happen to you or not is almost impossible to predict, all right'.*	*'There is one in three chance you may not get better after surgery, and there is a 10% risk of complication, so these are the risks of surgery. There are risks of medical management that there is a 0.7% chance per year for you to pick up a complication which might end up you having surgery at some point, so the arguments are quite balanced'*
Discussing preferences Use direct and indirect questions – e.g. why do you prefer this treatment?, acknowledge their reasons and balance their views – e.g. highlight advantages of less preferred treatment and disadvantages of preferred	*R: 'Are you interested in the study or what do you think?'* *P: 'Just happy to get it gone'* *R: 'You just want an operation?'* *P: 'Yeah, just get it gone'* *R: 'You understand that there obviously are risks and' (overspeaking)* *P: 'Yeah, yeah, yeah'* *R: 'Fair enough'.*	Patient: 'I would rather have the surgery' *'Okay, no, that's entirely your own choice … By having surgery it won't necessarily get rid of the pain. It may not be, the pain may not be coming from the gallbladder and the gallstones, this is another reason why they suggest the trial … But you do get a priority; if you went on the trial and you were randomised for the management side of it but then all of a sudden the pain was too unbearable, you have got an open appointment to come back whenever you want to. It's entirely – I'm not trying to persuade you one way or the other'*

TABLE 29 Feedback from consultations – attachment to e-mail (*continued*)

	Examples from your site	Examples of what works well at other sites
Discuss uncertainty Remind that worldwide there is no evidence that one treatment is better than the other and link this to why C-GALL being done	*I don't think there's a right and wrong answer, and I think that's the thing, it's difficult, almost impossible to predict in terms of whether you're going to have an issue, when and where, and that's why I'm talking to you about this study because I genuinely believe that I'm not … it's not clear exactly what we should do. But like I say, some patients come to me with a very clear thing, 'I want this doing, I don't want this happening again', and that's it … and some don't.*	*Now at the moment there is no good evidence to suggest which one is the better treatment and … because you fall in that criteria in which we are not sure about whether surgery is the right thing, or carrying on medical management is the right thing.* *We don't know what the right option is … So the outcomes are balanced between two treatments and we don't know which one is better than the other, and that's the reason of running this trial.*
Discussing crossovers	*Well one thing I would add about the trial is that it's not the kind of way we should … it's meant to be set up. But if things change it doesn't mean that you exclusively have to still stick with that allocation if that makes sense.* *So some patients for example I have involved the trial, they've been put into the medical management arm and watched for 18 months but then their symptoms have got quite a bit worse and they now want an operation, and that means that they just change, but they're still part of the study but they're part of the study that say, 'Well this approach didn't work for me'.*	*Explore preferences with patients, stress the need for participant to consider all treatments and equipoise, and don't give impression to be biased to one treatment or the other. Don't volunteer information about cross overs but give opportunity for participant to ask questions and address these.*

BOX 1 Tips for discussing with patients (attachment to e-mail)

- This document provides suggestions for the clinical researcher team to help explain the C-GALL trial to patients
- This document is *not* a script – it provides suggestions that should help to improve informed decision-making and the informed consent discussion
- Please consider using some of these suggestions to complement your own consultation style/approach.

Key strategies for structuring the conversation about participation in the C-GALL trial

- Mention and discuss the C-GALL trial with all eligible patients *early on* during consultation
- Explain what *participation involves* (randomisation, treatment allocation, completion of questionnaires)
- Provide *well-balanced information* about both treatments and *be consistent* across all patients
- Discuss *randomisation process* (e.g. we allocate treatment at random) *early on* during consultation as preferences are then often expressed
- Encourage patients to keep an *open mind* until you have had the chance to present the trial
- Remind patients about the *uncertainty* (e.g. *worldwide there is no evidence that one treatment is better than another*)
- Highlight that the patient will *not be advantaged or disadvantaged* if they were to select a treatment or be randomised
- Mention that the patient is suitable to receive both treatments and there are risk and benefit in both approaches (detailed in PIL)
- Explore reasons underlying a patient's preference *through direct and indirect questions* (e.g. Why do you prefer this treatment?), *acknowledge* their reasons, and *balance their views* about treatments (e.g. highlight the advantages of the less preferred treatment and disadvantages of the preferred treatment)
- *Stress the need* for participant to consider all treatments and equipoise but *don't give the impression* to be biased towards either treatment
- Give the patients an opportunity to ask questions
- Stay positive and reassuring all the time. Do not predict the outcome of the trial/study (if asked)

provide additional information regarding expectations about trial retention. Specifically, more information on when trial participants would receive questionnaires and how long they would receive them for that is frequency and duration. The importance of completion and return of the questionnaire was also reinforced as important to assess the impact of the different treatments, even if participants did not experience any symptoms. The importance of questionnaire data to produce reliable results was also emphasised. These amendments were made to the trial PIL and approved by the REC with an updated version implemented in June 2019 (see *Appendix 2*).

Updates to cover letters and newsletters to target questionnaire response

Recommendations to improve questionnaire return were made to the PMG based on the findings from the behavioural diagnosis findings in the Phase 1 interviews with trial participants. A list of relevant BCTs was generated from existing templates and C-GALL trial paperwork and compared to those identified as relevant from the patient interview findings. A total of 12 BCTs were identified as being relevant for return of participant-completed questionnaires in the C-GALL trial. Specific content was developed or enhanced within the cover letters and newsletter for trial participants to ensure these techniques were covered.

Table 30 presents the specific BCTs, the suggested text for inclusion in the letters, and the justification for inclusion based on interview findings. Changes to the questionnaire cover letters and to trial newsletters were implemented in June 2019.

Discussion

The purpose of this process evaluation was to inform the delivery of the trial, and in particular to focus on the critical components of recruitment and retention of trial participants. We applied existing methods used to improve recruitment to surgical trials and enhanced their practicality by also considering challenges to trial retention. In addition, we applied a behavioural science lens through which to view the trial process problems – a novel contribution to process evaluations in ongoing trials to directly inform and improve recruitment and retention.

TABLE 30 Proposed BCTs and suggested text for inclusion in letters/newsletters

BCT	Suggested text	Justification for inclusion from participant interview findings
1.4 Action planning 11.3 Conserving mental resources	'The questionnaire requires you to answer questions about your general health over the past 3 months. It can help to keep a note of your symptoms in a diary throughout your participation in this trial. This will help you to complete the questionnaires quickly and as accurately as possible'.	Participants suggested that they struggled to remember their symptoms over the time period specified on the questionnaire. Many participants also said that it helped/would have helped to keep a note of their symptoms to facilitate accurate reporting.
1.9 Commitment	'The C-GALL trial requires a strong commitment from you and [local hospital] to stay in the study until the end. This means your [local research team/consultant] has placed considerable trust in the patients they asked to join them in this research. Your [consultant] will not be able to fulfil their part in this study without the continued co-operation and participation of their patients'.	The majority of participants indicated the importance of committing to the study, which influenced questionnaire completion and return. Indicating that study participation requires commitment from participants to complete questionnaires at different time points may be effective.
5.1 Information about health consequences 5.2 Salience of consequences	'As you are aware, living with gallstones can be debilitating. Your participation in this study will help us gain insight into how different treatment options influence well-being. This will help us to identify ways we can support individuals with gallstones to improve their quality of life'.	Participants reported that they had hoped their participation in the study will help people with gallstones in the future. This was also cited as encouraging trial participation.
2.7 Feedback on outcome of behaviour	'The information you provide in these questionnaires will contribute to the integrity of the results'.	Some participants mentioned that they felt guilty (or would have felt guilty) for failing to return the questionnaire. Many acknowledged that this was because they knew the questionnaires were important for the study. Emphasising this could serve a potent reminder.
9.1 Credible source 6.3 Information about others' approval 10.4 Social reward 3.2 Social support (practical)	Newsletter: [Picture of Co-Chief Investigators] 'Co-Chief Investigators, Professor Craig Ramsay and Professor Irfan Ahmed would like to thank participants for their contribution to the trial to date. Without your support, we couldn't do our research' [Picture of Trial Manager and Data Coordinator] 'The trial management team would like to express sincere thanks to those who have responded to the questionnaires we have administered so far. This is a vital component of the trial and we appreciate that it takes time and effort to complete. If you have any queries about the questionnaires, please feel free to get in touch via e-mail [indicate e-mail] or telephone [insert number]'	Thanking participants within clinical trials could have an impact on their motivation to participate. A couple of participants in the qualitative study also indicated their appreciation for the support from the trial office. This was identified as a facilitator to questionnaire return: including pictures/quotes from the trial team and trial contact information will highlight the practical support available to them.
2.2 Feedback on behaviour	If possible, include information about how many questionnaires have been completed, how many are left to complete and an indication of when their participation in the trial ends, tailored to the individual. Could draw attention to it by including this information within a table.	Participants often reported that they could not remember how many questionnaires they had completed/returned. Many participants also reported being unaware of when their participation in the trial ends – including this information will provide them with knowledge of their participation status and could increase their motivation to complete the trial.

continued

TABLE 30 Proposed BCTs and suggested text for inclusion in letters/newsletters (*continued*)

BCT	Suggested text	Justification for inclusion from participant interview findings
4.1 Instruction on how to perform the behaviour	'As part of the C-GALL study you are asked to complete five questionnaires at different time points. These questionnaires are identical but are designed to measure your well-being over a certain amount of time. It is important to complete the questionnaires even if you have not had surgery or are not experiencing any symptoms around the time that you receive the questionnaire'.	Many participants reported that they felt the questionnaire contained items that were not relevant to their circumstances, either because they had not had surgery or they were not experiencing symptoms around the time of questionnaire administration. Some were also confused about receiving the same questionnaire on multiple occasions.

Audio recordings of trial consultations provide a rich data set to understand the complexities of decision-making around trial participation. Several studies using this approach have now demonstrated the value of content review and feedback on recruitment discussions.[82] Similar to these existing embedded efforts to improve recruitment in trials we also experienced challenges associated with encouraging healthcare professionals at all sites to audio-record the consultations in which the trial was discussed. Of the 22 recruiting centres in C-GALL, 16 provided at least one consultation audio recording. Exploring ways to best encourage and support sites to routinise the audio recording of the consultations in which the trial is discussed needs to be revisited for future trials.

Findings from the audio recordings and participant interviews led to the PMG reviewing and revising the PIL to enhance the information on trial retention. Existing research has demonstrated that information on the consequences of not completing a trial is not routinely provided in trial PILs.[87] When considered alongside our interview findings that indicate participants were not clear on the expectations of them during trial follow-up indicate the need for more directive information on trial retention is required from trial outset. This is further supported by research that highlighted potential trial participants deem understanding what they will have to do as a trial participant when making the initial decision to participate as highly important.[88] We amended the C-GALL PIL based on the findings of our embedded process evaluation, but the learning generated is transferable for application to future trials.

Rather than being firmly prescriptive in its design, the embedded process evaluation allowed a targeted adaptive approach in order to be responsive to the needs of the trial. This flexibility was a core strength of the activity and allowed additional value to be gleaned from the investigations of trial process, for example, through exploring discussion of trial retention during recruitment consultations. This also allowed the process evaluation to contribute to the wider methodological literature through its investigation of trial retention in this way.[68] The other strength of the activity was the use of behavioural theory to inform data collection, analysis and solution development. The application of behavioural science to problems of trial process is gaining attention.[73] Our process evaluation built on these existing developments but elevated the approach by conducting a series of linked activities from multiple data sources to inform the development of participant-centred strategies.

There were a number of weaknesses. Notably, the sample of interviews with trial participants who were not able to complete the trial till the end constituted a small proportion of those who were initially invited. As such, the views represented by those interviewed may not represent the majority of participants who did not complete and return trial questionnaires. The other potential weakness of the study is that we did not formally evaluate any of the strategies which were implemented based on the findings of the process evaluation. Preliminary evaluation of the impact of the e-mail feedback (using before and after analysis) suggested it was effective but more robust evaluation (e.g. interrupted time series or randomised evaluation) is required and will be a focus of future activities. As mode of follow-up changed (from postal questionnaire to telephone follow-up) due to the pandemic, it was not possible to evaluate the participant-facing strategies focusing on trial retention (PIL, cover letter, newsletter).

Conclusion

The embedded process evaluation in C-GALL allowed exploration and targeted solution-focused strategies that aimed at improving recruitment and retention across the trial. The adaptive nature of the design of the evaluation allowed issues to be investigated in both a proactive and reactive response, improving efficiencies and outputs for the trial. There are some generalisable findings and recommendations for good practice for ongoing and future trials. These include ensuring that details of follow-up processes are shared with participants in both written and verbal information that is shared during informed consent, and highlighting the importance of completing questionnaires (including at multiple time points). The innovative adoption of a behavioural science approach to much of the process evaluation contributed novel insights into trial conduct challenges. Future efforts should focus on the formal evaluation of the approach and solutions developed.

Chapter 7 A core outcome set for symptomatic uncomplicated gallstone disease

Introduction

Recommendations from the recent NIHR guideline on Gallstone Disease[66] have clearly demonstrated that insufficient high-quality information has been produced for patients on the effect of cholecystectomy on patient outcomes. The guideline recommends that 'research is needed to establish the long-term patient benefits and harms, so that appropriate information can be provided to patients to aid decision-making and long-term management of their condition'.[66] However, due to the way they have been designed and conducted many completed trials are not as informative on patient care, as they could be due to lack of standardisation across studies, outcome definition, collection and reporting. This heterogeneity of outcomes across studies also hampers useful synthesis of primary studies in meta-analyses and ultimately negatively impacts on decision-making by all stakeholders. In addition to the heterogeneity of outcomes currently reported and the problems this causes, measuring the wrong outcomes (i.e. those that are not valued by clinicians or, more importantly, patients) could also be a real risk for many studies if stakeholders are not consulted during the trial design process. One way that these problems with heterogeneity of trials and the consequent relevance of their results to stakeholders can be addressed is through the development and use of COSs.[90,91] At study inception there were no registered or published COS for symptomatic uncomplicated gallstone disease.

Core outcome set aims to define a minimum set of outcomes that should be considered 'core' for the evaluation and reporting of specific interventions or conditions (i.e. the set of outcomes that should always be considered and ideally measured in effectiveness trials).[90,91] There is a growing body of literature to provide support for development of COS.[90] Specifically, they are developed using consensus methods involving stakeholder groups, such as health professionals and patients, so as to ensure that the outcomes being defined are both clinically and personally relevant for the individuals involved.[90,91] Generation of a COS is not expected to be mutually exclusive to the measurement of other outcomes. However, a core set will foster greater consistency in outcome reporting between studies and lead to more meaningful data being available to contribute to future meta-analysis.[90,91] Moreover, COS can minimise the threat of outcome reporting bias by ensuring consistency between what is measured and what is reported.[90,91] Ultimately, they should improve the overall efficiency of trials through focused outcome collection and reporting, and the quality of the evidence they produce, enabling better informed healthcare decisions to be made.

The general methodology describing the development of a COS (a systematic review of the literature to identify outcomes reported to date; interviews with stakeholders to explore additional outcomes of importance; and a consensus-based approach to determine which outcomes should be considered core) was adopted in this research to develop a COS for symptomatic uncomplicated gallstone disease. Details of this project were registered in advance and included in the Core Outcome Measures in Effectiveness Trials (COMET) Initiative database.[92]

Aims and objectives

Aim

The aim of this study was to develop a COS for uncomplicated symptomatic gallstone disease effectiveness trials, to recommend which outcomes should be measured and reported as a minimum and reflecting the interests of all relevant stakeholders.

Objectives

The study-specific objectives are:

- to identify a list of outcomes from those: reported in studies of the treatment of uncomplicated symptomatic gallstones from a systematic review of the literature; or reported in qualitative studies that have explored the lived experience of those with a diagnosis of uncomplicated symptomatic gallstones with specific reference to outcomes of importance
- to explore additional outcomes relevant to the treatment of uncomplicated symptomatic gallstones using semistructured interviews with stakeholders (which may include patients, clinicians, specialist nurses, RNs)
- to prioritise outcomes from both a clinician and patient perspective
- to compare patient and clinician prioritised outcomes
- to integrate patient and clinician outcomes into a combined COS.

Methods

Study overview

The development of the COS followed best practice and involved two sequential stages.[93] The methods for each stage are outlined below and presented in *Figure 16*. The scope was restricted to interventions (surgical and non-surgical) that treat symptomatic uncomplicated gallstone disease in adults.

Stage 1. Identification of outcomes relevant for symptomatic uncomplicated gallstone disease: generation of the outcome list.

The identification of relevant outcomes was informed by two sources: existing evidence and new primary research. The specifics of these are detailed below.

Identification of outcomes from existing literature

To generate the initial long list of outcomes, we conducted three linked reviews of existing literature: (1) a systematic review to identify outcomes reported in trials of interventions for symptomatic uncomplicated gallstone disease; (2) a content analysis of individual items within disease-specific patient-reported outcome measures (PROMs); and (3) outcomes reported by patients with a lived experience of symptomatic uncomplicated gallstone disease in exploratory studies. This multipronged approach presents an efficient and robust approach to providing the first step towards generation of a COS.

Types of studies

The search strategies developed for the health technology assessment conducted previously by our team were updated.[35] In particular, we aimed to identify additional randomised trials, or studies reporting the development of PROMs, in symptomatic uncomplicated gallstone disease (e.g. trials of early vs. delayed laparoscopic cholecystectomy for uncomplicated biliary colic). We included RCTs from reference lists of relevant Cochrane reviews (early vs. delayed laparoscopic cholecystectomy for biliary colic; robot assistant vs. human or another robot in laparoscopic cholecystectomy; miniport vs. standard ports for laparoscopic cholecystectomy; fewer-than-4 vs. 4 ports for laparoscopic cholecystectomy; early vs. delayed laparoscopic cholecystectomy for acute cholecystitis).[94–98] Reference lists of all reviews

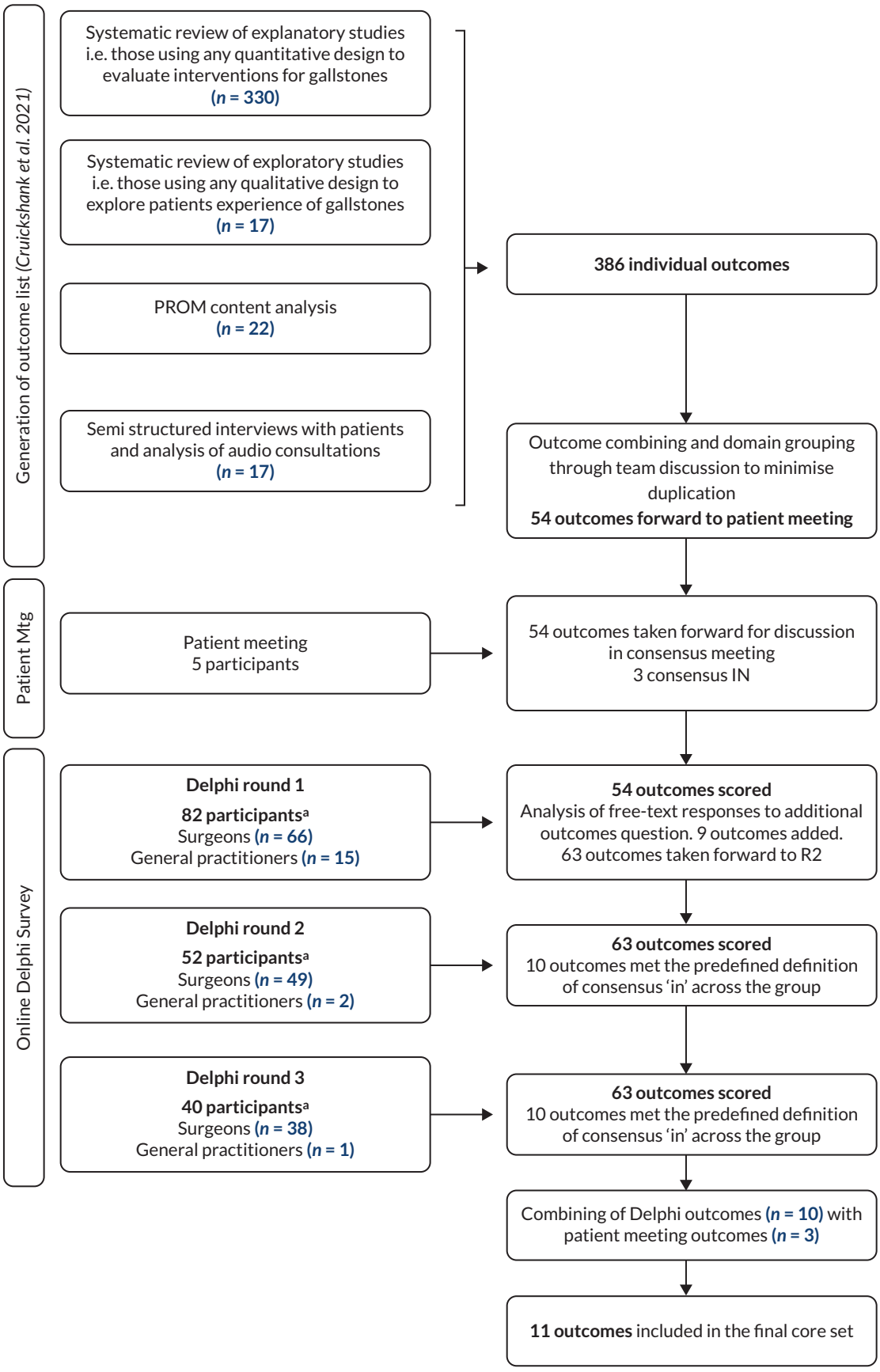

FIGURE 16 Core outcome set development overview. a, 1 Delphi participant did not specify a clinical role.

were searched for relevant studies. Exploratory studies (using interviews, focus groups and other methods) that have explored any aspect of living with symptomatic uncomplicated gallstone disease were also included.

Search methods for identification of studies

Extensive electronic searches were undertaken to identify trials for a project on the clinical effectiveness of cholecystectomy and these are reported in full elsewhere.[35] The databases searched included MEDLINE (1946 to week 37 2012), MEDLINE-in-process (10 September 2012), EMBASE (1974 to 10 September 2012), Science Citation Index (1970 to 12 September 2012), BIOSIS (1956 to 12 September 2012) and the Cochrane Central Register of Controlled Trials (Issue 9–12, 2012). Studies identified from these searches were used to elicit reported outcomes. The MEDLINE and EMBASE searches were updated in May 2016 (September 2012–May 2016) to identify more recent relevant trials.

A specific search of MEDLINE and EMBASE was undertaken to identify studies that report PROM outcome data for cholecystitis, biliary colic and gallstones, with records retrieved by the main search for trials excluded to avoid duplication of the results. This search was undertaken in May 2016 (1980 to May 2016).

In addition, a search for relevant qualitative studies was undertaken in August 2016 in the Ovid versions of MEDLINE (from 1966 to 2016). The search strategies combined search terms for cholecystitis, cholecystectomy, biliary colic and gallstones, with terms for qualitative research.

The search strategies for MEDLINE and EMBASE are reported in *Appendix 5*.

Inclusion criteria for eligible studies were as follows:

- Participants: Adults aged over 18 years with symptomatic uncomplicated gallstone disease.
- Intervention: Any intervention (surgical or non-surgical management, i.e. expectant management or dietary advice or medical therapy) used to manage symptomatic uncomplicated gallstone disease in adults.
- Outcomes: All reported outcomes well eligible for inclusion.

Excluded studies included those focusing on asymptomatic gallstone disease or on acute severe cholecystitis, cholangitis or pancreatitis, which were not considered suitable for inclusion. In addition, studies including ‘complex’ gallstone cases that is, empyema, ascending cholangitis and gallstone ileus, were excluded. Reports published in non-English languages for which a translation could not be organised were also excluded.

Study selection and data extraction

For all three reviews of existing literature conducted, citations identified through the search were independently assessed by one reviewer with another assessing a 10% random sample. All full-text papers considered potentially eligible were screened by one reviewer and checked by a second reviewer.

For the review of interventions for symptomatic uncomplicated gallstone disease, one reviewer extracted details of all outcomes reported (verbatim) and any reported definition of outcomes provided by the authors (e.g. operating time may have been defined and reported by some studies as ‘interval between initial skin incision and in closure’ others ‘duration of surgery’, etc.). Disease-specific outcomes identified as being assessed with validated tools (whether self-reported or not) were reviewed and the specific domains contained within the tool were used to present the outcomes rather than reporting the tool itself.

Data were recorded in Microsoft Excel for Microsoft 365. A 10% sample was checked by a second reviewer. Other relevant data (i.e. study and participant characteristics) were extracted by one reviewer

and checked by a second reviewer. At all stages, disagreement between reviewers was resolved by discussion.

From the list of outcomes reported in trials of interventions for symptomatic uncomplicated gallstone disease described above, disease-specific PROMS were identified and supplemented with the studies identified in the search. Data were extracted by one reviewer who recorded the name of the PROM(s), the reported PRO scales and individual verbatim items.

For the studies reporting participants' experiences, data on study characteristics such as author, publication date, country, focus of investigation, data collection methods, number of participants and details of sample size were extracted. Additionally, two authors independently extracted data from two main sources reporting study findings: (1) direct quotes from participants and (2) authors interpretations of participants' quotes. These data were recorded verbatim and analysed to identify 'descriptive' thematic codes.

Data analysis

All data are summarised and presented in tabular form. Data from the review of interventions for symptomatic uncomplicated gallstone disease are presented verbatim.

The individual verbatim items from each PROM were analysed using an inductive content analysis approach and informed by previous PROM coding work.[99] All PROM items were systematically categorised into conceptual health domains according to the aspect which they aim to capture. Health domains were generated inductively from the identified individual items. Domain mapping was conducted independently by two authors, with any conflicts resolved through discussion.

For the qualitative studies, the constant comparison method was used to compare findings across studies and an inductive thematic synthesis was undertaken to generate a list of themes and subthemes (focused on outcomes) from the data, to map across the presurgery and postsurgery timeline.[100,101] Throughout this process, the description and wording of the themes were continually revised, and notes made as to how themes and/or subthemes related and how some could be merged. These findings were discussed further with the research team to finalise the themes across the studies, and these were considered, where appropriate, as domains relevant for inclusion in the development of the COS.

Identification of outcomes of importance to stakeholders from new primary research

In addition to outcomes reported in existing literature, we conducted primary qualitative research to further inform the identification of outcomes of relevance to patients. Data were sought from interviews with patients with a diagnosis of symptomatic uncomplicated gallstone disease; a focus group with patients who had undergone cholecystectomy; and audio consultations from participants taking part in the C-GALL trial.

Participant identification and recruitment

Potential participants for the interviews were identified and recruited from those participating in the C-GALL trial. Participants were provided with a study PIL either in the clinic or posted to the participant if a decision about C-GALL trial entry was made later. The PIL contained a detachable reply slip to complete and return to the researcher (in a reply-paid envelope) indicating whether they would like to discuss participating in the interview study exploring various aspects of trial participation and included outcomes of importance. Patients being approached to participate in the C-GALL trial were asked (for trial purposes) if they would consent to their consultation being audio-recorded; more information in *Chapter 6*. If consent was obtained, these audio recordings were then analysed by the research team for the identification of outcomes.

A second group of participants took part in a focus group discussion, who had experienced gallstone disease, were identified through the Scottish Health Research register (SHARE www.registerforshare.org/) and sent an invitation letter asking them to contact the research team if they were interested in participating. Following initial contact, a researcher phoned the interested participants and ensured they understood what taking part in this study entailed and arranged a suitable time for the focus group.

Data collection

One author conducted the interviews over the telephone between April and August 2017. The focus group was conducted by two members of the trial team in July 2017. Trial consultations from the four sites open to recruitment (at time of COS development) were sampled to inform outcome identification. Informed (written and recorded) consent was obtained from all participants (for interviews, focus groups and consultations) prior to data collection, and confidentiality of the participants was assured.

Participants were encouraged to consider what aspects of their disease or treatment impacted them most, both in terms of physical and psychological functioning and what improvements they would wish to see in terms of their outcomes. The interviews, focus group and audio consultations were audio-recorded and transcribed verbatim (targeted transcription for the consultations) using a professional transcription service.

Data analysis

All transcripts were imported into NVivo (V.10, 2013: QSR International, Warrington, UK) and analysed using conventional content analysis (i.e. coding categories are derived directly from the text data and are used to interpret meaning from the content).[99] Various themes and subthemes were generated by one researcher based on the contents of the transcripts to identify the outcomes stated by the participants, and these were then further discussed with another member of the team to finalise the list of outcomes identified across the primary qualitative data. The analysis was oriented to address the aim of identifying the range of outcomes that might be considered important to patients and the reasons used to justify assessment of them as important.

Categorisation of identified outcomes into outcome domains

The list of potential outcomes generated from both the systematic evidence search and the primary qualitative research formed the basis of a 'long' list of outcomes used to refine the items into a final 'short' list for inclusion in the consensus stage of the COS development. Outcomes were first grouped and reduced according to original source, that is, the initial long list from the systematic evidence review was reduced for direct duplication by two members of the research team. A similar process was conducted to deduplicate the outcomes identified from the PROM coding, qualitative evidence synthesis and primary qualitative research. These outcome lists were then merged and reduced (through iterative group discussions) based on areas of overlap to produce a final shortlist of individual outcomes of relevance in this context.

These individual outcome items were further grouped into broader concept-level headings to categorise outcome domains. These concept-level headings were informed by other outcome categorisation work in the area of COS and supplemented through discussion with the project management group.[102–104]

Stage 2. Agreement of outcomes relevant for symptomatic uncomplicated gallstone disease: defining the core set.

As yet, there is no agreed methodology on the best approach to achieve consensus when developing a COS for use in effectiveness trials.[90] A range of consensus-generating approaches have been used in existing COS, ranging from survey-based approaches (e.g. Delphi) to consensus meetings (Nominal Group Technique) and variations in between.[90] This study chose to use a mixture of consensus meeting with patients and a Delphi consensus survey approach with healthcare professionals. These methods are detailed below.

Consensus meeting with patients

Patients who were members of the C-GALL PPI Group (established to provide patient input into the C-GALL trial) were invited to contribute to the consensus meeting. In advance of the meeting, these patients were e-mailed information, including a brief description of COS and the purpose of the activity. They were also provided with the list of identified outcomes from stage 1 of this study and asked to identify their top three outcomes to discuss at the meeting. The group discussion was facilitated by the Health Services Research Unit PPI co-ordinator and the C-GALL trial PPI grant holder. Verbal consent was sought at the start of the meeting, with all discussion audio-recorded. Everyone shared their top three identified outcomes with a short explanation of why they felt it was important. All outcomes identified as important were discussed further, allowing each person to put forward their feelings about their chosen outcomes. After discussion, the individuals then rated the outcome as 'high', 'medium' or 'lower' importance. It was stressed that naturally, these were all important, but the purpose was to focus on the single most important outcome, which should be collected every time. The outcomes were then ranked based on the average scoring and the consensus definition applied to determine inclusion of outcomes in the core set. After all initially identified outcomes were discussed, the group were given time to check whether there were any other outcomes they felt would be important to include.

Online Delphi survey with healthcare professionals

A Delphi survey was used to seek agreement on the relative importance of outcomes identified from all previous stages. Each outcome generated from Stage 1 was listed together with a plain language definition on the online Delphi Manager platform.[105] Although there is no formal guidance on sample size for the panel in Delphi surveys, a minimum of 10–18 has been suggested, and as such, we aimed to have a final sample larger than 18.[106,107] Healthcare professionals were invited to participate through e-mail distribution lists of professional societies and through social media, The Association of Surgeons of Great Britain, Association of Upper Gastrointestinal Surgery of Great Britain and Ireland and Association of Laparoscopic Surgeons of Great Britain and Ireland (ALSGBI) were specifically requested to send the e-mail invite to their memberships. General Practitioners at Scottish sites involved in the C-GALL trial were also sent invites through the Scottish Primary Care Research Network.

The e-mail invite contained a brief information sheet and a link to the Delphi website, providing further information and allowing interested participants to register. Potential participants were reminded of the importance of completing all rounds of the process. Those who agreed to participate were asked to register their details online, which generated a unique identifier against which their data was stored and reminder e-mails for non-response were generated. Participants were not able to identify other participants or individual responses from other participants. Consent was implicit by completion and return of the questionnaire.

Three rounds of the survey (R1, R2 and R3) were completed. In R1 the outcomes were listed alphabetically by domain (to minimise potential weighting due to ordering effects), and participants were asked to consider what outcomes should be measured and reported as a minimum to develop a COS for uncomplicated symptomatic gallstone disease effectiveness trials. They then scored outcome importance using a 9-point Likert scale, where 1–3 was not important and 7–9 was essential, as per recommended methodology for COS.[93] Participants were also invited to submit any additional outcomes during R1, which were reviewed by the study team and considered for inclusion in R2. All participants were then asked to complete R2. Participants were provided with their R1 score and an anonymised distribution of the group's scores. Participants were asked to consider this information when scoring the outcome again in R2. A final round of rating (R3) was completed providing feedback as per R2. However, R3 also provided responders with patient scores (from the patient's consensus meeting) and asked them to consider the importance of the outcome considering their own scores, the scores of the other responders and the scores of the patients. No outcomes were removed between the rounds.

Analysis of round 3

For the final analysis, for each outcome the number of participants who scored the item and the distribution of scores were summarised, alongside the number of respondents who scored the items across all three rounds. For each outcome, the proportion of respondents scoring 1–3, 4–6 and 7–9 on the Likert scale was calculated for each outcome. Each outcome was classified as: 'consensus in' (i.e. consensus that the outcome should be included in a core set), 'consensus out' (i.e. consensus that the outcome should not be included in a core set) or 'no consensus' (i.e. items that are equivocal and require further research for clarification).

The original consensus definition was based on existing COS studies that required 70% or more of the group to agree an outcome as important (or not) and < 15% score in the opposite direction. However, while blinded to outcome identity, the consensus definition was amended given the large number of outcomes meeting the original consensus definition and considered 'consensus in' (i.e. 21). As such the stringency for consensus on outcomes being considered 'consensus in' and included in the core set was increased to more than 95% of respondents scoring outcomes as 7–9 (i.e. essential) and 'consensus out' considered as 50% or less scoring 7–9. Following the Delphi, the outcomes that had reached consensus from the patient meeting were compared and combined with the outcomes meeting consensus for being included in the core set from the Delphi survey (*Table 31*).

Results

Generation of outcome list

Sample characteristics

The search of existing literature identified 137 studies (129 from trials and PROMs and 8 qualitative studies) eligible for inclusion (*Figure 17*). A total of 137 publications, from 129 RCTs with 4 reporting gallstone disease-specific PROMs and 8 qualitative studies were included. The majority of the included trials involved one type of surgery versus another type of surgery.[108–113] There were only a few non-surgical interventions.

The PROMS varied in the aspect of focus of the measure with only two PROMs focused on gallstone disease,[114–116] one on gastrointestinal diseases,[117] and the other on quality of life after abdominal surgery.[118] All of the PROMs reported to measure multidimensional constructs, for example, QoL. The PROMs varied in the number of constructs they aimed to assess (ranging from 4 to 8) and the number of items they asked participants to report on (ranging from 5 to 41, median = 27).

Of the eight included studies using qualitative methods, seven used interviews [face to face (n = 5) or telephone (n = 2)] and one used focus groups for data collection. All of the treatments being explored in the studies were surgical, but the types of surgery varied.[117,119,120] *Table 32* provides a summary of the characteristics of participants included in the quantitative and qualitative literature review.

TABLE 31 Definition of consensus

Consensus classification	Description	Original definition	Amended definition
Consensus in	Consensus that outcome should be included in the COS	≥ 70% scoring 7–9 AND < 15% scoring 1–3	≥ 95% scoring 7–9
Consensus out	Consensus that outcome should not be included in the COS	≥ 70% scoring 1–3 AND < 15% scoring 7–9	≤ 50% scoring 7–9 and no new compelling reasons in the comment boxes
No consensus	Uncertainty about importance of outcome	≤ 50% scoring 7–9 and no new compelling reasons in the comment boxes	

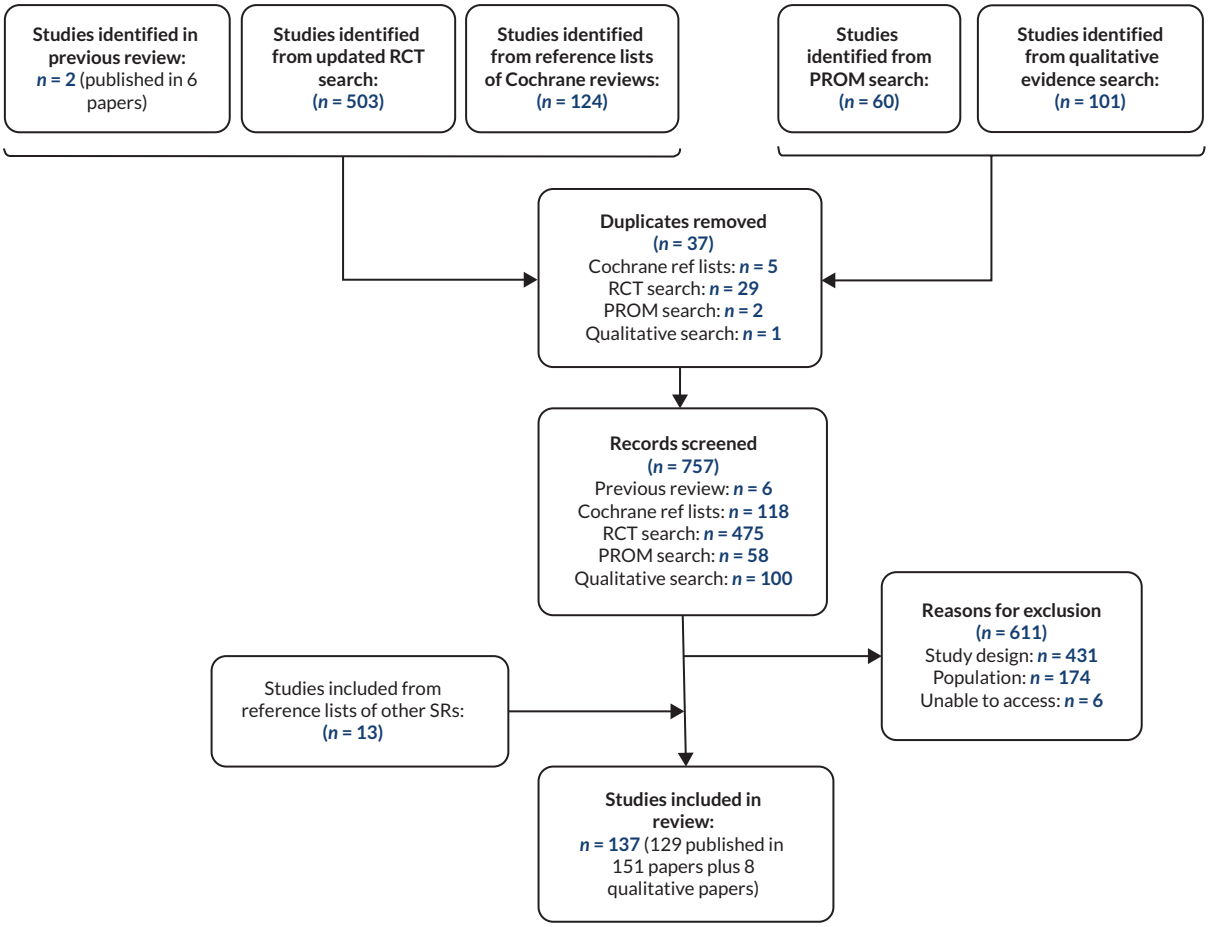

FIGURE 17 PRISMA systematic review diagram for evidence synthesis.

The primary qualitative research included 6 individual interviews, 1 focus group (n = 5 participants) and analysis of 20 consultations (5 each from 4 different hospital sites) which provided data from a total sample of 31 patients (*Table 33*). They included 26 women and 5 men, 26 of whom had been approached to take part in the C-GALL trial (12 trial consenters and 14 trial non-consenters), and 5 patients who had not been approached about the trial but who had all had a cholecystectomy.

Outcomes identified from existing literature and new primary research

A total of 386 individual recorded outcomes were identified across the combined evidence from existing literature and new primary research: 330 outcomes (which were reported 1147 times) from trials evaluating interventions, 22 outcomes from studies reporting PROMs, 17 outcomes from existing qualitative studies and 17 outcomes from primary qualitative research.

The review of existing qualitative literature identified five additional outcomes for inclusion in the long list, not identified by the trials or studies reporting PROMS, namely: dizziness, fainting, trust, weight and prevention of additional disease. The primary qualitative research identified a further three additional outcomes (breathing problems, cough and mortality) that were not in the previous patient-focused evidence (studies reporting PROMs and qualitative literature), with two of these (breathing problems, cough) making unique contributions to the overall outcome list.

The 390 individual reported outcomes across the four data sources were reduced into a 'short' list of outcomes which could be measured in comparative effectiveness trials (i.e. phase III pragmatic effectiveness trials) of interventions to treat uncomplicated symptomatic gallstone disease (see *Appendix 6*). Several outcomes were dropped from the long list as they were deemed not eligible as

TABLE 32 Summary characteristics and demographics of included studies

Characteristic	Quantitative review	Qualitative review
Number of study participants	(*n* = 119 studies)	(*n* = 8 studies)
Median	75	19.5
Range	14–618	6–100
Total	10,757	256
No of males/females	(*n* = 96 studies)	(*n* = 7 studies)
Median	21/53.5	4/16
Range	0–255/15–415	2–15/4–37
Total	2,632/6,166	43/113
Age (years)	(*n* = 14 studies)	–
Median (years)	46.1	–
Range (years)	40–53	19–81
Country	(*n* = 128 studies)	(*n* = 8 studies)
UK	7[a]	1
EEA (excluding UK)	44[a]	3
USA	12[a]	1
Other	66	3
No. of centres	(*n* = 129 studies)	(*n* = 8 studies)
Single centre	113	8
Multicentre	16	
Total no of outcomes reported	330	17

a One study (Marks 2013) was conducted in the UK, USA and Italy and is included in the count for each country.

Note
Number of studies (*n* = xx studies) in table relates to how many studies reported each relevant characteristic.

TABLE 33 Demographics of participants included in the research to identify outcomes for Delphi

Research type and number of participants (*n*)	Gender	Approached to take part in the C-GALL trial, Yes/No and number (*n*)	Trial consenters (allocated to receive surgery or observation/conservative management)[a] or decliners
Interviews at recruitment (*n* = 6)	Female = 6	Yes (*n* = 6)	Allocated surgery = 4 Allocated medical management = 1 Trial non-consenter = 1
Audio consultations (*n* = 20)	Female = 16 Male = 4	Yes (*n* = 20)	Allocated surgery = 4 Allocated medical management = 3 Trial non-consenter = 13
Focus group participants (*n* = 5)	Female = 4 Male = 1	No	Not applicable

a During the course of the trial observation/conservative management was defined as medical management.

clinical endpoint outcomes for use in trials of this type (e.g. system and process outcomes such as duration of surgery which might be important in earlier phase trials). Therefore, the final list covered 27 broad outcome domains that contained 54 distinct outcomes.

Defining the core set

Consensus meeting with patients

Five patients with uncomplicated gallstone disease, three of whom had previously had surgery (two of whom had resolution of symptoms, one who did not) and two patients who had not had surgery, participated in the consensus meeting. Following sharing of the prioritised outcomes, the group agreed that three outcomes should be considered as having 'high' importance for inclusion (i.e. in a COS), namely: QoL, overall pain and overall health state. When discussing the overall health state, there were also discussions about anxiety. The group felt that any measurement of overall health state would be more important than anxiety alone, but that it was important to take into consideration when measuring overall health state, the anxiety aspects of the disease and overall mental strain when dealing with the symptoms and length of time waiting for diagnosis and treatment.

Online Delphi survey with healthcare professionals

Round 1 of the Delphi survey was completed by 82 participants (66 surgeons and 15 GPs and 1 participant who did not specify a clinical role) and R2 by 54 (65% of those from R1). The R3 participant sample comprised 38 surgeons, 1 GP and 1 respondent who did not specify a clinical role (a total of 70% of those from R2) (*Table 34*). The majority were male (82.5%), aged 45–64 (55%) and based in

TABLE 34 Delphi sample demographics (round 3)

	N = 40
Subgroup	
Surgeon	38 (95.0)
GP	1 (2.5)
Unknown	1 (2.5)
Gender	
Male	33 (82.5)
Female	7 (17.5)
Age (years)	
18–44	17 (42.5)
45–64	22 (55.0)
65–84	1 (2.5)
Place of residence	
England	34 (85.0)
Scotland	4 (10.0)
Wales	2 (5.0)
Years of treating uncomplicated gallstones	
1–3 years	33 (82.5)
Missing	7 (17.5)

England (85%) (see *Table 34*). The Delphi survey was open to responses across the three rounds from June 2019 to January 2021.

Following R1 scoring, 41 additional outcomes were suggested for consideration. Following discussion with the COS study team, nine were agreed as new and taken forward for scoring in R2 and R3. These additional nine included: re-admission, reoperation, the need for ERCP, percutaneous drain, use of analgesia, need for an outpatient appointment, steatorrhea, pancreatitis and appetite. The other suggested outcomes were excluded due to being out of scope or duplicates of existing outcomes. Scores from 40 participants in R3 were included in the final analysis, an attrition rate of 37% from R1, and 23% from R2. Attrition analysis indicated no significant difference between responders and non-responders between the Delphi rounds.

Core outcome set

At the end of R3 of the Delphi, 10 outcomes achieved consensus for inclusion in the COS. Two of these 10 outcomes had also been identified at the consensus meeting with patients. An additional outcome (overall pain) was also identified as consensus during the patient meeting but did not achieve consensus during the healthcare professional Delphi. The final COS includes eleven outcomes grouped across four domains (*Table 35*). These were QoL, overall health state, overall satisfaction, overall pain, common bile duct injury, biliary leak, haemorrhage, need for endoscopic retrograde cholangiopancreatography, intra-abdominal collections, admission/re-admission for problems and reoperation.

Discussion

COSs for use in effectiveness trials comparing surgical approaches (to a range of comparators) have seen a significant increase in development over recent years. For example, the COMET registry contains 183 entries including the word 'surgery' in its database and includes a broad range of specialties from cancer to trauma and cosmetic surgery.[121] The value of the COS developed in the project reported here is for trials of interventions to treat uncomplicated gallstone disease, of which there are many trials evaluating various interventions.[35] While the COS reported in this paper was designed for use in

TABLE 35 COS for uncomplicated symptomatic gallstone disease

Outcome (*n* = 11)	Domain
QoL[a]	Generic health
Overall health state[a]	
Overall satisfaction[b]	
Overall pain[c]	
Common bile duct injury[d]	Intraoperative AE
Biliary leak[d]	
Haemorrhage[d]	
Need for ERCP[d]	
Intra-abdominal collections[d]	Intra- and postoperative AE
Admission/re-admission for problems[b]	Cost-effectiveness
Reoperation	

a Identified as core by patient meeting and healthcare professional Delphi survey.
b Identified as core by healthcare professional Delphi survey only.
c Identified as core by patient meeting only.
d Outcome relevant for surgical interventions only.

effectiveness trials, like many other COS (both within surgery and other specialties), it could have value for research and clinical practice.[122,123] However, further work to determine international relevance, identification of measures to collect the outcomes, and further validation with additional stakeholders is also of importance.

Core outcomes set focus on the 'what' of outcome measurement and reporting rather than the 'how' which would include defining time points at which the measurement should be made. Therefore, future research should determine how to measure the outcomes identified in this COS. Some of this work has been conducted. A recent systematic review assessing the methodological quality of PROMs in patients undergoing laparoscopic cholecystectomy identified six validated measures.[124] That review was not able to recommend a specific PROM for use in laparoscopic cholecystectomy due to the limited number of studies and poor quality of the measures identified.[124] A more recent review also identified considerable variation in the measurement and reporting of patient-reported outcomes after laparoscopic cholecystectomy. This more recent review identified the need for a COS that would incorporate patient-reported outcomes and the consideration of longer-term outcomes.[125]

Strengths and limitations

Given the challenges with recruitment and sustained participation of the Delphi, it will be important to engage with a wide range of healthcare professionals and trialists to ensure this COS for uncomplicated symptomatic gallstone disease is consulted on as widely as possible prior to implementation. Further work may also be required among additional clinical stakeholder groups (e.g. general practitioners, service commissioners) to ensure the core outcomes represent outcomes they would also consider core when evaluating a range of interventions for the treatment of uncomplicated symptomatic gallstones. In addition, determining the transferability of this COS to other settings (i.e. low- and middle-income countries) would also be important to ensure wide applicability and implementation.

A key strength of the work is the extensive outcome mapping exercise on which the COS builds. Even though a small number of patients contributed to the consensus meeting, the outcomes that had been previously identified as of importance to participants were informed by an extensive synthesis of patients reported relevant outcomes through systematic literature review and primary research.[89]

Conclusions

The study has developed the first COS for use in effectiveness trials evaluating interventions for uncomplicated symptomatic gallstone disease. It was prospectively registered on the COMET Initiative database and development, and reporting has been informed by existing standards for COSs.[90,126] The final COS includes 11 outcomes deemed critically important by both patients and healthcare professionals. Further work is required to identify appropriate ways to measure the core outcomes and to verify its findings in a wider stakeholder group.

Chapter 8 Discussion

To the best of our knowledge, the C-GALL trial is the largest multicentre, pragmatic trial to evaluate the clinical and cost-effectiveness of observation/conservative management compared with laparoscopic cholecystectomy to prevent recurrent symptoms and complications in adults presenting with uncomplicated symptomatic gallstones in a secondary care setting.

The primary outcome was QoL measured by AUC at up to 18 months using the SF-36 bodily pain domain. Secondary outcomes included the Otago gallstones CSQ, SF-36 domains (excluding bodily pain), AUC at up to 24 months for the SF-36 bodily pain domain, persistent symptoms (pain and dyspepsia) and complications.

Principal findings

Clinical effectiveness

The trial was conducted in the UK NHS setting, across 20 secondary care sites, and randomised 434 participants. There was no evidence of a difference in the primary outcome of bodily pain at 18 months and no evidence of differences in QoL, complications or in the need for further treatment between the strategies at up to 24 months follow-up. There was statistically significant evidence that gall bladder-specific measures of QoL improved in the cholecystectomy randomised group at 24 months.

Before the trial, it was envisaged that clinicians would be reluctant to recruit patients due to preference in favour of surgery. Moreover, it was thought that patients would be unwilling to consent to recruitment because it was likely that they would feel surgical options were needed after being referred to the hospital for that purpose. During the trial, it was noted that many patients opted for non-surgical treatment after being provided with detailed information about the alternative options available. Almost one-third of patients in the conservative management group subsequently received surgery and 30% of those randomised to surgery had not received surgery by 24 months.

Prespecified sensitivity analyses demonstrated clearly that compliance with treatment allocation, missing data and the impact of COVID-19 did not change the findings.

The profile of complications in the two groups was as expected in routine care. There was no evidence of any difference in total complications between the groups. In the observation/conservative management group, presurgical complications of cholecystitis/biliary colic were the most prevalent complications, and for the surgical management group, bile/stone spillage from gallbladder was the main complication.

Cost-effectiveness

Cost analysis shows that observation/conservative management group was less costly than the laparoscopic cholecystectomy group. The trial did not demonstrate a significant difference in QALYs between the groups. The ICER was found to be high, meaning significant potential savings to the NHS with limited QALY loss by following an observation/conservative management approach in the short term. Longer-term modelling suggested that following an observation/conservative management approach might be cost-effective but there was greater uncertainty due to limited information on subsequent surgeries in the randomised groups and differences in QoL beyond 24 months could reverse this finding. Sensitivity analysis incorporating longer-term QoL scores reduced the potential saving to just £14,700 per QALY lost. The current decision uncertainty could be reduced with a long-term follow-up of the C-GALL trial participants.

Process evaluation

The embedded process evaluation in C-GALL explored ways of improving recruitment and retention across the trial. A total of 16 sites provided 180 audio recordings of consultations for analysis. Analysis of the transcripts identified four core challenge areas for recruiters: (1) providing a balanced presentation about both treatments; (2) discussing and exploring preferences; (3) discussing uncertainty; and (4) discussing participants who did not receive their treatment allocation (crossovers in treatments). A subset of 38 audio recordings of consultations from 4 sites were included in the analysis of discussion of retention. Thirty (79%) of these consultations did not include any discussion of trial retention. Interviews with participants (n = 9) to explore challenges in returning postal questionnaires identified six themes influencing retention: unclear expectations of trial participation; personal attributes for questionnaire completion; significance of questionnaire non-return; commitment to returning questionnaires given other priorities; individual preferences for presentation mode and timing of the questionnaires; and internal and external strategies to encourage questionnaire return. The innovative adoption of a behavioural science approach to the process evaluation led to structured changes in written and verbal information across the trial including e-mail feedback, amendments to trial information leaflet and updates to cover letters and newsletters.

Core outcome set

The final COS for symptomatic uncomplicated gallstone disease included 11 critically important outcomes from both patients and healthcare professionals. These were QoL, overall health state, overall satisfaction, overall pain, common bile duct injury, biliary leak, haemorrhage, need for endoscopic retrograde cholangiopancreatography, intra-abdominal collections, admission/re-admission for problems and reoperation.

Comparison of findings with other studies

Two small Norwegian RCTs, involving 201 participants, demonstrated that 55% of people randomised to observation did not require an operation during the 14-year follow-up period and 12% of people randomised to cholecystectomy did not undergo the scheduled operation.[35] This contrasts with 70% randomised to observation/conservative management not undertaking surgery at 24 months in C-GALL and 30% in the surgery group not undergoing cholecystectomy. Scrutinising (in)efficient use of cholecystectomy: a randomized trial concerning variation in practice (SECURE)[30] conducted a non-inferiority multicentre RCT in the Netherlands of immediate cholecystectomy versus a restrictive strategy, whereby participants underwent cholecystectomy only when the participants fulfilled five prespecified criteria for surgery at any clinic visits. The SECURE trial noted that 30% (non-compliance) of the participants allocated to the restrictive policy underwent cholecystectomy, despite not fulfilling criteria for surgery, and attributed this to either the surgeon or the patient deciding that cholecystectomy was the best therapeutic option despite the outcome of the triage instrument. The SECURE study concluded that the current surgical management of patients with gallstones and abdominal symptoms is suboptimal, that a restrictive policy was not a solution and that physicians need to be more careful in advising a surgical approach in patients with gallstones and abdominal symptoms. The C-GALL trial findings were consistent with the SECURE conclusion and, in addition, provided stronger evidence from a broader range of patients since C-GALL included participants with biliary colic or acute cholecystitis (SECURE only included biliary colic participants).

Strengths of the trial

C-GALL's strengths included the pragmatic RCT design and methodological rigour. The benefit of the sample size is reflected in the precision with which outcomes were estimated. This multicentre trial also gives confidence in the generalisability of findings to the NHS. Our recruited sample had a mean age of 51 years and was representative of patients presenting with uncomplicated symptomatic gallstones in

a secondary care setting.[127] Among our sample, participants were predominantly female, white (86%) with Asian/Asian British (6.9%) and black/African/Caribbean/black British (2.8%). This contrasts well with English/Welsh statistics, 86%, 7.5% and 3.3%, respectively.[128] We are therefore confident that the results were representative of the UK population with gallstones.

This trial was pragmatic where patients in the UK may not always receive the treatment they are offered and waiting lists for surgical treatment existed. We carefully tracked treatment after randomisation and monitored compliance. A major strength included our sensitivity analyses, including compliance analysis, imputation for missing data and potential impact of COVID-19. These analyses did not change our findings.

The cost-effectiveness analysis had several strengths. Firstly, the RCT design allowed the collection of data on resource use and QoL collected prospectively for comparable groups. Secondly, in cost analysis, critical model data were supported by RCT data (e.g. survival analysis; QoL data for the reduction in QoL weight before and after surgery).

We incorporated an embedded process evaluation with qualitative interviews to better understand and reduce recruitment and retention challenges. The process evaluation gave some generalisable findings and recommendations for good practice for ongoing and future trials including how to improve the informed consent and follow-up processes.

A key strength of the COS was the extensive outcome mapping exercise, including both existing and primary research, on which the COS builds. The development of a COS for uncomplicated symptomatic gallstone disease will help to ensure that important outcomes to patients and the NHS are collected in the future.

Reporting equality, diversity and inclusion

As described earlier, the participant population were closely matched in age and ethnicity to the UK population with gallstones. We optimised participation through the use of recordings of screening/recruitment appointments to provide feedback to the recruiters' potential conscious/unconscious biases. The research team included a group of PPI partners from across the UK that were actively involved throughout the study contributing to COS development, study materials and contributing to discussions of the study results with the TSC and the trial investigators as well as preparation of the plain language summary. The study team represented a broad range of expertise in quantitative and qualitative methodologies. Less experienced members of the team were encouraged to lead discrete components of the study (under senior supervision), lead ancillary publications from that work and lead the study's contribution to SWAT. For example, publications related to the COS were authored by less experienced members of the team.

Limitations of the trial

An unexpected difficulty was the longer than expected time on the waiting list for the delivery of surgery for those patients who were allocated to cholecystectomy. When designing the trial, it was anticipated that this wait would be, on average, 6 months. Therefore an 18-month follow-up was chosen as the primary outcome follow-up time to reflect a time equivalent to 12 months after surgery. However, during the study, we observed that patients often experienced longer times to surgery initially due to existing NHS resources that resulted in long waiting lists. To address this, we added a 24-month follow-up time point. Our sensitivity analyses on compliance with the treatment suggested that the waiting list was unlikely to be biasing the study findings. The existence of the waiting list may limit the generalisability to some other countries' jurisdictions. A further limitation was the non-blinding

of participants and treating surgeons to allocation. In this trial, the pragmatic research question was testing the most effective treatment strategy in real-life setting, leading to an inevitable lack of blinding. Although statistically significant differences in gall bladder-specific measures were found, the importance of the size of the differences was unclear with no previously published estimates of important differences. Nevertheless, considering Cohen's standardised effect sizes of 0.25 of a SD to be important, the observed study differences were larger and about 0.4 of a SD. Lastly, non-responders to the questionnaire tended to be younger and have worse self-reported pain and quality of life at recruitment compared to responders. Such non-response is often found in trials. The MI approach and sensitivity analyses limited the effect of this possible non-response bias.

Implications for health care

Current clinical guidelines recommend offering laparoscopic cholecystectomy for biliary pain or acute cholecystitis and radiological evidence of gallstones.[30] Hence, surgical management remains the default option for people with symptomatic gallstone disease and one of the most common elective surgical procedures performed in the NHS.[30] In the UK, many people with uncomplicated symptomatic gallstone disease are put on a waiting list and operated on electively after several months. The C-GALL trial demonstrates that in adults presenting with uncomplicated symptomatic gallstones in a secondary care setting observation/conservative management may be as effective and cost-effective as laparoscopic cholecystectomy in the short term. The crossover between groups suggests that it remains key to identify patients who will and will not benefit from surgery and that discussion about the options, benefits and risks of both approaches should be part of the consent process.[129]

Implications for research

Costs and benefits will continue to be incurred in both groups beyond 24 months, so future research should focus on (1) long-term follow-up data to establish the lifetime cost-effectiveness (e.g. using a simple cohort design) and (2) identification of the cohort of patients that should be offered surgery.

Conclusion

Overall, our results suggested that in the short term (up to 24 months) observation/conservative management may be a cost-effective use of NHS resources in selected patients but subsequent surgeries in the randomised groups and differences in quality of life beyond 24 months could reverse this finding, and longer-term follow up is needed to verify the safety and cost-effectiveness of this approach.

Additional information

Contributions of authors

Karen Innes (https://orcid.org/0000-0001-8512-4368) (Trial Manager) contributed to the design of the trial, was responsible for the day-to-day management of the trial and contributed to the interpretation of results and writing/editing the report.

Irfan Ahmed (https://orcid.org/0000-0002-2172-4177) (Co-Chief Investigator, Professor) contributed to the conception and design of the trial, was involved in recruitment, recruited, and was responsible for the day-to-day oversight of the trial, interpretation of results and writing/editing the report.

Jemma Hudson (https://orcid.org/0000-0002-6440-6419) (Statistician) conducted the statistical analysis and contributed to the interpretation of results, writing/editing the report and reviewing the final report.

Rodolfo Hernández (https://orcid.org/0000-0003-2619-8230) (Health Economist) contributed to the conception and design of the study, the analysis of the health economics data, the development of the health economics model, the drafting of the health economics chapters, the interpretation of the trial results, plus commented on other chapters of the report.

Katie Gillies (https://orcid.org/0000-0001-7890-2854) (Qualitative Researcher) contributed to the conception and design of the trial, oversaw the process evaluation and core outcome set development and analysis, contributed to the interpretation of results and writing/editing the report.

Rebecca Bruce (https://orcid.org/0000-0001-8508-1206) (Data Co-ordinator) was responsible for aspects of the day-to-day management of the trial and contributed to the interpretation of results and editing the report.

Victoria Bell (https://orcid.org/0000-0001-9518-9296) (Assistant Trial Manager) was responsible for aspects of the day-to-day management of the trial and contributed to the interpretation of results and writing/editing the report.

Alison Avenell (https://orcid.org/0000-0003-4813-5628) (Clinical Chair in Health Services Research) contributed to the conception and design of the trial, the interpretation of results and writing/editing the report.

Jane Blazeby (https://orcid.org/0000-0002-3354-3330) (Professor of Surgery) contributed to the design of the trial, conduct of the trial, interpretation of results and writing/editing the report.

Miriam Brazzelli (https://orcid.org/0000-0002-7576-6751) (Senior Research Fellow) contributed to the conception and design of the trial, the interpretation of results and writing/editing the report.

Seonaidh Cotton (https://orcid.org/0000-0002-7883-0608) (Senior Trail Manager) contributed to oversight of the day-to-day management of the trial and contributed to the interpretation of results and writing/editing the report.

Bernard Croal (https://orcid.org/0000-0002-6375-3507) (Patient and Public Representative) as a named grant holder, contributed to the conception and design of the trial, conduct of the trial, recruitment and follow-up of participants, interpretation of results and writing/editing the report. He also focused on the

patient and public involvement (PPI) components of the trial which fed important information back to the trial.

Mark Forrest (**https://orcid.org/0000-0002-2395-8823**) (Senior IT Development Manager) contributed to the design of the trial software and data curation, interpretation of results and writing/editing the report.

Graeme MacLennan (**https://orcid.org/0000-0002-1039-5646**) (CHaRT Director) contributed to the statistical analysis, the interpretation of results and writing/editing the report.

Peter Murchie (**https://orcid.org/0000-0001-9968-5991**) (Academic GP) contributed to the conception and design of the trial, the interpretation of results and writing/editing the report.

Samantha Wileman (**https://orcid.org/0000-0002-1031-1449**) (Quality Assurance Manager) contributed to the conception and design of the trial, the interpretation of results and writing/editing the report.

Craig Ramsey (**https://orcid.org/0000-0003-4043-7349**) (Co-Chief Investigator, Professor) contributed to the conception and design of the trial, was responsible for the day-to-day oversight of the trial, interpretation of results and writing/editing the report.

Acknowledgements

The authors wish to thank the men and women who participated in C-GALL. We also thank the CHaRT data co-ordinators and trial managers who helped support the study: Zoe Batham, Louise Campbell, Janice Cruden, Dianne Dejean, Jackie Ellington, Andrea Fraser and Bev Smith (Data Co-ordinators), Tracey Davidson and Alison McDonald (Trial Managers).

We are grateful to Kirsty McCormack and John Norrie for their help and advice in developing the grant proposal, to the Programming Team in CHaRT for developing and maintaining the study website. We thank Juliette Snow and Rachael West for their help with contracting, and Louise Cotterell, Kerry Duffus and Anne Buckle for their assistance in managing the budget. Our thanks also go to the Research Governance team (Louise King, Stacey Dawson, Lynn McKay) at the University of Aberdeen for their advice and support during the study.

Thanks to Jamie McAllister (NHS Grampian) for providing unit cost data for the within-trial economic analysis.

Thanks to Chen *et al.* for allowing the C-GALL group the use of the Otago Condition-Specific Questionnaire (OCSQ) for gallstone disease, developed by Chen *et al*. in the University of Otago, New Zealand.[1,2]

1. Chen TY, Landmann MG, Potter JC, van Rij AM. Questionnaire to aid priority and outcomes assessment in gallstone disease. *ANZ J Surg*. 2006;76(7):569–74.

2. Chen TY. A novel set of condition-specific quality of life questionnaires in elective general surgical patient prioritization and outcome assessment [dissertation]. Dunedin (NZ): University of Otago; 2012. Retrieved from http://hdl.handle.net/10523/2588

Members of the PMG for their ongoing support and advice. The independent members of the TSC and DMC, and the staff at the recruiting sites (listed below) who facilitated recruitment, treatment and follow-up of trial participants.

Project Management Group

Craig Ramsay, Irfan Ahmed, Alison Avenell, Victoria Bell, Jane Blazeby, Miriam Brazzelli, Rebecca Bruce, Bernard Croal (Grant Holder PPI Representative), Mark Forrest, Katie Gillies, Rodolfo Hernandez, Jemma Hudson, Karen Innes, Graeme MacLennan, Peter Murchie, John Norrie, Samantha Wileman.

Trial Steering Group

David Beard (Chair), Ian Beckingham (Independent Member), John Leeds (Independent Member TSC), Dee McDonald (Patient Representative) and PMG members.

Data Monitoring Committee

Catherine Hewitt (Chair), Jonathan Lund (Independent Member), Tim McAdam (Independent Member) and Amir Nisar (Independent Member).

Embedded Process Evaluation Team

Eilidh Duncan, Jennifer Dunsmore, Louisa Lawrie, Rumana Newlands.

Core Outcome Set Team

Mohammed Bekheit, Moira Cruickshank.

Patient and Public Involvement (PPI) group

Special thanks to those who participated in the PPI group: Edna Bell, Anna Glassey, Denise Gorman, Carolyn Hutchison, Marion Morton, Alec Paterson, Sheena Rendall, Wendy Stuart, Annamaria Wolf and to Katie Banister for her involvement in the PPI activities.

Members of the C-GALL trial group responsible for recruitment in the clinical centres were as follows:

Aberdeen: James Milburn (Principal Investigator), Irfan Ahmed (Principal Investigator), Momin Malik, Sandra Mann, Amy Nicol (Nee – Angus), Tracy Scott, Lorna Van Lierop

Bedford: Dipankar Chattopadhyay (Principal Investigator), Beena David, Carina Galpin, Mel Penacerrada, Kanapathi Rajaratnam, Retno Walandari

Birmingham: Robert Sutcliffe (Principal Investigator), Elizabeth Cornes, Sharon Garner, Manijeh Ghods, Ewen Griffiths, Ravi Marudanayagam, Andrew McDarby, Claire McNeill, Sonia Puig-Compano, Keith Roberts, Arlo Whitehouse

Borders: Jamie Young (Principal Investigator), Joy Dawson, Dave Gribble, Laura Stephen

Cardiff: David O'Reilly (Principal Investigator), Rezwana Chowdhury, Antonio Foliaki, Martin Grigg, Laura Jones, Lucy Knibbs, Gail Williams, Jolene Witherspoon, Caroline Young

Coventry: Jawad Ahmad (Principal Investigator), Susan Dale, Natalie Kenny, Edward Lopez, Gabriele Marangoni

Durham: Zaher Toumi (Principal Investigator), Sarah Clark, Louise Duncan, Stef Hobson, Andrea Kay, Ami Laidlaw, Diana Nayman, Ami Wilkinson

Fife: Peter Driscoll (Principal Investigator), Karlygash Ausel, Laura Beveridge, Keith Boath, Geert Koffeman, Sue Pick

Glasgow: Simon Gibson (Principal Investigator), Anissa Benchiheub, Carol Dalton, Paul Glen, Steven Henderson, Gillian Rice

Heart of England: Markos Daskalakis (Principal Investigator), Sally Abbott, Safia Begum, Maryam Hussain, Tabinda Kharodia, Julie Markley, Faye Moore, Raj Nijjar, VJ Shabangu, Karim Sillah, Rishi Singhal, Samantha Stafford, Paul Super, Olga Tucker

Liverpool: Christopher Halloran (Principal Investigator), Lorna Fleming-Bird, Julie Haydock, Francesca Westwell

Newport: Andrew Beamish (Principal Investigator), Michael Nutt (Principal Investigator), Ashraf Rasheed (Principal Investigator), Lindianne Aitken, Evelyn Baker, Nancy Hawkings, Georgia Mallison, Heeam Nassa, Grace Okoro, Claire Price, Sandapa Punchihewa, Laura Smith, Jemma Tuffney, Elaine Wall, Cerian Williams

North Tees: Talvinder Gill (Principal Investigator), Mazuin Abutalib, Ish Ahmed, Rajendirin Ashwin, Liz Baker, Leo Brown, Bussa Gopinath, Deena Harji, Siddek Isreb, Aya Musbahi, Milind Rao, Murugappan Shanmugam, Kyaw Toe, Suvi Virupaksha, Debbie Wilson

Nottingham: Ravinder Vohra (Principal Investigator), Tracy Brear, Debra Campion, Jody Carvell, James Catton, David Chadwick, Sam Dutta, Tanvir Hossain, Maria Laye, Kate Marks, Simon Parsons, Shirley Pyke, Judith Rozentals, John Saunders, Nilanjana Tewari, Rhian Warman, Neil Welch, Fady Yanni

Plymouth: David Stell (Principal Investigator), Abigail Patrick, Aleasha Udal, Natasha Wilmshurst

Royal Free: Brian Davidson (Principal Investigator), Christine Eastgate, Helder Filipe, Yvonne Gleeson, Kurinchi Gurusamy

Sandwell: Andrew Torrance (Principal Investigator), Veronica Campbel, Julie Colley, Mokhtar Eltair, Samia Hussain, Yogesh Kumar, Rajnish Mankotia, Jenny Porter, Liam Preece, Elizabeth Williams

Taunton: Marianne Hollyman (Principal Investigator), Sarah Brown, Katrina Butcher, Viktoria Cripps, Rebecca Goodchild, Libby Graham, Ami Mackay, Matthew Mason, Peter Mekhail, Alison Moss, Hamish Noble, Maria O'Leary, Corinne Pawley, Andrew Robertson, Charmaine Shovelton, Muwaffaq Telfah, Frances Venn, Richard Welbourn, Esther Zebracki

Warwick: Paul Marriott (Principal Investigator), Indy Atwal, Bridget Campbell, Vasileios Charalampakis, Angela Day, Angela Doyle, Sophie Mason, Penny Parsons, Alexandra Smith, Karen Spicer, Frances Walsh, Kathryn Webb

Yeovil: John Spearman (Principal Investigator), Joanna Allison, Kate Beesley, Nathan Curtis, Godwin Dennison, Linda Howard, Alison Lewis, Jess Perry

Sites opened but never recruited

Aintree: Hassan Malik (Principal Investigator), Matthew Caffrey, Shirley Cooper, Rafael Diaz-Nieto, Declan Dunne, Joanne Earley, Stephen Fenwick, Melanie Harrison, Tanya Ingram, Robert Jones, Sophie Marsh, Sarah Stevenson, Cath Toohey

Dundee: Iain Tait (Principal Investigator), Mairead Tennent

Patient data statement

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 they are 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

Data-sharing statement

All data requests should be submitted to the corresponding author for consideration. Access to available anonymised data may be granted following review.

Ethics statement

C-GALL received favourable ethics opinion from North of Scotland Research Ethics Committee (REC) A, on 23 May 2016 (REC reference number 16/NS/0053).

Information governance statement

The University of Aberdeen and NHS Grampian are committed to handling all personal information in line with the UK Data Protection Act (2018) and the General Data Protection Regulation (EU GDPR) 2016/679. Under the Data Protection legislation, University of Aberdeen and NHS Grampian are joint Data Controllers, and you can find out more about how we handle personal data, including how to exercise your individual rights and the contact details for our Data Protection Officer here: https://www.abdn.ac.uk/about/privacy/

Disclosure of interests

Full disclosure of interests: Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at https://doi.org/10.3310/MNBY3104.

Primary conflicts of interest: Katie Gillies reports other financial or no financial interest as a member of the NIHR HTA CET Committee since 2020; Jane Blazeby reports grants from NIHR Bristol Biomedical Research Centre and being a member of the NIHR Clinical Trials Unit Standing Advisory Committee 2015–9; Seonaidh Cotton is a coinvestigator on unrelated grants from NIHR (HTA and EME programmes: NIHR129819, 15/130/95, 15/130/20) for which her institution has received payment; Bernard Croal reports a Leadership or fiduciary role in the Association of Clinical Biochemistry and Laboratory medicine as president 2021–2023 and Royal College of Pathologists as Trustee and Scottish Chair; Craig Ramsay reports no financial interest as a member of the NIHR HTA General Committee 2017–22.

Additional funding

Chapter 7 – Core Outcome Set: This work was also supported by NHS Grampian Endowment (project grant no. 16/11/006). KG held a Medical Research Council UK Methodology Fellowship during the delivery of this project (MR/L01193X/1).

Publications

Tunji-Ajayi P, Duncan EM, Gillies K. An embedded mixed-methods study highlighted a lack of discussions on retention in clinical trial consultations. *J Clin Epidemiol*. 2020;**123**:49–58. https://doi.org/10.1016/j.jclinepi.2020.03.011

Ahmed A, Innes K, Brazzelli M, *et al.* Protocol for a randomised controlled trial comparing laparoscopic cholecystectomy with observation/conservative management for preventing recurrent symptoms and complications in adults with uncomplicated symptomatic gallstones (C-GALL trial). *BMJ Open* 2021;**11**:e039781.

Cruickshank M, Newlands R, Blazeby J, *et al.* Identification and categorisation of relevant outcomes for symptomatic uncomplicated gallstone disease: in-depth analysis to inform the development of a core outcome set. *BMJ Open* 2021;**11**:e045568.

Lawrie L, Duncan EM, Dunsmore J, *et al*. Using a behavioural approach to explore the factors that affect questionnaire return within a clinical trial: a qualitative study based on the theoretical domains framework. *BMJ Open* 2021;**11**:e048128. https://doi.org/10.1136/bmjopen-2020-048128

Innes K, Hudson J, Banister K, *et al.* Core outcome set for symptomatic uncomplicated gallstone disease. *Br J Surg* 2022;**109**:539–44. https://doi.org/10.1093/bjs/znac095

References

1. Ahmed I, Innes K, Brazzelli M, Gillies K, Newlands R, Avenell A, *et al*. Protocol for a randomised controlled trial comparing laparoscopic cholecystectomy with observation/conservative management for preventing recurrent symptoms and complications in adults with uncomplicated symptomatic gallstones (C-Gall trial). *BMJ Open* 2021;**11**(3):e039781.
2. American College of Physicians. Guidelines for the treatment of gallstones. *Ann Intern Med* 1993;**119**(7 Pt. 1):620–2.
3. Portincasa P, Moschetta A, Palasciano G. Cholesterol gallstone disease. *Lancet* 2006;**368**(9531):230–9.
4. Shaffer EA. Gallstone disease: epidemiology of gallbladder stone disease. *Best Pract Res Clin Gastroenterol* 2006;**20**(6):981–96.
5. Stinton LM, Myers RP, Shaffer EA. Epidemiology of gallstones. *Gastroenterol Clin North Am* 2010;**39**(2):157–69, vii.
6. Bateson MC. Gallstones and cholecystectomy in modern Britain. *Postgrad Med J* 2000;**76**(901):700–3.
7. Barbara L, Sama C, Morselli Labate AM, Taroni F, Rusticali AG, Festi D, *et al*. A population study on the prevalence of gallstone disease: the Sirmione study. *Hepatology* 1987;**7**(5):913–7.
8. Everhart JE, Khare M, Hill M, Maurer KR. Prevalence and ethnic differences in gallbladder disease in the United States. *Gastroenterology* 1999;**117**(3):632–9.
9. Heaton KW, Braddon FE, Mountford RA, Hughes AO, Emmett PM. Symptomatic and silent gall stones in the community. *Gut* 1991;**32**(3):316–20.
10. Williams JG, Roberts SE, Ali MF, Cheung WY, Cohen DR, Demery G, *et al*. Gastroenterology services in the UK: the burden of disease, and the organisation and delivery of services for gastrointestinal and liver disorders: a review of the evidence. *Gut* 2007;**56**(Suppl. 1):1–113.
11. Festi D, Dormi A, Capodicasa S, Staniscia T, Attili AF, Loria P, *et al*. Incidence of gallstone disease in Italy: results from a multicenter, population-based Italian study (the MICOL project). *World J Gastroenterol* 2008;**14**(34):5282–9.
12. Festi D, Reggiani MLB, Attili AF, Loria P, Pazzi P, Scaiolo E, *et al*. Natural history of gallstone disease: expectant management or active treatment? Results from a population-based cohort study. *J Gastroenterol Hepatol* 2010;**25**(7):719–24.
13. Sakorafas GH, Milingos D, Peros G. Asymptomatic cholelithiasis: is cholecystectomy really needed? A critical reappraisal 15 years after the introduction of laparoscopic cholecystectomy. *Dig Dis Sci* 2007;**52**(5):1313–25.
14. Beckingham I. ABC of diseases of liver, pancreas, and biliary system: gallstone disease. *BMJ* 2001;**322**:91–4.
15. Friedman GD, Raviola CA, Fireman B. Prognosis of gallstones with mild or no symptoms: 25 years of follow-up in a health maintenance organization. *J Clin Epidemiol* 1989;**42**(1986):127–36.
16. Thistle J, Cleary P, Lachin J, Tyor M, Hersh T. The natural history of cholelithiasis: the National Cooperative Gallstone Study. *Ann Intern Med* 1984;**101**:171–5.
17. Attili AF, De Santis A, Capri R, Repice AM, Maselli S. The natural history of gallstones: the GREPCO experience. The GREPCO Group. *Hepatology* 1995;**21**(3):655–60.

18. McSherry CK, Ferstenberg H, Calhoun WF, Lahman E, Virshup M. The natural history of diagnosed gallstone disease in symptomatic and asymptomatic patients. *Ann Surg* 1985;**202**(1):59–63.

19. Berhane T, Vetrhus M, Hausken T, Olafsson S, Sondenaa K. Pain attacks in non-complicated and complicated gallstone disease have a characteristic pattern and are accompanied by dyspepsia in most patients: the results of a prospective study. *Scand J Gastroenterol* 2006;**41**(1):93–101.

20. Schmidt M, Dumot JA, Soreide O, Sondenaa K. Diagnosis and management of gallbladder calculus disease. *Scand J Gastroenterol* 2012;**47**(11):1257–65.

21. CholeS Study Group. West Midlands Research Collaborative: population-based cohort study of outcomes following cholecystectomy for benign gallbladder diseases. *Br J Surg* 2016;**103**(12):1704–15.

22. Larsen TK, Qvist N. The influence of gallbladder function on the symptomatology in gallstone patients, and the outcome after cholecystectomy or expectancy. *Dig Dis Sci* 2007;**52**(3):760–3.

23. Jorgensen T, Teglbjerg JS, Wille-Jorgensen P, Bille T, Thorvaldsen P. Persisting pain after cholecystectomy: a prospective investigation. *Scand J Gastroenterol* 1991;**26**(1):124–8.

24. Ahmed R, Freeman JV, Ross B, Kohler B, Nicholl JP, Johnson AG. Long term response to gallstone treatment: problems and surprises. *Eur J Surg* 2000;**166**(6):447–54.

25. Lamberts MP, Lugtenberg M, Rovers MM, Roukema AJ, Drenth JP, Westert GP, van Laarhoven CJHM. Persistent and de novo symptoms after cholecystectomy: a systematic review of cholecystectomy effectiveness. *Surg Endosc* 2013;**27**(3):709–18.

26. Girometti R, Brondani G, Cereser L, Como G, Del Pin M, Bazzocchi M, Zuiani C. Post-cholecystectomy syndrome: spectrum of biliary findings at magnetic resonance cholangiopancreatography. *Br J Radiol* 2010;**83**(988):351–61.

27. Luman W, Adams WH, Nixon SN, Mcintyre IM, Hamer-Hodges D, Wilson G, Palmer KR. Incidence of persistent symptoms after laparoscopic cholecystectomy: a prospective study. *Gut* 1996;**39**(6):863–6.

28. Schmidt M, Sondenaa K, Dumot JA, Rosenblatt S, Hausken T, Ramnefjell M, *et al*. Post-cholecystectomy symptoms were caused by persistence of a functional gastrointestinal disorder. *World J Gastroenterol* 2012;**18**(12):1365–72.

29. Bisgaard T, Rosenberg J, Kehlet H. From acute to chronic pain after laparoscopic cholecystectomy: a prospective follow-up analysis. *Scand J Gastroenterol* 2005;**40**(11):1358–64.

30. van Dijk AH, de Reuver PR, Besselink MG, van Laarhoven KJ, Harrison EM, Wigmore SJ, *et al*. Assessment of available evidence in the management of gallbladder and bile duct stones: a systematic review of international guidelines. *HPB (Oxford)* 2017;**19**(4):297–309.

31. NHS England. 2019/20 National Cost Collection for the NHS data (Version 2). URL: www.england.nhs.uk/publication/2019-20-national-cost-collection-data-publication/?msclkid=a-717fabdba4111ecae08168a6a0b8189 (accessed 1 December 2021).

32. National Institute for Health and Care Excellence. Adult laparoscopic cholecystectomy 01/09/2010-31/10/2020. https://doi.org/10.5061/dryad.k9q7h

33. Somasekar K, Shankar PJ, Foster ME, Lewis MH. Costs of waiting for gall bladder surgery. *Postgrad Med J* 2002;**78**(925):668–9.

34. Schmidt M, Hausken T, Glambek I, Schleer C, Eide GE, Sondenaa K. A 24-year controlled follow-up of patients with silent gallstones showed no long-term risk of symptoms or adverse events leading to cholecystectomy. *Scand J Gastroenterol* 2011;**46**(7–8):949–54.

35. Brazzelli M, Cruickshank M, Kilonzo M, Ahmed I, Stewart F, McNamee P, *et al*. Clinical effectiveness and cost-effectiveness of cholecystectomy compared with observation/conservative management for preventing recurrent symptoms and complications in adults presenting with uncomplicated symptomatic gallstones or cholecystitis: a systematic review and economic evaluation. *Health Technol Assess* 2014;**18**(55):1–101, v.
36. StataCorp. *Stata Statistical Software: Release 16*. College Station, TX: StataCorp LLC; 2019.
37. Zou G. A modified Poisson regression approach to prospective studies with binary data. *Am J Epidemiol* 2004;**159**(7):702–6.
38. WHO. WHO Director-General's opening remarks at the media briefing on COVID-19: 11 March 2020. URL: www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 (accessed 4 November 2022).
39. Treweek S, Gillies K, Innes K, MacLennan G. Sending Christmas cards to trial participants to improve retention. *Trials* 2017;**18**:406. URL: www.qub.ac.uk/sites/TheNorthernIrelandNetworkforTrialsMethodologyResearch/FileStore/Filetoupload,846275,en.pdf (accessed 4 November 2022).
40. Coleman E, Arundel C, Clark L, Doherty L, Gillies K, Hewitt C, *et al*. Bah humbug! Association between sending Christmas cards to trial participants and trial retention: randomised study within a trial conducted simultaneously across eight host trials. *BMJ* 2021;**375**:e067742.
41. Duncan A, Innes K, Goulao B, Ramsay C. SWAT 145: The sticker SWAT: does a trial logo sticker placed on the outside of the envelope encourage the return of postal questionnaires. URL: www.qub.ac.uk/sites/TheNorthernIrelandNetworkforTrialsMethodologyResearch/FileStore/Filetoupload,1140958,en.pdf (accessed 4 November 2022).
42. Drummond M, Sculpher M, Claxton K, Stoddart G, Torrance G. *Methods for the Economic Evaluation of Health Care Programmes*. 4th edn. Oxford: Oxford University Press; 2015.
43. Glick HA, Doshi JA, Sonnad SS, Polsky D. *Economic Evaluation in Clinical Trials*. 2nd edn. Oxford: Oxford University Press; 2014.
44. Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs AH, Carswell C, *et al*. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. *Appl Health Econ Health Policy* 2022;**20**(2):213–21.
45. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002;**21**(2):271–92.
46. School of Health and Related Research. *Measuring and Valuing Health*. URL: www.sheffield.ac.uk/scharr/research/themes/valuing-health (accessed 4 November 2022).
47. NHS Digital. *Hospital Admitted Patient Care Activity 2020–21: Procedures and interventions*. URL: https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity/2020-21 (accessed 4 November 2022).
48. Cooper K, Breeman S, Scott NW, Scotland G, Hernández R, Clark TJ, *et al*. Laparoscopic supracervical hysterectomy compared with second-generation endometrial ablation for heavy menstrual bleeding: the HEALTH RCT. *Health Technol Assess* 2019;**23**(53):1–108.
49. Public Health Scotland. *Data and Intelligence*. URL: https://beta.isdscotland.org/topics/finance/file-listings-fy-2019-to-2020/ (accessed 4 November 2022).
50. NHS Institute for Innovation and Improvement. *Improving Quality and Efficiency in the Operating Theatre*. URL: https://alesi-surgical.com/wp-content/uploads/2019/09/Improving-quality-and-efficiency-in-the-operating-theatre.pdf (accessed 4 November 2022).

51. NHS Improvement. *Archived Reference Costs*. URL: https://webarchive.nationalarchives.gov.uk/ukgwa/20200501111106/https://improvement.nhs.uk/resources/reference-costs/ (accessed 4 November 2022).
52. Curtis L, Burns A. *Unit Costs of Health and Social Care 2020*. 2020th edn. Canterbury: Personal Social Services Research Unit, University of Kent; 2020.
53. National Institute for Health and Care Excellence. *British National Formulary*. URL: http://bnf.nice.org.uk/drug/ (accessed 4 November 2022).
54. Cooper K, Breeman S, Scott NW, Scotland G, Clark J, Hawe J, *et al.*; HEALTH Study Group. Laparoscopic supracervical hysterectomy versus endometrial ablation for women with heavy menstrual bleeding (HEALTH): a parallel-group, open-label, randomised controlled trial. *Lancet* 2019;**394**(10207):1425–36.
55. *EUROQOL*. URL: https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/faqs-2/ (accessed September 2022).
56. StataCorp. *Stata Statistical Software: Release 17*. College Station, TX: StataCorp LLC; 2021.
57. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley; 1987.
58. Brand J, van Buuren S, le Cessie S, van den Hout W. Combining multiple imputation and bootstrap in the analysis of cost-effectiveness trial data. *Stat Med* 2019;**38**(2):210–20.
59. National Institute for Health and Care Excellence. *Technology Appraisal Processes*. URL: www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-technology-appraisal-guidance/process (accessed 4 November 2022).
60. Schmidt M, Sondenaa K, Vetrhus M, Berhane T, Eide GE. A randomized controlled study of uncomplicated gallstone disease with a 14-year follow-up showed that operation was the preferred treatment. *Dig Surg* 2011;**28**(4):270–6.
61. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal*. National Institute for Health and Care Excellence, 2013. URL: www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781?msclkid=015aa049ba4211eca67d6ef9d3815df7 (accessed December 2017).
62. Gary A, Clarke P, Wolstenholme S, Wordsworth S. *Applied Methods of Cost-effective Analysis in Health Care*. Oxford: Oxford University Press; 2011.
63. Office for National Statistics. *National Life Tables – Life Expectancy in the UK: 2018 to 2020*. 23 September 2021. URL: www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/2018to2020?msclkid=599e4de4ba4211ec9b9ec1408d1c6ed5/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/2018to2020 (accessed 12 November 2022).
64. van den Berg B. Sf-6D population norms. *Health Econ* 2012;**21**(12):1508–12.
65. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value Health* 2010;**13**(5):509–18.
66. National Institute for Health and Care Excellence. *Gallstone Disease: Diagnosis and Management [CG188]*. 2014. URL: www.nice.org.uk/guidance/cg188 (accessed 4 November 2022).
67. Lawrie L, Duncan E, Dunsmore J, Newlands R, Gillies K. Using a behavioural approach to explore the factors that affect questionnaire return within a clinical trial: a qualitative study based on the theoretical domains framework. 2021. URL: https://bmjopen.bmj.com/content/11/4/e048128 (accessed 4 November 2022).

68. Tunji-Ajayi P, Duncan EM, Gillies K. An embedded mixed-methods study highlighted a lack of discussions on retention in clinical trial consultations. *J Clin Epidemiol* 2020;**123**:49–58.

69. Abrams P, Shearer K. The MASTER trial: artificial urinary sphincter versus male sling. *Trends Urol Men's Health*. 2015;**6**(2):37–8.

70. Donovan J, Mills N, Smith M, Brindle L, Jacoby A, Peters T, *et al*. Quality improvement report: improving design and conduct of randomised trials by embedding them in qualitative research: ProtecT (prostate testing for cancer and treatment) study. Commentary: presenting unbiased information to patients can be difficult. *BMJ* 2002;**325**(7367):766–70.

71. Rogers CA, Welbourn R, Byrne J, Donovan JL, Reeves BC, Wordsworth S, *et al*. The By-Band study: gastric bypass or adjustable gastric band surgery to treat morbid obesity: study protocol for a multi-centre randomised controlled trial with an internal pilot phase. *Trials* 2014;**15**:53.

72. Donovan JL, Rooshenas L, Jepson M, Elliott D, Wade J, Avery K, *et al*. Optimising recruitment and informed consent in randomised controlled trials: the development and implementation of the Quintet Recruitment Intervention (QRI). *Trials* 2016;**17**(1):283.

73. Gillies K, Brehaut J, Coffey T, Duncan EM, Francis JJ, Hey SP, *et al*. How can behavioural science help us design better trials? *Trials* 2021;**22**(1):882.

74. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. *Implement Sci* 2012;**7**:37.

75. Newlands R, Duncan E, Presseau J, Treweek S, Lawrie L, Bower P, *et al*. Why trials lose participants: a multitrial investigation of participants' perspectives using the theoretical domains framework. *J Clin Epidemiol* 2021;**137**:1–13.

76. Lalu MM, Foster M, Presseau J, Dowlatshahi D, Castillo G, Cardenas A, *et al*. What are potential barriers and enablers to patient and physician participation in Canadian cell therapy trials for stroke? A stakeholder interview study. *BMJ Open* 2020;**10**(3):e034354.

77 Castillo G, Lalu MM, Asad S, Foster M, Kekre N, Fergusson DA, *et al*. Navigating choice in the face of uncertainty: using a theory informed qualitative approach to identifying potential patient barriers and enablers to participating in an early phase chimeric antigen receptor T (CAR-T) cell therapy trial. *BMJ Open* 2021;**11**(3):e043929.

78. Lawrie L, Duncan EM, Jansen JO, Campbell MK, Brunsdon D, Skea Z, *et al*. *Behavioural Optimisation to Address Trial Conduct Challenges: Case Study in the UK-REBOA Trial (Under Review with Trials)*. URL: www.researchsquare.com/article/rs-1166980/v1 (accessed 4 November 2022).

79. Atkins L, Francis J, Islam R, O'Connor D, Patey A, Ivers N, *et al*. A guide to using the theoretical domains framework of behaviour change to investigate implementation problems. *Implement Sci* 2017;**12**(1):77.

80. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, *et al*. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013;**46**(1):81–95.

81. Wilson C, Rooshenas L, Paramasivan S, Elliott D, Jepson M, Strong S, *et al*. Development of a framework to improve the process of recruitment to randomised controlled trials (RCTs): the SEAR (Screened, Eligible, Approached, Randomised) framework. *Trials* 2018;**19**(1):50.

82. Rooshenas L, Elliott D, Wade J, Jepson M, Paramasivan S, Strong S, *et al*.; ACST-2 Study Group. Conveying equipoise during recruitment for clinical trials: qualitative synthesis of clinicians' practices across six randomised controlled trials. *PLOS Med* 2016;**13**(10):e1002147.

83. Donovan JL, Paramasivan S, de Salis I, Toerien M. Clear obstacles and hidden challenges: understanding recruiter perspectives in six pragmatic randomised controlled trials. *Trials* 2014;**15**:5.

84. Paramasivan S, Strong S, Wilson C, Campbell B, Blazeby JM, Donovan JL. A simple technique to identify key recruitment issues in randomised controlled trials: Q-QAT – quanti-qualitative appointment timing. *Trials* 2015;**16**:88.

85. Malterud K, Siersma VD, Guassora AD. Sample Size in Qualitative Interview Studies: Guided by Information Power. *Qual Health Res* 2016;**26**(13):1753–60.

86. Goulao B, Duncan A, Floate R, Clarkson J, Ramsay C. Three behavior change theory-informed randomized studies within a trial to improve response rates to trial postal questionnaires. *J Clin Epidemiol* 2020;**122**:35–41.

87. Kearney A, Rosala-Hallas A, Bacon N, Daykin A, Shaw ARG, Lane AJ, *et al*. Reducing attrition within clinical trials: the communication of retention and withdrawal within patient information leaflets. *PLOS ONE* 2018;**13**(10):e0204886.

88. Innes K, Cotton S, Campbell MK, Elliott J, Gillies K. Relative importance of informational items in participant information leaflets for trials: a Q-methodology approach. *BMJ Open* 2018;**8**(9):e023303.

89. Cruickshank M, Newlands R, Blazeby J, Ahmed I, Bekheit M, Brazzelli M, *et al*. Identification and categorisation of relevant outcomes for symptomatic uncomplicated gallstone disease: in-depth analysis to inform the development of a core outcome set. *BMJ Open* 2021;**11**(6):e045568.

90. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, Tugwell P. Developing core outcome sets for clinical trials: issues to consider. *Trials* 2012;**13**:132.

91. Davis K, Gorst SL, Harman N, Smith V, Gargon E, Altman DG, *et al*. Choosing important health outcomes for comparative effectiveness research: an updated systematic review and involvement of low and middle income countries. *PLOS ONE* 2018;**13**(2):e0190695.

92. Comet Initiative. *Development of Core Outcome Set for Symptomatic Uncomplicated Gallstone Disease*. URL: www.comet-initiative.org/Studies/Details/927 (accessed 4 November 2022).

93. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, *et al*. *The COMET Handbook*: version 1.0. *Trials* 2017;**18**(Suppl. 3):280.

94. Gurusamy KS, Samraj K, Fusai G, Davidson BR. Early versus delayed laparoscopic cholecystectomy for biliary colic. *Cochrane Database Syst Rev* 2008;**2008**(4):CD007196.

95. Gurusamy KS, Koti R, Fusai G, Davidson BR. Early versus delayed laparoscopic cholecystectomy for uncomplicated biliary colic. *Cochrane Database Syst Rev* 2013;**2013**(6):CD007196.

96. Gurusamy KS, Davidson C, Gluud C, Davidson BR. Early versus delayed laparoscopic cholecystectomy for people with acute cholecystitis. *Cochrane Database Syst Rev* 2013;**2013**(6):CD005440.

97. Gurusamy KS, Samraj K, Fusai G, Davidson BR. Robot assistant versus human or another robot assistant in patients undergoing laparoscopic cholecystectomy. *Cochrane Database Syst Rev* 2012;**2012**(9):CD006578.

98. Gurusamy KS, Vaughan J, Ramamoorthy R, Fusai G, Davidson BR. Miniports versus standard ports for laparoscopic cholecystectomy. *Cochrane Database Syst Rev* 2013;**2013**(8):CD006804.

99. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res* 2005;**15**(9):1277–88.

100. Hewitt-Taylor J. Use of constant comparative analysis in qualitative research. *Nurs Stand* 2001;**15**(42):39–42.

101. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Res Methodol* 2008;**8**:45.

102. Macefield RC, Jacobs M, Korfage IJ, Nicklin J, Whistance RN, Brookes ST, *et al*. Developing core outcomes sets: methods for identifying and including patient-reported outcomes (PROs). *Trials* 2014;**15**:49.

103. Hopkins JC, Howes N, Chalmers K, Savovic J, Whale K, Coulman KD, *et al*.; By-Band Trial Management Group. Outcome reporting in bariatric surgery: an in-depth analysis to inform the development of a core outcome set, the BARIACT study. *Obes Rev* 2015;**16**(1):88–106.

104. Coulman KD, Hopkins J, Brookes ST, Chalmers K, Main B, Owen-Smith A, *et al*.; BARIACT Working Group. A core outcome set for the benefits and adverse events of bariatric and metabolic surgery: the BARIACT project. *PLOS Med* 2016;**13**(11):e1002187.

105. Comet Initiative. *DelphiManager*. URL: www.comet-initiative.org/delphimanager/ (accessed 4 November 2022).

106. Okoli C, Pawlowski SD. The Delphi method as a research tool: an example, design considerations and applications. *Inform Manag* 2004;**42**(1):15–29.

107. Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CF, Askham J, *et al*. Consensus development methods, and their use in clinical guideline development. *Health Technol Assess* 1998;**2**(3):i1–iv88.

108. Aprea G, Coppola Bottazzi E, Guida F, Masone S, Persico G. Laparoendoscopic single site (LESS) versus classic video-laparoscopic cholecystectomy: a randomized prospective study. *J Surg Res* 2011;**166**(2):e109–12.

109. Artis T, Kucuk C, Akay A, Zararsiz G, Sozuer E. *Prospective Randomized Study Comparing Single Incision vs. Standard Laparoscopic Cholecystectomy*. URL: www.cochranelibrary.com/central/doi/10.1002/central/CN-01063839/full# (accessed 4 November 2022).

110. UMIN-CTR Clinical Trial. *Randomized Controlled Trial to Evaluate the Superiority of the Single Port Chlecystectomy (SPC) on Reduction of Postoperative Pain, Compared with the Conventional Laparoscopic Cholecystectomy at the Single Institution*. URL: https://center6.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000003103 (accessed 4 November 2022).

111. ClinicalTrials.gov. *Single Port Laparoscopic Cholecystectomy Versus Four Port Laparoscopic Cholecystectomy: Impact on Postoperative Pain*. URL: https://clinicaltrials.gov/ct2/show/NCT01348620 (accessed 4 November 2022).

112. Bignell M, Lewis MP, Cheong EC, Rhodes M. A prospective, randomized, single-blind trial of 5-mm versus 3-mm ports for laparoscopic cholecystectomy: is smaller better? *Surg Endosc* 2013;**27**(10):3616–21.

113. Bingener J, Skaran P, McConico A, Novotny P, Wettstein P, Sletten DM, *et al*. A double-blinded randomized trial to compare the effectiveness of minimally invasive procedures using patient-reported outcomes. *J Am Coll Surg* 2015;**221**(1):111–21.

114. Russell ML, Preshaw RM, Brant RF, Bultz BD, Page SA. Disease-specific quality of life: the Gallstone Impact Checklist. *Clin Invest Med* 1996;**19**(6):453–60.

115. Chen TY, Landmann MG, Potter JC, van Rij AM. Questionnaire to aid priority and outcomes assessment in gallstone disease. *ANZ J Surg* 2006;**76**(7):569–74.

116. Eypasch E, Wood-Dauphinée S, Williams JI, Ure B, Neugebauer E, Troidl H. The Gastrointestinal Quality of Life Index: a clinical index for measuring patient status in gastroenterologic surgery. *Chirurg* 1993;**64**(4):264–74.

117. Kleinbeck SV, Hoffart N. Outpatient recovery after laparoscopic cholecystectomy. *AORN J* 1994;**60**(3):394–402.

118. Urbach DR, Harnish JL, McIlroy JH, Streiner DL. A measure of quality of life after abdominal surgery. *Qual Life Res* 2006;**15**(6):1053–61.

119. Lindseth GN, Denny DL. Patients' experiences with cholecystitis and a cholecystectomy. *Gastroenterol Nurs* 2014;**37**(6):407–14.

120. Psaila J, Agrawal S, Fountain U, Whitfield T, Murgatroyd B, Dunsire MF, *et al*. Day-surgery laparoscopic cholecystectomy: factors influencing same-day discharge. *World J Surg* 2008;**32**(1):76–81.

121. Comet Initiative. *Advanced Search*. URL: www.comet-initiative.org/Studies (accessed 12 November 2021).

122. Remus A, Smith V, Gutke A, Mena JJS, Mørkved S, Wikmar LN, *et al*. A core outcome set for research and clinical practice in women with pelvic girdle pain: PGP-COS. *PLOS ONE* 2021;**16**(2):e0247466.

123. Chiarotto A, Ostelo RW, Turk DC, Buchbinder R, Boers M. Core outcome sets for research and clinical practice. *Braz J Phys Ther* 2017;**21**(2):77–84.

124. Daliya P, Gemmill EH, Lobo DN, Parsons SL. A systematic review of patient reported outcome measures (PROMs) and quality of life reporting in patients undergoing laparoscopic cholecystectomy. *Hepatobiliary Surg Nutr* 2019;**8**(3):228–45.

125. Alexander HC, Nguyen CH, Moore MR, Bartlett AS, Hannam JA, Poole GH, Merry AF. Measurement of patient-reported outcomes after laparoscopic cholecystectomy: a systematic review. *Surg Endosc* 2019;**33**(7):2061–71.

126. Comet Initiative. *Identification and Categorisation of Relevant Outcomes for Symptomatic Uncomplicated Gallstone Disease: In-Depth Analysis to Inform the Development of a Core Outcome Set*. URL: www.comet-initiative.org/Studies/Details/1909 (accessed 4 November 2022).

127. SWORD Methods. *SWORD Database Accessed Through AUGIS Website (Due to Lack of Funding This Database Closed in 2020)*. 2019.

128. GOV.UK. *Population of England and Wales*. URL: www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/national-and-regional-populations/population-of-england-and-wales/latest (accessed 8 March 2022).

129. Sacks GD, Dawes AJ, Ettner SL, Brook RH, Fox CR, Maggard-Gibbons M, *et al*. Surgeon perception of risk and benefit in the decision to operate. *Ann Surg* 2016;**264**(6):896–903.

Appendix 1 Summary of breaches in the C-GALL trial

TABLE 36 Summary of breaches

Breach number, date	Summary of breach
1, 16 November 2016	Breach to inclusion/exclusion criteria. Breach assessed as non-serious and closed out on 16 December 2016.
2, 26 February 2021	Double randomisation. Participant were randomised at both Birmingham (site 30) and Sandwell (site 35) which are both hospitals based in Birmingham. Participant was randomised to surgical management in both instances. Breach assessed as non-serious and closed out on 1 March 2021.

Appendix 2 Summary of amendments to the C-GALL trial

TABLE 37 Summary of amendments

Protocol version, date	Summary of revisions
1, 8 April 2016	New document
2, 20 May 2016	Amendments to: • Amended as part of REC provisional opinion
3, 10 October 2016	Amendments to: • Revision to protocol following initial recruitment and first trial steering committee meeting including • Administrative changes • Clarification around exclusion criteria to include time periods • Section 3.5.1 Identifying participants – updated to reflect the changes made to Appendix 1 Qualitative study section • Section 4 Safety – the safety section has been revised to ensure clarity for research staff at sites, making it clear what requires reporting • Section 6 Schedule of data collection – schedule of data collection table updated to reflect the flow diagram • Section 7 Recruitment projections – recruitment projection graph updated due to the change in start date • Appendix 1 Embedded Qualitative Evaluation – updates to the qualitative study section
4, 14 June 2017	Amendments to: • Revision to document replacing conservative/observation management to medical management throughout.
5, 14 December 2017	Amendments to: • Revisions to protocol relating to consent procedures for audio recording of recruitment conversations (Eilidh Duncan) • Updates to safety section (see *Section 4*) • Reorder reference sections
6, 15 March 2019	Amendments to: • Revision to follow-up period • Revision to recruitment period
7, 8 August 2019	Amendments to: • Addition of text about questionnaire reminders
8, 11 October 2019	Amendments to: • Revisions to Gannt chart in protocol. Extending recruitment by 1 month in the Aberdeen site
9, 16 January 2020	Amendments to: • Update to Appendix 2 – COS • The COS will now be generated using a Delphi instead of a face-to-face meeting
10, 7 May 2020	Amendments to: • Update to the way PQ data are collected during follow-up
11, 6 July 2020	Amendments to: • Add in case note review at 24 months
12, 9 February 2021	Amendments to: • Add appendix – Sticker SWAT

Appendix 3 Baseline and trial results

TABLE 38 Reasons for ineligibility and preference

Reason	*N* (%)
Ineligible[a]	***N = 647***
Adult patient with unconfirmed symptomatic gallstones not electively referred to a secondary care setting for consultation	233 (36.0)
Clinical diagnosis of symptomatic gallstone disease not confirmed by imaging	144 (22.3)
Medically unfit for surgery	105 (16.2)
Gallstone in common bile duct or evidence of previous choledocholithiasis	87 (13.4)
Evidence of obstructive jaundice	64 (9.9)
Unable to consent	50 (7.7)
Previous open major upper abdominal surgery	45 (7.0)
History of acute pancreatitis	41 (6.3)
Evidence of empyema of the gallstone with sepsis	21 (3.2)
Suspicion of gallbladder cancer	15 (2.3)
Haemolytic disease	15 (2.3)
Currently pregnant	13 (2.0)
Perforated gallbladder	11 (1.7)
Preference	***N = 910***
Patient preferred laparoscopic cholecystectomy	538 (59.1)
Patient preferred to have observation/conservative management	167 (18.4)
Did not want to be randomised	91 (10.0)
Surgeon had preference for laparoscopic cholecystectomy	29 (3.2)
Other reason	24 (2.6)
Preference reason unknown	48 (5.3)
Surgeon had preference for observation/conservative management	13 (1.4)

a More than one reason is possible.

TABLE 39 Response rates for follow-up time points

Time point	Observation/conservative management, *N* = 217		Laparoscopic cholecystectomy, *N* = 217	
	Responded	Provided SF-36 bodily pain outcome	Responded	Provided SF-36 bodily pain outcome
Baseline	217 (100.0)	215 (99.1)	217 (100.0)	216 (99.5)
3 months	176 (81.1)	176 (81.1)	175 (80.6)	174 (80.2)
9 months	145 (66.8)	144 (66.4)	162 (74.7)	160 (73.7)
12 months	156 (71.9)	156 (71.9)	150 (69.1)	149 (68.7)
18 months	168 (77.4)	167 (77.0)	161 (74.2)	161 (74.2)
24 months	136 (62.7)	135 (62.2)	138 (63.6)	138 (63.6)
30 months	132 (60.8)	132 (60.8)	130 (59.9)	130 (59.9)

Note
Values are *n* (%).

TABLE 40 Surgery details up to 18 months

Surgical details	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Received surgery		
Yes	54 (24.9)	146 (67.3)
No	163 (75.1)	71 (32.7)
	Received surgery, *N* = 54	Received surgery, *N* = 146
Time to surgery (months) – median (IQR); *n*	8.1 (4.0–10.6); 53	4.5 (2.7–6.9); 146
Time between surgery and 18 months' follow-up (months) – median (IQR); *n*	9.9 (7.4–13.0); 53	13.5 (11.1–15.5); 146
Length of hospital stay (days) – median (IQR); *n*	1 (0–1); 51	0 (0–1); 143
Operation time (minutes) – median (IQR); *n*	65 (50–97); 47	62 (50–86); 132
Elective surgery		
Yes	46 (85.2)	142 (97.3)
No	6 (11.1)	2 (1.4)
Missing	2 (3.7)	2 (1.4)
Procedure type		
Laparoscopic	51 (94.4)	142 (97.3)
Open	1 (1.9)	1 (0.7)
Laparoscopic converted to open	1 (1.9)	1 (0.7)
Missing	1 (1.9)	2 (1.4)
Grade of operating surgeon		
Consultant	31 (57.4)	95 (65.1)
Consultant supervised by another consultant	4 (7.4)	7 (4.8)

TABLE 40 Surgery details up to 18 months (*continued*)

Surgical details	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Registrar	2 (3.7)	6 (4.1)
Registrar supervised by a consultant	7 (13.0)	19 (13.0)
Specialty (specialty and associate specialist grade) supervised by a consultant	2 (3.7)	–
Senior House Officer supervised by a consultant	1 (1.9)	2 (1.4)
Specialist trainee	–	2 (1.4)
Specialist trainee supervised by a consultant	1 (1.9)	4 (2.7)
Other	–	2 (1.4)
Other supervised by a consultant	2 (3.7)	5 (3.4)
Unknown operating surgeon but supervised by a consultant	–	1 (0.7)
Missing	4 (7.4)	3 (2.1)
Prophylactic antibiotic used in the operation		
Yes	31 (57.4)	77 (52.7)
No	18 (33.3)	61 (41.8)
Missing	5 (9.3)	8 (5.5)
Difficulty of surgery[a]		
Straightforward	28 (51.9)	92 (63.0)
Mildly difficult	5 (9.3)	12 (8.2)
Moderately difficult	3 (5.6)	15 (10.3)
Extremely difficult	3 (5.6)	–
Missing	15 (27.8)	27 (18.5)
Admitted to ICU or HDU		
No	48 (88.9)	138 (94.5)
ICU	–	2 (1.4)
HDU	1 (1.9)	–
Missing	5 (9.3)	6 (4.1)
Time in ICU (hours) – median (IQR); *n*	–	30 (24–36); 2
Time if HDU (hours) – value; *n*	47 (47, 47); 1	–
Required additional pain relief	12 (22.2)	30 (20.5)
Histopathology		

continued

TABLE 40 Surgery details up to 18 months (*continued*)

Surgical details	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Normal gallbladder		
Yes	3 (5.6)	7 (4.8)
No	47 (87.0)	136 (93.2)
Missing	4 (7.4)	3 (2.0)
Cholecystitis for abnormal gallbladder		
No	4 (8.5)	7 (5.2)
Acute	4 (8.4)	5 (3.7)
Chronic	38 (80.9)	124 (91.2)
Missing	1 (2.1)	–
Incidental biliary cancer		
No	50 (92.6)	142 (97.3)
Missing	4 (7.4)	4 (2.7)

a Completed by the operating surgeon.

Note
Values are *n* (%) unless otherwise stated.

TABLE 41 Primary outcome – AUC SF-36 norm-based bodily pain score over 18 months

Primary outcome	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217	Observation/conservative, *N* = 217		Laparoscopic cholecystectomy, *N* = 217	
	ITT	ITT	Complied[a]	Not-complied[b]	Complied[c]	Not-complied[d]
Sensitivity analysis						
Baseline	44.5 (11.7); 199	43.4 (11.2); 201	46.3 (11.2); 140	40.4 (11.9); 59	43.0 (11.1); 146	44.6 (11.6); 55
3 months	44.6 (11.5); 176	42.6 (11.0); 174	46.4 (10.8); 124	40.1 (12.0); 52	42.0 (11.0); 126	44.2 (10.7); 48
9 months	46.6 (11.4); 144	47.9 (12.7); 160	46.0 (10.8); 103	48.1 (12.9); 41	48.7 (12.5); 119	45.6 (13.1); 41
12 months	48.6 (11.6); 156	49.0 (11.4); 149	47.8 (11.0); 118	50.8 (13.3); 38	50.0 (10.6); 112	45.9 (13.1); 37
18 months	49.4 (11.7); 167	50.4 (11.6); 161	48.9 (11.6); 121	50.6 (11.8); 46	51.5 (11.1); 119	47.3 (12.3); 42
AUC over 18 months	46.7 (8.8); 200	46.8 (8.7); 201	47.2 (8.6); 141	45.7 (9.2); 59	47.1 (8.2); 146	45.8 (9.8); 55
MD, 95% CI; *p*-value	−0.0	(−1.8 to 1.7); 0.966			−0.1	(−4.7 to 4.5); 0.968
Complete case analysis						
Baseline	45.0 (11.6); 167	44.2 (11.2); 161	46.6 (11.0); 121	40.8 (12.0); 46	44.0 (10.9); 119	44.6 (12.2); 42
3 months	45.5 (11.5); 147	42.6 (10.4); 145	46.9 (10.9); 106	41.8 (12.2); 41	42.5 (10.4); 106	43.0 (10.6); 39
9 months	46.7 (11.3); 132	47.9 (12.8); 133	46.2 (10.9); 95	48.1 (12.3); 37	49.0 (12.2); 101	44.7 (14.1); 32

TABLE 41 Primary outcome – AUC SF-36 norm-based bodily pain score over 18 months (*continued*)

Primary outcome	Observation/ conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217	Observation/conservative, *N* = 217		Laparoscopic cholecystectomy, *N* = 217	
	ITT	ITT	Complied[a]	Not-complied[b]	Complied[c]	Not-complied[d]
12 months	48.7 (11.5); 147	49.1 (11.2); 130	48.2 (10.9); 112	50.5 (13.5); 35	50.7 (10.1); 98	44.1 (13.0); 32
18 months	49.4 (11.7); 167	50.4 (11.6); 161	48.9 (11.6); 121	50.6 (11.8); 46	51.5 (11.1); 119	47.3 (12.3); 42
AUC over 18 months	47.3 (8.8); 167	47.0 (8.4); 161	47.5 (8.8); 121	46.8 (8.8); 46	47.7 (7.6); 119	44.8 (10.2); 42
MD, 95% CI; *p*-value	0.4	(–1.5 to 2.2); 0.690			0.8	(–3.9 to 5.5); 0.737

a Received observation/conservative management.
b Received laparoscopic cholecystectomy.
c Received laparoscopic cholecystectomy.
d Received observation/conservative management.

Note
Compliance was defined as participants who received their allocated treatment within 24 months (apart from the laparoscopic cholecystectomy group if surgery was an emergency); values are mean (SD); *n* unless otherwise stated.

TABLE 42 Secondary outcome – AUC SF-36 norm-based bodily pain score over 24 months

Primary outcome	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Sensitivity analysis		
Baseline	44.5 (11.7); 200	43.4 (11.2); 203
3 months	44.6 (11.5); 176	42.6 (11.0); 174
9 months	46.6 (11.4); 144	47.9 (12.7); 160
12 months	48.6 (11.6); 156	49.0 (11.4); 149
18 months	49.4 (11.7); 167	50.4 (11.6); 161
24 months	48.0 (12.0); 135	49.1 (12.3); 138
AUC over 24 months	47.2 (8.6); 201	47.4 (8.7); 203
MD, 95% CI; *p*-value	–0.2	(–1.9 to 1.5); 0.803
Complete case analysis		
Baseline	45.3 (11.8); 135	44.0 (10.6); 138
3 months	46.0 (11.5); 124	42.9 (10.7); 121
9 months	46.5 (11.7); 108	47.8 (12.8); 116
12 months	48.1 (11.8); 115	48.4 (11.8); 112
18 months	48.9 (11.9); 125	49.5 (11.7); 127
24 months	48.0 (12.0); 135	49.1 (12.3); 138
AUC over 24 months	47.4 (8.9); 135	47.3 (8.8); 138
MD, 95% CI; *p*-value	0.2	(–1.9 to 2.2); 0.868

Note
Values are mean (SD), *n* unless otherwise stated.

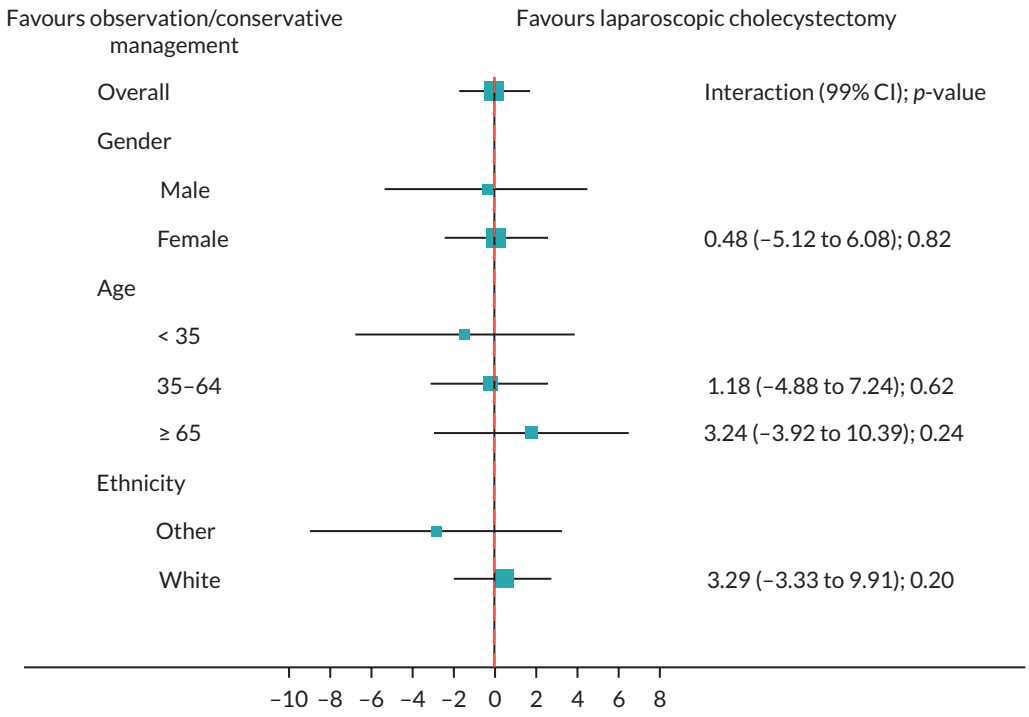

FIGURE 18 Subgroups for observation/conservative management vs. laparoscopic cholecystectomy up to 18 months for SF-36 bodily pain sensitivity analysis.

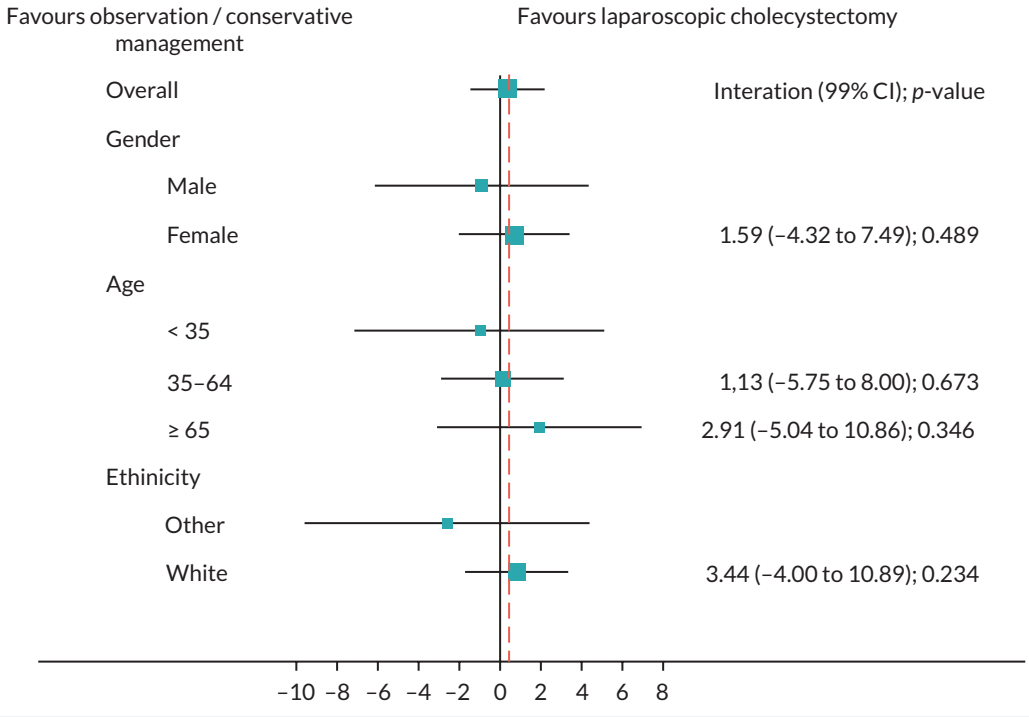

FIGURE 19 Subgroups for observation/conservative management vs. laparoscopic cholecystectomy up to 18 months for SF-36 bodily pain complete case analysis.

TABLE 43 Secondary outcome – complications up to 18 months

Clinical secondary outcome	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Number of participants	32 (14.7)	44 (20.3)
RR (95% CI); *p*-value	0.72	95% CI (0.46 to 1.14); *p*-value 0.17
Number of complications		
1	18	31
2	8	5
3	4	8
4	2	–
Presurgery complications		
Number of participants	25 (11.5)	11 (5.1)
Number of complications		
1	20	9
2	4	–
3	1	1
4	–	1
Details of presurgery complications		
Cholecystitis	14	8
Biliary colic	8	2
Pancreatitis	2	3
Choledocholithiasis	2	–
Cholecystitis and jaundice	1	–
Choledocholithiasis and pancreatitis	1	–
Cholecystitis, choledocholithiasis and jaundice	–	1
Cholecystitis and pancreatitis	1	–
Bouveret syndrome[a]	1	–
Cholecystitis, choledocholithiasis and pancreatitis	–	1
Jaundice	–	1
RUQ pain	1	–
Intraoperative complications		
Number of participants	9 (4.1)	24 (11.1)
Number of complications		
1	8	23
2	1	1

continued

TABLE 43 Secondary outcome – complications up to 18 months (*continued*)

Clinical secondary outcome	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Details of intraoperative complications		
Bile/stone spillage from gallbladder	6	16
Injury to abdominal viscera (including liver tear or laceration)	1	5
Bleeding > 500 ml	1	2
Bile leak from the bile duct, hepatic duct or ducts at base of liver	1	1
Injury to bile duct	1	–
Ruptured empyema	–	1
Postoperative complications		
Number of participants	7 (3.2)	14 (6.5)
Number of complications		
1	5	9
2	1	4
3	1	1
Details of postoperative complications		
Bleeding > 500 ml	1	2
Bile leak that required no treatment	2	3
Bowel obstruction requiring no treatment	1	3
Bowel obstruction requiring surgery	–	1
Wound infection	2	–
Intraperitoneal – collection/abscess requiring no treatment	1	3
Intraperitoneal – collection/abscess requiring percutaneous drainage	1	–
Vomiting	–	3
Dizziness and hypotension	1	–
Haematoma	–	1
Missed stone in the bile duct	–	1
Renal failure	–	1
Residual gallbladder inflamed	1	–
Wound dehiscence	–	1
Postsurgery complications within 30 days of discharge		
Number of participants	2 (0.9)	3 (1.4)
Details of postsurgery complications within 30 days of discharge		
Cholangitis	–	1
Surgical site infection	1	1

TABLE 43 Secondary outcome – complications up to 18 months (*continued*)

Clinical secondary outcome	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Bile leak	–	1
Postcholecystectomy syndrome[b]	1	–
Postsurgery complications after 30 days of discharge		
Number of participants	1 (0.5)	1 (0.5)
Details of postsurgery complications after 30 days of discharge		
RUQ pain	–	1
Incisional hernia	1	–
Death – cardiovascular event		
Number of participants	–	1 (0.5)

a Bouveret syndrome occurs when a gallstone enters the small bowel via a bilioenteric fistula and is impacted in the duodenum or stomach, causing gastric outlet obstruction.
b Persistence of same symptoms reported by the patient.

Note
Values are *n* (%) or *n*.

TABLE 44 Secondary outcome – complications up to 18 months by treatment received

Clinical secondary outcome	Observation/conservative management, *N* = 234	Laparoscopic cholecystectomy, *N* = 200
Number of participants	8 (3.4)	68 (34.0)
Number of complications		
1	6	43
2	1	12
3	–	11
4	1	2
Presurgery complications		
Number of participants	8 (100.0)	28 (29.2)
Number of complications		
1	6	23
2	1	3
3	–	2
4	1	–
Details of presurgery complications		
Cholecystitis	5	17
Biliary colic	4	6
Pancreatitis	–	5
Choledocholithiasis	1	1

continued

TABLE 44 Secondary outcome – complications up to 18 months by treatment received (*continued*)

Clinical secondary outcome	Observation/conservative management, *N* = 234	Laparoscopic cholecystectomy, *N* = 200
Cholecystitis and jaundice	–	1
Choledocholithiasis and pancreatitis	1	–
Cholecystitis, choledocholithiasis and jaundice	–	1
Cholecystitis and pancreatitis	–	1
Bouveret syndrome[a]	–	1
Cholecystitis, choledocholithiasis and pancreatitis	–	1
Jaundice	1	–
RUQ pain	–	1
Intraoperative complications		
Number of participants	0 (0.0)	33 (34.4)
Number of complications		
1	–	31
2	–	2
Details of intraoperative complications		
Bile/stone spillage from gallbladder	–	22
Injury to abdominal viscera (including liver tear or laceration)	–	6
Bleeding > 500 ml	–	3
Bile leak from the bile duct, hepatic duct or ducts at base of liver	–	2
Injury to bile duct	–	1
Ruptured empyema	–	1
Postoperative complications		
Number of participants	0 (0.0)	21 (10.5)
Number of complications		
1	–	14
2	–	5
3	–	2
Details of postoperative complications		
Bleeding > 500 ml	–	3
Bile leak that required no treatment	–	5
Bowel obstruction requiring no treatment	–	4

TABLE 44 Secondary outcome – complications up to 18 months by treatment received (*continued*)

Clinical secondary outcome	Observation/conservative management, *N* = 234	Laparoscopic cholecystectomy, *N* = 200
Bowel obstruction requiring surgery	–	1
Wound infection	–	2
Intraperitoneal – collection/abscess requiring no treatment	–	4
Intraperitoneal – collection/abscess requiring percutaneous drainage	–	1
Vomiting	–	3
Dizziness and hypotension	–	1
Haematoma	–	1
Missed stone in the bile duct	–	1
Renal failure	–	1
Residual gallbladder inflamed	–	1
Wound dehiscence	–	1
Postsurgery complications within 30 days of discharge		
Number of participants (and complications)	0 (0)	5 (5.2)
Details of postsurgery complications within 30 days of discharge		
Cholangitis	–	1
Surgical site infection	–	2
Bile leak	–	1
Postcholecystectomy syndrome[b]	–	1
Postsurgery complications after 30 days of discharge		
Number of participants	0 (0)	2 (2.1)
Details of postsurgery complications after 30 days of discharge		
RUQ pain	–	1
Incisional hernia	–	1
Death – cardiovascular event		
Number of participants	0 (0)	1 (1.0)

a Bouveret syndrome occurs when a gallstone enters the small bowel via a bilioenteric fistula and is impacted in the duodenum or stomach, causing gastric outlet obstruction.
b Persistence of same symptoms reported by the patient.

Note
Values are *n* (%) or *n*.

TABLE 45 Secondary outcome – complications up to 24 months by treatment received

Clinical secondary outcome	Observation/ conservative management, *N* = 234	Laparoscopic cholecystectomy, *N* = 200
Number of participants	7 (3.2)	71 (32.7)
Number of complications		
1	5	45
2	1	12
3	–	12
4	1	2
Presurgery complications		
Number of participants	7 (100.0)	29 (29.3)
Number of complications		
1	5	24
2	1	3
3	–	2
4	1	–
Details of presurgery complications		
Cholecystitis	5	17
Biliary colic	4	6
Pancreatitis	–	5
Choledocholithiasis	1	1
Cholecystitis and jaundice		
Choledocholithiasis and pancreatitis	–	1
Cholecystitis, choledocholithiasis and jaundice	–	1
Cholecystitis and pancreatitis	–	1
Bouveret syndrome[a]	–	1
Cholecystitis, choledocholithiasis and pancreatitis	–	1
Jaundice	1	–
RUQ pain	–	1
Intraoperative complications		
Number of participants	0 (0.0)	33 (33.3)
Number of complications		
1	–	31
2	–	2
Details of intraoperative complications		
Bile/stone spillage from gallbladder	–	22
Injury to abdominal viscera (including liver tear or laceration)	–	6
Bleeding > 500 ml	–	3
Bile leak from the bile duct, hepatic duct or ducts at base of liver	–	2

TABLE 45 Secondary outcome – complications up to 24 months by treatment received (*continued*)

Clinical secondary outcome	Observation/ conservative management, *N* = 234	Laparoscopic cholecystectomy, *N* = 200
Injury to bile duct	–	1
Ruptured empyema	–	1
Postoperative complications		
Number of participants	0 (0.0)	23 (10.6)
Number of complications		
1	–	15
2	–	5
3	–	3
Details of postoperative complications		
Bleeding > 500 ml	–	3
Bile leak that required no treatment	–	5
Bowel obstruction requiring no treatment	–	5
Bowel obstruction requiring surgery	–	1
Wound infection	–	4
Intraperitoneal – collection/abscess requiring no treatment	–	5
Intraperitoneal – collection/abscess requiring percutaneous drainage	–	1
Vomiting	–	3
Dizziness and hypotension	–	1
Haematoma	–	1
Missed stone in the bile duct	–	1
Renal failure	–	1
Residual gallbladder inflamed	–	1
Wound dehiscence	–	1
Postsurgery complications within 30 days of discharge		
Number of participants	0 (0.0)	5 (5.1)
Number of complications		
1	–	5
Details of postsurgery complications within 30 days of discharge		
Cholangitis	–	1
Surgical site infection	–	2
Bile leak	–	1
Postcholecystectomy syndrome[b]	–	1
Postsurgery complications after 30 days of discharge		
Number of participants	0 (0.0)	2 (2.0)

continued

TABLE 45 Secondary outcome – complications up to 24 months by treatment received (*continued*)

Clinical secondary outcome	Observation/ conservative management, *N* = 234	Laparoscopic cholecystectomy, *N* = 200
Details of postsurgery complications after 30 days of discharge		
RUQ pain	–	1
Incisional hernia	–	1
Death – cardiovascular		
Number of participants	–	1 (1.0)

a Bouveret syndrome occurs when a gallstone enters the small bowel via a bilioenteric fistula and is impacted in the duodenum or stomach, causing gastric outlet obstruction.
b Persistence of same symptoms reported by the patient.

Note
Values are *n* (%) or *n*.

TABLE 46 Secondary outcome – further treatment up to 18 months

Clinical secondary outcome	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Number of participants	9/202 (4.5)	12/203 (5.9)
RR (95% CI); *p*-value	0.75	95% CI (0.31, 1.78); *p*-value 0.509
Number of treatments		
1	7	8
2	2	2
3	–	1
7	–	1
Details		
Pain relief	3	8
Antibiotics	2	3
ERCP	3	4
Antisickness	1	–
Gas and air	1	–
Catheter fitted for urinary retention	–	2
Bowel	–	1
Blood transfusion	–	1
Laparotomy washout and haemostasis	–	1
Fluids	–	1
Pancreatitis treatment	–	1
Unknown	1	–

Note
Values are *n* (%) or *n*.

TABLE 47 Professional appointments, medication prescribed and further investigation up to 18 months

Details	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Number of participants who required appointments	103/202 (51.0)	118/203 (58.1)
Details		
GP	69	89
District nurse	3	8
Practice nurse	13	32
NHS hospital outpatients	50	59
NHS hospital A&E	38	28
Private health care	2	3
Appointment with other care provider	8	–
Consultant	5	–
Physiotherapist	2	–
Paramedics	1	–
Medications prescribed		
Number of participants who were prescribed and taken medication	63/200 (31.5)	56/201 (27.9)
Details		
Pain relief	95	67
Paracetamol	35	23
Ibuprofen	20	18
Dihydrocodeine	30	23
Tramadol	7	2
Other pain medication	3	1
Antispasmodic	23	17
Antisickness	–	1
Antibiotic	–	3
Antireflux	17	20
Antidiarrhoea	–	1
Other	3	7
Bile salts	1	5
Iron	2	2
Further investigations		
Number of participants	10/200 (5.0)	10/201 (5.0)
Number of investigations		
1	8	9
2	2	1

continued

TABLE 47 Professional appointments, medication prescribed and further investigation up to 18 months (*continued*)

Details	Observation/conservative management, *N* = 217	Laparoscopic cholecystectomy, *N* = 217
Details		
MRI	3	2
Ultrasound scan	3	4
C-scet	1	2
CT	1	1
Endoscopy	2	1
Sonographer	–	1
Unknown	2	–

Note
Values are *n* (%) or *n*.

TABLE 48 Proportion of participants in the COVID period or not (COVID defined as from 11 March 2020) for people who responded to the questionnaires

	Time point (months)				
In COVID period	3	9	12	18	24
Overall					
No	350 (100.0)	284 (93.4)	248 (81.3)	199 (60.7)	97 (35.5)
Yes	–	20 (6.6)	57 (18.7)	129 (39.3)	176 (64.5)
Observation/conservative management group					
No	176 (100.0)	136 (94.4)	128 (82.1)	97 (58.1)	43 (31.9)
Yes	–	8 (5.6)	28 (17.9)	70 (41.9)	90 (66.7)
Laparoscopic cholecystectomy group					
No	174 (100.0)	148 (92.5)	120 (80.5)	102 (63.4)	52 (37.7)
Yes	–	12 (7.5)	29 (19.5)	59 (36.6)	86 (62.3)

Note
Values are *n* (%).

Appendix 4 Health economics

Within-trial economic evaluation: further details on the statistical analysis

Test results for the generalised linear model (GLM) family and link selection for the within RCT cost–utility analysis are reported in *Tables 49* and *50*. A modified Park test was conducted, showing a Gaussian family as a better model specification for Cost and QALY analyses. Identity link was selected for both analyses following results for the Pearson Correlation, Pregibon Link and a Modified Hosmer and Lemeshow tests.

Within-trial economic evaluation: multiple imputation regression results for the subgroup analyses

Table 51 shows the coefficient values for the subgroup indicator/dummy and treatment interaction terms from the subgroup regression analyses. The 95% CIs cross zero for all subgroups with *p*-values well above the 5% significance threshold. Therefore, we found no evidence that the composition of these subgroups has any differential effect on costs or QALYs.

Model-based economic evaluation: further details on the survival analysis on time to cholecystectomy

Akaike information criterion (AIC) and Bayesian information criterion (BIC) show very little difference between the alternative time to event distributions (see *Table 52*). Therefore, survival function selection was based on the visual inspection of the survival curves (see *Figure 20*) and the modelled proportion of individuals going through cholecystectomy by 10 years.

TABLE 49 Model selection test results: total costs

Family	Chi²	*p*-value
Gaussian NLLS	0.8768	0.3491
Poisson	27.3204	0.0000
Gamma	129.7354	0.0000
Inverse Gaussian	308.1218	0.0000
Results of tests for link; *p*-values		
GLM, Gaussian family		
	Identify link	**Log-link**
Pearson correlation test	1.0000	0.9296
Pregibon link test	0.0000	0.2073
Modified Hosmer and Lemeshow	0.0000	0.0000

Note
Fitted model: Link = Identity; Family = Gaussian.
Family: chi-squared and *p*-value in descending order of likelihood.

TABLE 50 Model selection test results: total QALYs

Family	Chi²	*p*-value
Gaussian NLLS	0.2625	0.6084
Poisson	1.2216	0.269
Gamma	2.8838	0.0895
Inverse Gaussian or Wald	5.249	0.022
Results of tests for link; *p*-values		
GLM, Gaussian family		
	Identify link	
Pearson correlation test	1.0000	
Pregibon link test	1.0000	
Modified Hosmer and Lemeshow	0.6841	

Note
Fitted model: Link = Identity; Family = Gaussian.
Family: chi-squared and *p*-value in descending order of likelihood.

TABLE 51 Multiple imputation regression analysis

Subgroup		Coefficient	Standard error	*p*-value	95% CI	
Gender (males = 1)	Costs	−607.92	493.17	0.22	−1574.53	358.68
	QALYs	0.006	0.044	0.89	−0.080	0.092
Age (65 and over = 1)	Costs	199.26	418.69	0.63	−621.40	1019.92
	QALYs	0.030	0.057	0.60	−0.082	0.142
Ethnicity (white = 1)	Costs	56.99	388.44	0.88	−704.37	818.34
	QALYs	0.058	0.052	0.27	−0.045	0.161

Note
Interaction between treatment dummy variable (observation/conservative management = 1) and subgroup indicator.

TABLE 52 Survival analysis for time to cholecystectomy

	From month 9 onwards		From month 12 onwards	
Number of subjects; number of failures; time at risk (days)	**271; 81; 448.9**		**243; 54; 385.5**	
Distribution	**AIC**	**BIC**	**AIC**	**BIC**
Exponential	587.8	595.0	395.4	402.4
Weibull	568.6	579.4	393.9	404.4
Gompertz	571.6	582.4	393.5	404.0
Loglogistic	565.6	576.4	392.9	403.4
Generalised	563.0	577.4	394.1	408.1

Note
AIC, AKAIKE Information Criterion; BICS, Bayesian Information Criterion.

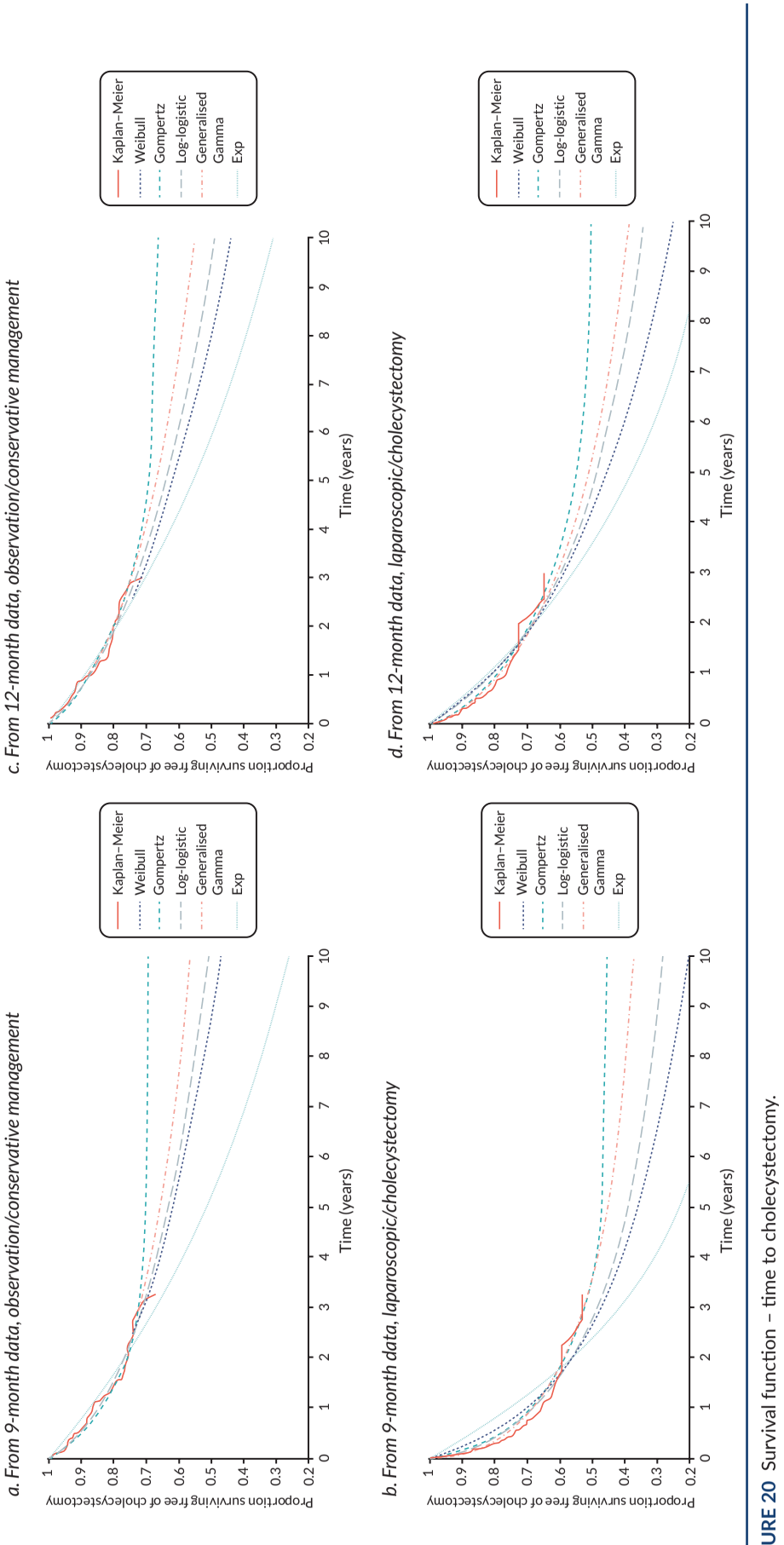

FIGURE 20 Survival function – time to cholecystectomy.

Model-based economic evaluation: model validation; within RCT and 24-month decision model run results

Table 53 reports the within-trial analysis and model-based analysis results. For this comparison, the Markov model was run for a 24-month time horizon using the unit cost for the day-case cholecystectomy as this unit cost was similar to the cost of the cholecystectomy episode resulting from the within-trial analysis. The difference in costs and QALYs for the two analyses produce consistent results. That is, ICERs are well above the usual threshold applied in the UK for decision-making (i.e. £20,000). However, these are substantially different: £55,235 and £88,930 for the within-trial and model-based analyses, respectively. The cost difference between the two analyses reflects the fewer cost categories incorporated in the modelling analysis. Moreover, the difference in the incremental QALYs between the two analyses is the result of the alternative approach used to obtain the AUC (linear interpolation for the within-trial analysis vs. simple extrapolation for the model) and the fact that neither cost nor QALYs were adjusted for minimisation factors in the model-based analysis.

Model-based economic evaluation: SF-6D repeat measure analysis results

The C-Gall trial follow-up was 24 months. However, the trial office continue issuing PQs beyond 24 months, every 6 months. SF-36 data were collected and SF-6D scores obtained and analysed. A mixed-effects regression model for repeat measures with adjustment for the minimisation covariates (gender, age) and including centre as a random effect was used to obtain the SF-6D utility score difference between trial groups by data collection time point. The results of this analysis are reported in *Table 54*. Consistent with the trial main statistical analysis of no significant difference in the quality of life (measured by AUC over 18 months using the SF-36 bodily pain domain), non-significant utility score differences between randomised groups were observed for up to 24 months. Nevertheless, statistically significant differences are seen at 30 and 36 months post randomisation with a non-significant overall treatment effect through the period of 0.0176 favouring laparoscopic cholecystectomy.

TABLE 53 C-GALL trial within-trial analysis and model-based results (over 24 months follow-up/time horizon)

Intervention	Total cost (£)	Δ cost (£)	Total QALYs	Δ QALY	ICER
Within-trial analysis					
Laparoscopic cholecystectomy	2510		1.413		
Observation/conservative management	1477	−1033	1.395	−0.019	55,235
Model-based analysis for 24-month time horizon					
Laparoscopic cholecystectomy	1966		1.366		
Observation/conservative management	871	−1095	1.354	−0.012	88,930

Δ increment.

TABLE 54 SF-6D utility score. Treatment effect by time point

		Treatment effect				
Time point (months)	***N***	**Coefficient**	**Standard error**	***z***	***P* > *z***	**(95% CI)**
3	294	0.0105	0.0142	0.74	0.46	(−0.0173 to 0.0383)
9	255	−0.0201	0.015	−1.35	0.178	(−0.0495 to 0.0092)
12	236	−0.0186	0.0154	−1.21	0.227	(−0.0487 to 0.0115)
18	255	−0.0286	0.015	−1.91	0.056	(−0.0579 to 0.0007)
24	225	−0.0088	0.0157	−0.56	0.573	(−0.0396 to 0.0219)
30	223	−0.0342	0.0157	−2.18	0.029	(−0.065 to −0.0034)
36	187	−0.0345	0.0167	−2.06	0.039	(−0.0673 to −0.0017)
42	136	−0.0299	0.019	−1.58	0.115	(−0.0671 to 0.0073)
48	77	0.0085	0.0243	0.35	0.725	(−0.0391 to 0.0561)
Overall treatment effect[a]		−0.0176	0.0101	−1.74	0.08	(−0.0374 to 0.0022)

a Dummy variable; Observation/conservative management.

Appendix 5 Updated search strategy to identify clinical effectiveness studies

Ovid MEDLINE(R) Epub Ahead of Print, In-Process and Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 – 3 May 2016 EMBASE 1980 to 2016 Week 18

Date of search: 3 May 2016

1. cholecystitis/
2. cholecystitis, acute/
3. cholecystolithiasis/
4. gallstones/
5. cholelithiasis/
6. biliary colic/
7. z(gall?bladder adj3 (empyema or inflam$)).tw.
8. (biliary colic or gall?stone$ or cholecystitis or cholecystolithiasis).tw.
9. ((pain or biliary symptom$) adj5 (cholecystitis or cholecystolithiasis or gall?bladder)).tw.
10. or/1-9
11. exp Cholecystectomy/
12. cholecystectom$.tw.
13. ((excis$ or remov$) adj4 gall?bladder).tw.
14. ((surgery or surgical) adj5 (cholecystitis or cholecystolithiasis or gall?bladder)).tw.
15. or/11-14
16. exp clinical trial/
17. randomized controlled trial.pt.
18. controlled clinical trial.pt
19. randomi?ed.ab.
20. randomly.ab.
21. trial.ab.
22. placebo.ab.
23. drug therapy.fs.
24. groups.ab.
25. comparative study/
26. (prospective$ or retrospective$).tw.
27. (compare$ or compara$).ti,ab.
28. or/16-27
29. 10 and 15 and 28
30. (review or editorial or case report$ or letter).pt.
31. 29 not 30
32. limit 31 to human

SEARCH STRATEGY TO IDENTIFY PROMS **IN CHOLECYSTITIS**

Ovid MEDLINE(R) Epub Ahead of Print, In-Process and Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 – 3 May 2016 EMBASE 1980 to 2016 Week 18

Date of search: 3 May 2016

1. exp cholecystitis/
2. cholecystolithiasis/
3. gallstones/

4. biliary colic/
5. (gall?bladder adj3 (empyema or inflam$)).tw,kw.
6. (biliary colic or gall?stone$ or cholecystitis or cholecystolithiasis).tw,kw.
7. ((pain or biliary symptom$) adj5 (cholecystitis or cholecystolithiasis or gall?bladder)).tw,kw.
8. or/1-7
9. (core adj3 outcome?).tw,kw.
10. (patient reported adj3 outcome?).tw,kw.
11. prom.tw,kw.
12. 8 and (9 or 10 or 11)
13. *outcome assessment/
14. *'Outcome Assessment (Health Care)'/
15. 8 and (13 or 14)
16. 12 or 15

SEARCH STRATEGY TO IDENTIFY QUALITATIVE STUDIES

Epub Ahead of Print, In-Process and Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) < 1946 to 2 May 2016

Date of search: 18 August 2016

1. exp cholecystitis/
2. cholecystolithiasis/
3. gallstones/
4. exp Cholecystectomy/
5. biliary colic/
6. (gall?bladder adj3 (empyema or inflam$)).tw,kw.
7. (biliary colic or gall?stone$ or cholecystitis or cholecystolithiasis).tw,kw.
8. ((pain or biliary symptom$) adj5 (cholecystitis or cholecystolithiasis or gall?bladder)).tw,kw.
9. or/1-8
10. qualitative research/
11. exp interviews as topic/
12. focus groups/
13. grounded theory/
14. (qualitative or interview$ or focus group?).tw,kw.
15. (ethno$ or grounded or thematic or interpretive or narrative or discourse analysis or discursive or mixed method$).tw,kw.
16. or/10-15
17. 9 and 16
18. exp animals/not human/
19. 17 not 18

Appendix 6 List of outcomes for consensus stage

How important is it that an intervention to treat uncomplicated symptomatic gallstone disease can be shown to affect:

#	Domain	Outcome	Definition
1	*Physical*	Physical activity	Activities such as walking, running, swimming, cycling, physical labour, climbing stairs, gardening, etc.
2		Exercise	Being able to do activities requiring physical effort, carried out to sustain or improve health and fitness (strength and endurance)
3	*Role*	No. of days sick leave	Length of time off work after the operation in days
4		Time to everyday life	Length of time taken to return to usual everyday activities
5		Impact on others	Impact of your gallstone condition or your gallstone surgery on relationships with people surrounding you
6	*Pain*	Overall pain	Overall pain
7		Abdominal pain	General pain occurring at rest and/or when coughing, originating in the abdominal area
8		Umbilical pain	Pain around the belly button scar (this is where the main port is that removed the stones)
9		Shoulder pain	Pain relating to or affecting the right shoulder region
10	*Bowel movements*	Diarrhoea	Watery stools, loose bowel motion
11		Constipation	Difficulty passing stool
12	*Thirst/ dehydration*	Resumption of orals	Starting to eat and drink after treatment
13	*Appetite/eating/ taste*	Time to resume eating	Length of time taken to return to oral food intake
14	*Fatigue*	Fatigue	Feeling physically or mentally tired or lacking in energy
15	*Sleep*	Length of night sleep	Length of night's sleep
16	*Cognitive*	Difficulty concentrating	Inability to focus attention on one task or problem
17	*Emotional*	Anxiety	A feeling of worry, nervousness or unease
18		Distress	A feeling of extreme anxiety, stress or anguish
19		Trust	Belief in the reliability, truth or ability of someone or something
20	*Generic health*	QoL	How well you feel physically and emotionally because of a combination of: • Your gallstones • The prospect of treatment • The result of treatment (treatment might include surgery or painkillers)
21		Overall health state	Overall state of your physical and mental condition
22		Overall satisfaction	The degree to which expectations or needs have been fulfilled

#	Domain	Outcome	Definition
23	*Dietary habits*	Food intolerance	An adverse physical reaction by the body to certain foods
24	*Social*	Time away from recreational activities	Time spent away from enjoyable activities as a result of your gallstone condition or gallstone surgery
25	*Belching/ bloating/gas*	Flatulence	Belching, farting, bloating or gas
26		Bloating	Abdominal swelling as a result of excess fluid or gas
27		Abdominal discomfort	Pain or discomfort in the stomach area
28	*Service use*	Hospital stay	Length of time spent in the hospital from admission to discharge
29	*Vomiting/ nausea*	Vomiting	Being sick
30		Nausea	Feeling sick
31	*Reflux*	Heartburn	A form of indigestion that presents as a burning sensation in the chest, caused by acid reflux
32	*Body image*	Satisfaction with body image	A feeling of satisfaction with your physical appearance
33		Satisfaction with the cosmetic outcome	The extent to which you are content with the cosmetic results of gallstone surgery
34	*Sexual function*	Satisfaction in the context of sexual intercourse	The extent to which you are satisfied with experiences of sexual intercourse in relation to your gallstone condition or your gallstone surgery *Note: This outcome is particularly relevant to women having Natural Orifice Transluminal Endoscopic Surgery (NOTES)*
35		Pain in the context of sexual intercourse	The extent to which you are experiencing pain during or after sexual intercourse in relation to your gallstone condition or your gallstone surgery *Note: This outcome is particularly relevant to women having Natural Orifice Transluminal Endoscopic Surgery (NOTES)*
36	*Regurgitation*	Regurgitation	Bringing swallowed food back up to the mouth
37	*Dysphagia/ swallowing*	Trouble swallowing food	Problems swallowing food
38	*Generic symptoms*	General discomfort	An unpleasant feeling and/or low-level pain which is hard to define
39		Residual symptoms	Continuing to have symptoms (such as pain, bloating, etc.) after removal of the gallbladder
40		Dizziness	Feeling light-headed or dizzy
41		Fainting	Fainting (short-term loss of consciousness)
42	*Mortality*	Mortality	Death from any cause
43	*Intra-op AE*	Common bile duct stones	Stones in the common bile duct
44		Common bile duct injury	During surgery the common bile duct is damaged
45		Biliary leak	The liver produces bile which is stored in the gallbladder (see diagram). If this is damaged, the bile can leak and cause complications
46		Haemorrhage	Bleeding or the abnormal flow of blood; the release of blood from a ruptured blood vessel
47	*Intra- and post-op AE*	Intra-abdominal collections	After surgery, any type of fluid collecting in the abdomen

#	Domain	Outcome	Definition
48	*Post-op AE*	Hernia occurrence	Internal hernia – displacement of an organ within the abdomen through a potential defect
49		Port-site complications	Complications such as infection, hernia, pain or bleeding at or within the 'keyholes' characteristic of keyhole surgery
50		Wound infections	An infection at the wound site
51		Patient-perceived success of the operation	How patients perceive the success of the operation
52	*Cost-effectiveness*	Hospital cost	Total hospital costs, taking into account the total length of hospital stay, operating room charges, medical and surgical supplies, pharmacy, laboratory and pathology, recovery room, anaesthesia and ICU/observation rooms
53		Overall cost	Cost of use of healthcare services; for example contact with a GP, in- or outpatient contact, prescribed medications
54		Cost-effectiveness ratio	Cost-effectiveness of treatment route (medical management or surgery to remove the gallbladder), calculated by dividing cost by success rate (defined by the quality of life after treatment)

EME
HSDR
HTA
PGfAR
PHR

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

Published by the NIHR Journals Library