NIHR | National Institute for Health and Care Research

Journals Library

Check for updates

Health Technology Assessment

Volume 28 • Issue 33 • July 2024

ISSN 2046-4924

Treatment options for patients with pilonidal sinus disease: PITSTOP, a mixed-methods evaluation

Steven Brown, Daniel Hind, Emily Strong, Mike Bradburn, Farhat Vanessa Nasim Din, Ellen Lee, Matthew J Lee, Jonathan Lund, Christine Moffatt, Jonathan Morton, Asha Senapati, Philip Shackley, Peter Vaughan-Shaw, Arkadiusz Peter Wysocki, Tia Callaghan, Helen Jones and Nyantara Wickramasekera on behalf of the PITSTOP Management Group

DOI 10.3310/KFDQ2017

Treatment options for patients with pilonidal sinus disease: PITSTOP, a mixed-methods evaluation

Steven Brown,[1*] Daniel Hind,[2] Emily Strong,[2] Mike Bradburn,[2] Farhat Vanessa Nasim Din,[3] Ellen Lee,[2] Matthew J Lee,[4] Jonathan Lund,[5] Christine Moffatt,[6] Jonathan Morton,[7] Asha Senapati,[8] Philip Shackley,[9] Peter Vaughan-Shaw,[10] Arkadiusz Peter Wysocki,[11] Tia Callaghan,[2] Helen Jones[12] and Nyantara Wickramasekera[9] on behalf of the PITSTOP Management Group

[1]Department of General Surgery, Northern General Hospital, Sheffield, UK
[2]Sheffield Clinical Trials Research Unit, School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, UK
[3]Academic Coloproctology, Institute of Genetics and Cancer, University of Edinburgh, Western General Hospital, Edinburgh, UK
[4]Department of Oncology and Metabolism, The Medical School, University of Sheffield, Sheffield, UK
[5]Derby Royal Infirmary, University Hospitals of Derby and Burton, Derby, UK
[6]Nottingham University Hospitals NHS Trust, Nottingham, UK
[7]Addenbrookes Hospital, Cambridge University Hospitals, Cambridge, UK
[8]St Mark's Hospital, London, UK; Queen Alexandra Hospital, Portsmouth, UK
[9]School of Health and Related Research, Regent Court, Sheffield, UK
[10]Department of Colorectal Surgery, Western General Hospital, Edinburgh, UK
[11]Griffiths University Medical School, Brisbane, Australia
[12]Oxford University Hospitals NHS Foundation Trust, Oxford, UK

[*]Corresponding author

Disclaimer: This report contains transcripts of interviews conducted in the course of the research and contains language that may offend some readers.

Published July 2024
DOI: 10.3310/KFDQ2017

This report should be referenced as follows:

Brown S, Hind D, Strong E, Bradburn M, Din FVN, Lee E, *et al*. Treatment options for patients with pilonidal sinus disease: PITSTOP, a mixed-methods evaluation. *Health Technol Assess* 2024;**28**(33). https://doi.org/10.3310/KFDQ2017

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 17/17/02. The contractual start date was in September 2018. The draft manuscript began editorial review in April 2023 and was accepted for publication in November 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

Treatment options for patients with pilonidal sinus disease: PITSTOP, a mixed-methods evaluation

Steven Brown[1*] Daniel Hind[2] Emily Strong[2] Mike Bradburn[2]
Farhat Vanessa Nasim Din[3] Ellen Lee[2] Matthew J Lee[4]
Jonathan Lund[5] Christine Moffatt[6] Jonathan Morton[7]
Asha Senapati[8] Philip Shackley[9] Peter Vaughan-Shaw[10]
Arkadiusz Peter Wysocki[11] Tia Callaghan[2]
Helen Jones[12] and Nyantara Wickramasekera[9]
on behalf of the PITSTOP Management Group

[1]Department of General Surgery, Northern General Hospital, Sheffield, UK
[2]Sheffield Clinical Trials Research Unit, School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, UK
[3]Academic Coloproctology, Institute of Genetics and Cancer, University of Edinburgh, Western General Hospital, Edinburgh, UK
[4]Department of Oncology and Metabolism, The Medical School, University of Sheffield, Sheffield, UK
[5]Derby Royal Infirmary, University Hospitals of Derby and Burton, Derby, UK
[6]Nottingham University Hospitals NHS Trust, Nottingham, UK
[7]Addenbrookes Hospital, Cambridge University Hospitals, Cambridge, UK
[8]St Mark's Hospital, London, UK; Queen Alexandra Hospital, Portsmouth, UK
[9]School of Health and Related Research, Regent Court, Sheffield, UK
[10]Department of Colorectal Surgery, Western General Hospital, Edinburgh, UK
[11]Griffiths University Medical School, Brisbane, Australia
[12]Oxford University Hospitals NHS Foundation Trust, Oxford, UK

[*]Corresponding author steven.brown13@nhs.net

Background: There is no consensus on optimal management of pilonidal disease. Surgical practice is varied, and existing literature is mainly single-centre cohort studies of varied disease severity, interventions and outcome assessments.

Objectives: A prospective cohort study to determine:

- disease severity and intervention relationship
- most valued outcomes and treatment preference by patients
- recommendations for policy and future research.

Design: Observational cohort study with nested mixed-methods case study. Discrete choice experiment. Clinician survey. Three-stage Delphi survey for patients and clinicians. Inter-rater reliability of classification system.

Setting: Thirty-one National Health Service trusts.

Participants: Patients aged > 16 years referred for elective surgical treatment of pilonidal disease.

Interventions: Surgery.

Main outcome measures: Pain postoperative days 1 and 7, time to healing and return to normal activities, complications, recurrence. Outcomes compared between major and minor procedures using regression modelling, propensity score-based approaches and augmented inverse probability weighting to account for measured potential confounding features.

Results: *Clinician survey*: There was significant heterogeneity in surgeon practice preference. Limited training opportunities may impede efforts to improve practice.

Cohort study: Over half of patients (60%; *N* = 667) had a major procedure. For these procedures, pain was greater on day 1 and day 7 (mean difference day 1 pain 1.58 points, 95% confidence interval 1.14 to 2.01 points, *n* = 536; mean difference day 7 pain 1.53 points, 95% confidence interval 1.12 to 1.95 points, *n* = 512). There were higher complication rates (adjusted risk difference 17.5%, 95% confidence interval 9.1 to 25.9%, *n* = 579), lower recurrence (adjusted risk difference −10.1%, 95% confidence interval −18.1 to −2.1%, *n* = 575), and longer time to healing (>34 days estimated difference) and time to return to normal activities (difference 25.9 days, 95% confidence interval 18.4 to 33.4 days).

Mixed-methods analysis: Patient decision-making was influenced by prior experience of disease and anticipated recovery time. The burden involved in wound care and the gap between expected and actual time for recovery were the principal reasons given for decision regret.

Discrete choice experiment: The strongest predictors of patient treatment choice were risk of infection/persistence (attribute importance 70%), and shorter recovery time (attribute importance 30%). Patients were willing to trade off these attributes. Those aged over 30 years had a higher risk tolerance (22.35–34.67%) for treatment failure if they could experience rapid recovery. There was no strong evidence that younger patients were willing to accept higher risk of treatment failure in exchange for a faster recovery. Patients were uniform in rejecting excision-and-leave-open because of the protracted nursing care it entailed.

Wysocki classification analysis: There was acceptable inter-rater agreement (κ = 0.52, 95% confidence interval 0.42 to 0.61).

Consensus exercise: Five research and practice priorities were identified. The top research priority was that a comparative trial should broadly group interventions. The top practice priority was that any interventions should be less disruptive than the disease itself.

Limitations: Incomplete recruitment and follow-up data were an issue, particularly given the multiple interventions. Assumptions were made regarding risk adjustment.

Conclusions and future work: Results suggest the burden of pilonidal surgery is greater than reported previously. This can be mitigated with better selection of intervention according to disease type and patient desired goals. Results indicate a framework for future higher-quality trials that stratify disease and utilise broad groupings of common interventions with development of a patient-centred core outcome set.

Trial registration: This trial is registered as ISRCTN95551898.

Funding: This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: 17/17/02) and is published in full in *Health Technology Assessment*; Vol. 28, No. 33. See the NIHR Funding and Awards website for further award information.

Contents

List of tables **xi**

List of figures **xiii**

List of boxes **xv**

List of abbreviations **xvii**

Plain language summary **xix**

Scientific summary **xxi**

Chapter 1 Introduction **1**
Background 1
Pattern of disease and management options 1
Monetary and humanistic burden 1
Current evidence base 2
Rationale 2
Research objectives 3

Chapter 2 Consultant surgeon survey **5**
Methods 5
Survey design and development 5
Data collection 5
Sampling 5
Data analysis 5
Results 5
Respondent practice overview 5
Operative strategies employed 6
Case vignettes 6
Training 6
Discussion 7
Overview 7
Implications for policy-makers 8
Implications for researchers 9
Conclusion 9

Chapter 3 Cohort study **11**
Methods 11
Aims 11
Participants 11
Study procedures 11
Statistical methods 12
Patient and public involvement 15
Results 16
Recruitment and participant flow 16
Baseline characteristics 16
Data completion 18
Treatment decisions 21

Shared decision-making and decision regret 22
Outcome summaries 22
Risk-adjusted treatment comparisons 23
Surgeon variation 28
Adverse events 28
Discussion 28
Conclusions 32

Chapter 4 Mixed-methods substudy **33**
Methods 33
Design 33
Participants 33
Data collection 33
Patient and public involvement 33
Data analysis 33
Results 34
Sample 34
Coping in deliberation framework: health threat 34
Coping in deliberation framework: choice 34
Coping in deliberation framework: key outcomes at the time of decision-making 37
Coping in deliberation framework: consolidation 37
Acceptability framework: key outcomes at 6-month follow-up 37
Discussion 41

Chapter 5 Discrete choice experiment **43**
Methods 43
Design and theoretical/conceptual framework 43
Development of survey 43
Patient and public involvement 44
Sampling 44
Recruitment 44
Data analysis 44
Results 45
Participants 45
Patient preferences 45
Maximum acceptable risk 45
Preference heterogeneity 47
Ranking of treatments 47
Participants' understanding of and engagement with the survey 49
Discussion 49

Chapter 6 Consensus exercise **53**
Methods 53
Stakeholders 53
Patient and public involvement 53
Generation of longlist 53
e-Delphi consensus 53
Virtual consensus meeting 54
Results 54
Longlisting of potential outcomes 54
Delphi round 1 54
Delphi round 2 54
Delphi round 3 55
Consensus meeting 55

Discussion 56
Overview 56
Patient and public involvement 57
Impact for policy-makers 57
Impact for researchers 57
Conclusion 57

Chapter 7 Wysocki classification validation exercise **59**
Background 59
Rationale 59
Aims 59
Methods 60
Participating surgeons 60
Participating patients 60
Assessment schedule 60
Statistical methods 61
Sample size 61
Results 61
Participants 61
Agreement among participating surgeons 62
Agreement between participating surgeons and original classification 63
Treatment choice 63
Discussion 64

Chapter 8 Discussion **65**
Overview 65
Implications for practice 67
Implications for research 68
Equality, diversity and inclusion 68
Patient and public involvement 68

Additional information **69**

References **75**

Appendix 1 PITSTOP Project Management Group **83**

Appendix 2 Cohort study participating sites **85**

Appendix 3 Additional figures and tables **87**

List of tables

TABLE 1 Summary of operations offered 6

TABLE 2 Case vignette responses 7

TABLE 3 Training in different procedures 8

TABLE 4 List of outcomes 13

TABLE 5 Participant characteristics 17

TABLE 6 Participant disease characteristics 19

TABLE 7 Treatment characteristics 21

TABLE 8 Summary of pre-op consultation collaboration and post-op DR 23

TABLE 9 Comparison of complications during follow-up between major and minor procedure 25

TABLE 10 Comparison of recurrence between major and minor procedures 26

TABLE 11 Comparison of time to return to normal activity between major and minor procedures 27

TABLE 12 Comparison of time to healing between major and minor procedures 28

TABLE 13 Adverse events and SAEs 29

TABLE 14 Participant characteristics: baseline, surgery and 6-month outcomes 35

TABLE 15 Decision-making – cases ordered by self-reported quality of SDM 38

TABLE 16 Participants' sociodemographic and clinical characteristics 46

TABLE 17 Estimated MAR 47

TABLE 18 Preference heterogeneity – latent class model results 48

TABLE 19 Estimated MAR based on subgroups 48

TABLE 20 Survey comprehension and internal validity 50

TABLE 21 Summary of Delphi consensus participant characteristics 55

TABLE 22 Final set of statements presented in order of sum 56

TABLE 23 Patients participating in the classification exercise 62

TABLE 24 Summary of agreement between assessors 62

TABLE 25 Outcome completion rates **87**

TABLE 26 Procedure information **88**

TABLE 27 Continuous outcomes measured at multiple time points **90**

TABLE 28 Repacking and re-dressing procedures during follow-up: characteristics **92**

TABLE 29 Postoperative complications during follow-up **93**

TABLE 30 Outcomes measured once during follow-up and outcomes incorporating the full follow-up period **95**

TABLE 31 Comparison of pain on day 1 and day 7 between major and minor procedures **96**

TABLE 32 Recurrence rate for surgeons who operated on at least 10 cohort participants (*N* = 13 surgeons) **96**

TABLE 33 Outcomes and reflections ordered by level of DR (0–100, high to low) **97**

TABLE 34 Discrete choice experiment modelled preferences **99**

TABLE 35 Ratings presented by policy statement within each stakeholder group **100**

TABLE 36 Ratings presented by research statement within each stakeholder group **102**

TABLE 37 Baseline demographics and disease characteristics by 6-month follow-up attendance **104**

TABLE 38 Comparison of pain on day 1 and day 7 between asymmetric closure and minimal excision **105**

TABLE 39 Comparison of complications during follow-up between asymmetric closure and minimal excision **106**

TABLE 40 Comparison of recurrence between asymmetric closure and minimal excision **107**

TABLE 41 Comparison of treatment failure between asymmetric closure and minimal excision **108**

TABLE 42 Comparison of time to return to normal activities between asymmetric closure and minimal excision **108**

TABLE 43 Comparison of time to wound healing between asymmetric closure and minimal excision **108**

List of figures

FIGURE 1 Participant flow for the PITSTOP study **16**

FIGURE 2 Treatment choice by disease characteristic **22**

FIGURE 3 Pain with major and minor procedures, baseline to 6 months after surgery **24**

FIGURE 4 Time to return to normal activities by major or minor procedure type **27**

FIGURE 5 Time to wound healing by major or minor procedure type **27**

FIGURE 6 Regression model 1 results reproduced in a diagram **45**

FIGURE 7 Ranking of treatments **49**

FIGURE 8 Flow of items through the e-Delphi **55**

FIGURE 9 Allocation of patients with PSD to surgical assessors **60**

FIGURE 10 Raw agreement between assessment and original **63**

FIGURE 11 Kappa chance-corrected agreement between assessment and original **63**

FIGURE 12 Procedure preferences of responding surgeons **87**

FIGURE 13 Treatment choice by number of pits (*N* = 640) **89**

FIGURE 14 CollaboRATE score at baseline by 6-month DR score **109**

FIGURE 15 Gwet AC1 chance-corrected agreement between assessment and original **109**

FIGURE 16 Preferred treatment by surgeon and actual treatment within the study **109**

FIGURE 17 Time to return to normal activities between asymmetric closure and minimal excision **110**

FIGURE 18 Time to wound healing between asymmetric closure and minimal excision **110**

FIGURE 19 Funnel plot of minor procedure use by site **111**

List of boxes

BOX 1 Example DCE task **112**

BOX 2 Treatment ranking exercise **113**

List of abbreviations

ADL	activities of daily living
AE	adverse event
AIC	Akaike information criterion
BIC	Bayesian information criterion
BMI	body mass index
CAIC	consistent Akaike information criterion
CI	confidence interval
CODE	coping in deliberation
CTRU	Clinical Trials Research Unit
CWIQ	Cardiff Wound Impact Questionnaire
DCE	discrete choice experiment
DR	decision regret
EPSiT	endoscopic pilonidal sinus treatment
EQ-5D	EuroQol-5 Dimensions
EQ-5D-5L	EuroQol-5 Dimensions, five-level version
FG	fibrin glue
GP	general practitioner
HES	hospital episode statistics
IPW	inverse probability weighting
IQR	interquartile range
LTFU	lost to follow-up
MAR	maximum acceptable risk
PI	principal investigator
PITSTOP	PIlonidal sinus Treatment - STudying the OPtions
PPI	patient and public involvement
PREM	patient-reported experience measure
PRO	patient-reported outcome
PROM	patient-reported outcome measure
PSD	pilonidal sinus disease
QoL	quality of life
RCT	randomised controlled trial
REC	Research Ethics Committee
SAE	serious adverse event
SD	standard deviation
SDM	shared decision-making

Plain language summary

Background

Pilonidal disease is caused by ingrowing hairs between the buttocks. It can cause pain and infection and may need surgery. We do not know which operation gives the best results, or who operations help.

Objectives

PITSTOP aimed to find out which operation is the best and what is important to patients when deciding on surgery, and to suggest ideas for better treatment and future research.

Methods

We looked at what operations were done and their outcomes. We interviewed patients about their experiences. Some completed a survey to help us understand what operations they might prefer based on risks and outcomes. Surgeons completed a survey about their experiences, and we explored whether a new tool could help us tell the difference between 'mild' and 'bad' disease. We used findings from these studies to help patients and surgeons give priorities for future practice and research.

Results

Six hundred and sixty-seven patients joined PITSTOP. People who had a major operation had more pain and took longer to return to normal activities. Some were still affected 6 months after surgery. However, disease recurrence was lower than after a minor procedure. Patients based decisions about treatment on the likelihood of success and the time to recover. The study and the surgeons' survey both showed marked differences in practice. Surgeons tended to offer one or two operations learned during training. A classification tool put cases in similar groups, but this did not influence treatment choices. The consensus exercise identified five research priorities, the top one being to put types of surgery into two groups. Of the five practice priorities, the top one was that surgery should not make the patient worse than the disease.

Conclusions

There is variation in the treatment of pilonidal disease. Wound issues and impact on daily living should be avoided. The highlighted research questions should be addressed to improve care.

Scientific summary

Background

Despite being a common condition, there is no clear consensus as to how pilonidal disease should be managed. Throughout the UK, surgical practice is varied, and existing literature largely consists of single-centre cohort studies using a range of disease classification systems, interventions and outcome assessments. There is a need to improve clinical management of this condition and define future research processes and priorities.

Objectives

PITSTOP aimed to investigate surgical options for the treatment of pilonidal sinus disease (PSD). The objectives were to:

- Follow patients with symptomatic pilonidal sinus referred to each collaborating site, prospectively recording details of their pit/track anatomy, surgical management, medical events and health-related quality of life (QoL) until 6 months after their operation.
- Describe the combination of interventions currently in use and quantify clinical and patient-reported outcomes (PROs) associated with each.
- Identify patient-specific disease features that might predict poor outcome in each treatment group by risk-modelling methods.
- Derive a case mix-adjusted estimate of the risks associated with common treatment options, using causal inference methods to provisionally rank the optimal management strategies among patients for whom more than one treatment is considered appropriate.
- Provide an overview of patient views and experiences.
- Collect the views of patients on which interventions they would rather avoid and which outcomes they most value.
- Reach a surgeon-based consensus on which subtypes of pilonidal disease may benefit from which treatment options.
- Reach a surgeon and patient-based consensus on research priorities.

Design and setting

PITSTOP was an observational cohort with nested mixed-methods and qualitative design which took place across 33 NHS Trust sites in the UK. The study had an additional clinician survey component and validation of a classification system and culminated in a three-stage Delphi exercise to identify research and practice priorities.

Participants

Eligible patients undergoing surgical management for PSD and interested in participating were consented to the study. Participants referred to a collaborating centre for definitive elective surgical treatment of PSD were required to meet the following criteria:

- Consenting patients over 16 years of age with PSD.

Participants were ineligible if any of the following conditions were met:

- Asymptomatic
- Pregnant
- Unable to give consent
- Acute abscess
- Hypersensitivity to the sealants.

Intervention

This was a non-intervention study with the choice of treatment being a shared decision made between the surgeon and patient. Treatment of PSD typically comprises two essential components (with the exception of phenol injection, seton and fistuloscope/diathermy – which aim to induce fibrosis with or without obliteration of the tracks):

1. Excision of the affected skin and fat (the amount of which varies among patients, and surgeons differ with respect to how the resultant wound is managed).
2. Closure of the wound, which is either left open and heals slowly by secondary intention or closed using sutures, also known as primary wound closure.

There are 18 excision–closure combinations that are theoretically possible. Therefore, the operative intervention was recorded by the following excision and closure techniques:

Excision procedure

- Local excision.
- Curettage.
- Phenol injection.
- Pit picking.
- Seton.
- Laser-assisted excision.
- EPSiT (endoscopic pilonidal sinus treatment).

Closure procedure

- Primary midline closure.
- Fibrin glue (FG).
- Marsupialisation.
- Lateral closure (Karydakis; Bascom's cleft).
- Flap (Limberg; rhomboid).

Follow-up

Participants were followed up 1 day and 7 days after surgery (with day 0 the day of elective surgical treatment), at the routine clinic visit appointment, at 6 months, and at the end of study.

Main outcome measures

As PITSTOP is a cohort study examining current practice, there are no primary outcome measures. The following data were collected:

- pain (numeric rating scale) on day 1 and day 7 postoperatively and at each follow-up
- EuroQol five dimensions five levels questionnaire (EQ-5D-5L) QoL at each follow-up

- interactions with primary and secondary care
- length of time to healing
- return to normal activities
- complications
- recurrence
- infection.

Statistical methods

The study aimed to recruit 800 patients, with at least 100 within each of the front-running management strategies. Two treatment comparisons were undertaken: any major excisional procedure versus any minor procedure, and asymmetric closure versus any minor procedure (minimal excision).

Regression modelling, propensity score-based approaches [inverse probability weighting (IPW) and nearest neighbour matching] and augmented IPW were used to account for measured potential confounding. Continuous outcomes (pain at day 1 and day 7) were modelled using linear regression; binary outcomes (recurrence, complication) were modelled using logistic regression, with marginal probabilities used to estimate absolute risk difference. Time to wound healing and time to return to normal activity were modelled using parametric accelerated survival time.

The following features were considered as potential risk factors for outcome and treatment choice: sex, body mass index, depth of natal cleft, presence and type of gluteal hair, smoking status, pit density, presence of unilateral or bilateral disease, distance from furthest lateral opening to the nearest pit, presence of pus and Wysocki disease classification. Features affecting treatment choice were assessed using logistic regression with treatment choice as the outcome, and the same propensity score adjustments were used for each outcome. For each outcome, three regression models were fitted: one that adjusted for all listed features, one that adjusted for features associated with the outcome (based on Akaike's information criteria, model *c*-statistic and clinical review) and one that adjusted for disease classification alone. The difference in predicted outcome and corresponding 95% confidence interval (CI) were estimated for each modelling strategy on each outcome.

Mixed-methods case studies

Mixed-methods case studies were undertaken to understand why people make and regret decisions on their treatment. Longitudinal semistructured interviews (20 cohort participants from 13 sites) were conducted at baseline and 6 months later with framework analysis mapped findings to Witt's coping in deliberation framework and Sekhon's acceptability framework. We triangulated findings with baseline scores CollaboRATE shared decision-making (SDM) survey scores and 6-month decision regret (DR) scores.

Discrete choice experiment

An online survey using the discrete choice experiment (DCE) method was used to ask patients to choose their favoured treatment when presented with competing hypothetical treatment profiles. Regression analyses were conducted on DCE responses.

Clinician survey

A survey was developed following the CHERRIES statement checklist of recommendations to identify the most frequently used interventions for specific clinical scenarios in current PSD practice. It was

disseminated via the surgical trainee research collaboratives to practising consultants throughout the UK. Data were captured and stored in the REDCap software.

Consensus exercise and validation of a classification system

Two separate surveys were undertaken. The first comprised a Delphi consensus exercise in which surgeons and patients were asked to recommend best practice and further research. In the second, 15 surgeons were each asked to retrospectively assess photographs to quantify agreement in the Wysocki classification tool, with 90 patients each assessed by 6 surgeons.

Results

Cohort study

Participants

Thirty-one UK sites recruited participants over a 46-month period from May 2019 to March 2022. Seven hundred and twenty-nine participants consented to the study; after exclusions due to no procedure (n = 45), incorrect diagnosis (n = 7) and insufficient treatment information (n = 10), there were 667 participants included in the analysis cohort. Six-month follow-up data were available for 71% of participants; recurrence and complication data were available for 94% and 96% of participants, respectively.

Main results

Sixty per cent of patients (n = 397) received a major procedure; this comprised 272 (41%) asymmetric closure, 49 (7%) leave open and 76 (11%) midline closure. The remaining participants received minimal excision (n = 270, 40%), most commonly glue (n = 106, 16%) or pit picking (n = 60, 9%). Pain on day 1 and day 7 was higher for patients that received major procedures compared to minor procedures (augmented IPW-adjusted mean difference in day 1 pain 1.58 points, 95% CI 1.14 to 2.01 points, n = 536; mean difference in day 7 pain 1.53 points, 95% CI 1.12 to 1.95 points, n = 512). The difference was broadly consistent regardless of adjustment method.

Complications were reported by 207/385 (54%) participants that had major procedures, and 94/258 (36%) participants that had minor procedures (augmented IPW-adjusted risk difference 17.5%, 95% CI 9.1 to 25.9%, n = 579). Recurrence was reported for 86/373 (23%) of major and 87/256 (34%) of minor procedures; the augmented IPW-adjusted risk difference was −10.1% (95% CI −18.1 to −2.1%, n = 575). The estimated difference between groups was smaller for treatment failure (a composite of recurrence, failure to heal, and failure to return to normal activities), where 45% of major procedure participants versus 47% of minor procedure participants experienced treatment failure (augmented IPW-adjusted risk difference −2.3%, 95% CI −10.9 to 6.2%). The estimates for treatment difference in recurrence and complication were consistent across the various adjustment methods.

Treatment differences for time to healing and time to return to normal activities were less consistent between adjustment methods. Participants receiving major procedures took an estimated minimum of 20 days longer to return to normal activities (augmented IPW difference 25.9 days, 95% CI 18.4 to 33.4 days). Difference in time to healing was estimated to be at least 34 days more in the major group, but the estimates and corresponding CIs were inconsistent across models. At 6 months, around 25% of participants in both groups had wounds that had not healed; 12% of major procedure participants and 4% of minor procedure participants were yet to return to normal activity.

Comparisons of asymmetric closure with minimal excision produced results similar to the major versus minor comparisons.

Clinician survey

The survey was completed by 109 participants (54.5% response) who routinely cared for patients with PSD. Respondents reported a median caseload of 15 patients per year [interquartile range (IQR) 10–20]. Of those estimating their recurrence rates, 28.8% predicted they were in the 16–30% range. A wide range of treatment strategies were employed, with 65.1% practising an excision-and-leave-open technique despite this being considered obsolete as a treatment option. Surgical training in this obsolete technique was experienced by 83.1%. Many recommended non-surgical treatments despite the lack of evidence.

Mixed-methods case studies

No choice of treatment was offered to 9/20 participants, although this was not always seen negatively on the CollaboRATE instrument. Decision-making was influenced by prior experience of pilonidal sinus and anticipated recovery time. Participants scoring highest on DR also had among the highest SDM scores. The burden involved in wound care and the gap between expected and actual time for recovery were the principal reasons given for DR.

Discrete choice experiment

The survey was completed by 111 participants. The strongest predictor of treatment choice was risk of infection/persistence (attribute importance 70%), followed by shorter recovery time (attribute importance 30%). Patients were willing to trade off recovery time against risk of infection/persistence. Patients aged over 30 years had a higher risk tolerance (22.35–34.67%) for treatment failure if they could experience rapid recovery. Younger patients were willing to accept smaller risks (maximum acceptable risk 1.51–2.15) in exchange for a faster recovery. Patients were uniform in rejecting excision-and-leave-open because of the protracted nursing care it entailed.

Wysocki tool classification

The Wysocki tool demonstrated acceptable agreement. Overall, the kappa statistic for agreement was 0.52 (95% CI 0.42 to 0.61), with five of the six surgeons reaching consensus in 53 (59%) patients. Agreement did not differ with regard to the surgeon's experience. This, along with the prognostic value of the tool, suggests this could be used in practice to classify PSD.

Consensus exercise

The top five practice statements included: any treatment should not be worse than the disease itself; minimally invasive procedures should be preferred when feasible; surgeons should have opportunities to learn new techniques; a classification tool will help inform treatment options; and delayed return to activities is an important outcome.

The top five research statements included: future trials should compare broad groups of operations – minimally invasive procedures versus major excisional techniques, with stratification by severity of disease; a core outcome set and PROs should be used; there should be an attempt to develop an algorithm or decision tree to aid surgeon decision-making.

Conclusions

The burden of surgery for PSD is significantly greater than that reported in the literature. Many surgeons perceive this but continue to practise outdated procedures. While minimally invasive procedures may reduce this burden in many patients, they are not always offered when they could be. This practice may be driven by the desire to achieve cure at the expense of protracted recovery, but this is not always what patients want. Many would trade reduced chance of cure for more rapid recovery. Future practice priorities should follow the ethos of not making the surgery worse than the disease itself and appreciate that patients need to be properly informed about the burden associated with wound care and the likelihood of recurrence associated with different procedures, to minimise DR. Future trials should compare broad groups of interventions (minimally invasive vs. major excisions) stratified by disease severity and utilising a reliable and validated Wysocki classification system. Such trials should incorporate a core set of PROs.

Trial registration

This trial is registered as ISRCTN95551898.

Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: 17/17/02) and is published in full in *Health Technology Assessment*; Vol. 28, No. 33. See the NIHR Funding and Awards website for further award information.

Study protocol

The original protocol for the PITSTOP study can be found here. https://figshare.shef.ac.uk/articles/journal_contribution/The_PITSTOP_Study_PIlonidal_sinus_Treatment_STudying_the_OPtions_-_Protocol/7578242. A list of the protocol amendments since the initial Research Ethics Committee (REC) approval can be found below.

Version number	Changes made	Date of REC/ HRA approval	Amendment number
1.2	Errors in the flow diagram Errors that were overlooked such as age where it stated 18 years instead of 16 years Typological errors in the text Changes to the call at day one post surgery – this is now a phone call by a research nurse	03 May 2019/03 April 2019	SA01
1.2	Changes of collaborator Christine Moffatt University affiliation – previously University of Nottingham, now Nottingham Trent University (p7)		
1.2	Addition of optional consent item described which is the taking of photos pre-surgery to assist with the development of the classification system for PSD (p19)		
1.2	A follow-up call on day 1 post surgery as opposed to a text message (p19, p34, table 1)		
1.2	Removal for the 4–6 week time frame from routine clinic visit follow-up as this varies from clinic to clinic (p34)		

Version number	Changes made	Date of REC/ HRA approval	Amendment number
1.2	Addition of infection as outcome measure (p33)		
1.3	Clarification: Patients must have a minimum of 24 hours between receiving the Patient Information Sheet and consenting to the study (p16)	N/A/19 November 2019	MA07
1.3	Clarification: The postoperative routine clinic visit may also be conducted by the telephone if considered routine at site (p19)		
1.3	Clarification: The recruitment end date is July 2020 (p29)		
1.4	An additional telephone follow-up can be made if the participant does not attend the pre-arranged face-to-face clinic visit (p19)	N/A/03 December 2019	MA08
1.4	Clarification: Data can be collected by trained research personnel or a delegated member of the research team (p19)		
1.4	Addition of the REC reference on the front page (p1)		
1.5	Clarification: The photographs will be used to aid the validation of the classification system for PSD (p19)	02 March 2020/07 April 2020	SA03
1.5	Clarification: Preoperative photographs of the surgical site will be uploaded on the REDCap data capture system (p20)		
1.5	Clarification: A copy of the consent form will be uploaded on to the REDCap data capture system for monitoring purposes (p20)		
2.0	Clarification: Sponsor and CI e-mail addresses have been updated (p6)	03 August 2020/05 August 2020	SA04
2.0	Clarification: Study coordinating staff list has been updated (p8)		
2.0	Response to COVID-19: Addition of postal consent if preoperative consultations are conducted remotely OR social distancing measures inhibit research personnel to gain consent in person (p15)		
2.0	Surgeon survey is also accessible via an online link hosted on REDCap (p18)		
2.0	Research personnel may text a participant to arrange a follow-up call (p18)		
2.0	Response to COVID-19: Baseline measures can be taken over the phone (p19)		
2.0	Response to COVID-19: Follow-up data can be collected by study coordinating team if research personnel at NHS trust do not have capacity (p19)		
2.0	Response to COVID-19: missed 6-month follow-up data can be collected at the end of study (p19)		
2.0	The DCE questionnaire can be completed by non-PITSTOP patients. The questionnaire can be advertised using a leaflet and/or social media (p24)		
2.1	Research personnel may e-mail a participant to arrange a follow-up call (p18)	N/A/27 October 2020	MA11
2.2	Update to study coordinating team (p7)	N/A/29 September 2022	NSA13
2.2	Clarification of study procedures for the PSD seton surgical procedure (p19)		

Version number	Changes made	Date of REC/ HRA approval	Amendment number
2.2	Clarification of follow-up procedures (p20)		
2.2	Update to recruitment end date (p30)		
3.0	Amendment to consensus technique (p25)	18 November 2021	SA05

Chapter 1 Introduction

Background

Pilonidal disease is a common condition that affects around 26/100,000 of the population – predominantly young, working people.[1] The term 'pilonidal' derives from the Latin words for hair (pilus) and nest (nidus). It is an acquired disease resulting in obstruction of hair follicles in the natal cleft (the anatomical groove between the buttocks). Subsequent rupture of the follicles leads to abscess and sinus formation. Risk factors for development of the condition include male gender, extensive body hair, young adulthood, family history, local trauma, sedentary lifestyle, poor hygiene, an anatomically deep natal cleft and obesity.[1–3] Once established, the condition persists and progresses through insertion of ingrown or loose hairs into the sinuses.[2,3] The term pilonidal sinus disease (PSD) encapsulates a wide spectrum of abnormalities ranging from relatively asymptomatic simple sinuses to complex abscess cavities with multiple sinus tracks that persist despite repeated surgical intervention. Individuals present either as an emergency with a painful abscess between the buttocks or electively with a chronic cycle of pain and discharge from the sinuses, causing significant disruption to employment, relationships and quality of life (QoL).[4]

The ideal management of PSD should be simple, safe, cost-effective, easy to perform and lead to a rapid return to normal activities, with low rates of acute wound complications (including infection, seroma, haematoma), recurrence and rapid wound healing. These aims are not reliably delivered by current surgical practice and there is no consensus on how to manage based on disease characteristics.

Pattern of disease and management options

Patients often present with acute infection and abscess formation. Abscesses usually require hospitalisation with incision and drainage of the abscess cavity. One in five patients present with recurrent symptoms following emergency surgery.[5] This picture of relapsing and remitting infections is typical of chronic PSD.

Treatment of chronic PSD is surgical, usually using two essential components: excision and closure. The exceptions are phenol injection and fistuloscope/diathermy as stand-alone treatments (which aim to induce fibrosis and obliterate the tracks) and seton insertion which may induce fibrosis, allowing the possibility of a simpler subsequent surgery. There is no clear consensus as to which approach for each component is superior. For those procedures that involve excision, the tissue removed may be minimal (e.g. curettage or excision of the 'pit') or there may be substantial excision of the affected area and surrounding tissue to ensure complete removal of disease. The resultant wound may be left open to heal slowly by secondary intention, or it is closed with glue[6–8] or sutures. The skin closure technique may be midline or off-midline. In the off-midline technique, the wound is positioned adjacent to the natal cleft, rather than in the natal cleft itself, in order to theoretically aid healing.[8,9] Examples include: the Karydakis flap, Bascom cleft closure (Bascom II), rhomboid and Limberg flaps.

Monetary and humanistic burden

Pilonidal disease is relatively common and represents a significant burden to primary and secondary care in the NHS. The 2012 hospital episode statistics (HES) data reported 13,239 hospital admissions for PSD.[10] At present, both emergency and the most common elective excisional surgical treatments

leave large open wounds that may take months to heal.[6,7] Patients consequently require prolonged wound care from community healthcare services. As the disease tends to affect young otherwise healthy adults, this prolonged need for dressings and general wound care impacts on education, work, intimacy and social life, pain, recurrent infection and fear of wound deterioration, all severely affecting QoL.[11,12] Alternative techniques including minimally invasive interventions that aim to close the wound away from the midline may reduce the burden to the patient, but their efficacy outside the care of dedicated enthusiasts is not clear.

Current evidence base

The optimum treatment that is both easy to perform and results in rapid healing and minimal complications is not clear. This is reflected in varied practice throughout the UK with a perceived random selection of the procedure techniques detailed above. Some of these procedures result in lengthy healing times and long periods of incapacity. The literature on PSD is large but mainly consists of single-centre cohort studies looking at individually favoured techniques. Many of these have reported very low recurrence and infection rates for almost all procedures.[13] It has proven difficult to replicate these results in 'real life'. In addition to the literature being mainly from single-centre cohorts, most studies make no attempt to stratify patients or detail the extent of disease or the adjuvant management (antibiotics, anaesthetic, postoperative care). There have been numerous randomised controlled trials (RCTs), and nearly 40 systematic reviews that focus on management – including two Cochrane Reviews. Most of these systematic reviews include meta-analyses of cohort studies only or analyse comparative RCTs and non-RCTs (often combining these data) for numerous interventions with varied controls. The methodological flaws of many individual studies and systematic reviews, coupled with the uncertainty of front-running interventions and an absence of a universally accepted control, make the value and interpretation of the data difficult.[14]

The first Cochrane Review demonstrated that healing through secondary intention had lower overall recurrence rates compared to primary closure but at the expense of longer healing times.[14] Another systematic review reached the same conclusion but also compared two types of closure, suggesting off-midline to be preferable to midline.[15] The authors also concluded that outcome measures, such as time to healing, were poorly analysed, and health economic data were lacking. They proposed that future trials should be adequately powered, multicentric and include valid methods of assessing surgical outcomes. A systematic review of wound care after excision found no best practice guidelines and only one clinical pathway.[16]

The second Cochrane Review focused on fibrin glue (FG) in the treatment of PSD.[17] The authors concluded this was a promising and appealing option as monotherapy given the non-invasive nature and that it could be performed as a day-case procedure, under local anaesthesia. These conclusions echo the conclusions of a previous meta-analysis, both suggesting a need for further research.[18] Nevertheless, the research to date has largely considered FG as an adjunct to surgery and although small, single-centre observational studies[6,7,19–21] have been published, there is no RCT of FG as monotherapy in treatment of PSD.

Rationale

Currently, there is a lack of evidence regarding classification of disease, what are the front-running interventions, whether there is clinical equipoise for these interventions and whether comparative studies for these interventions are feasible in terms of recruitment, and finally what outcome measures are relevant to patients, can be easily and reliably measured and are sufficiently sensitive to change. Given the efficacy uncertainty surrounding a multitude of operative techniques, compounded by the reported negative implications for recovery, there is a need to improve the evidence base to guide future pilonidal management.[14]

Research objectives

The aim of the PITSTOP study is to answer the following research questions:

1. What are the different subtypes of pilonidal disease for which the various treatment options are indicated?
2. What combinations of excision and closure techniques are used?
3. Which outcomes do patients value and which interventions do they prefer?
4. What further research is needed?

To answer the research questions posed, we aimed to complete the following:

1. Conduct a survey of clinicians to assess management preferences.
2. Follow patients with symptomatic pilonidal sinus referred to each collaborating site, prospectively record details of their pit/track anatomy, surgical management, medical events and health-related QoL until 6 months after their operation.
3. Describe the combination of interventions currently in use and quantify clinical and patient-reported outcomes (PROs) associated with each.
4. Identify patient-specific disease features that might predict poor outcome in each treatment group by risk-modelling methods.
5. Derive a case mix-adjusted estimate of the risks associated with common treatment options, using causal inference methods to provisionally rank the optimal management strategies among patients for whom more than one treatment is considered appropriate.
6. Provide an overview of patient views and experiences.
7. Collect the views of patients on which interventions they would rather avoid and which outcomes they most value.
8. Validate a classification system.
9. Reach a surgeon-based consensus on which subtypes of pilonidal disease may benefit from which treatment options.
10. Reach a surgeon and patient-based consensus on research priorities.

Chapter 2 Consultant surgeon survey

Methods

Survey design and development

A survey was developed to identify the most frequently used interventions for specific clinical scenarios in current PSD practice. As this was a novel survey, it was designed by study collaborators and followed the CHERRIES statement checklist of recommendations.[22] The survey included questions on the following: the mean number of primary elective procedures performed annually, factors affecting choice of procedure, treatment choice for recurrent disease presentation and the factors affecting treatment choice for recurrent disease treatment. The survey was piloted to determine the clinical sensibility.

Data collection

To maximise completion rates, the survey could be completed online or on paper. The online survey was hosted on the Research Electronic Data Capture™ (REDCap) system managed centrally by the University of Sheffield Clinical Trials Research Unit (CTRU). REDCap (Research Electronic Data Capture) is a secure, web-based software platform designed to support data capture for research studies, providing 1) an intuitive interface for validated data capture; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for data integration and interoperability with external sources.[23,24] Paper surveys were returned by post or via e-mail. The questionnaires were anonymised at the respondent level.

Sampling

The survey was disseminated via the UK surgical trainee research collaboratives, led jointly by the South Yorkshire Surgical Research Group and the North-West Research Collaborative. Collaborators were asked to deliver the questionnaire to consultant colorectal surgeons in their units. The first point of contact was made through the National Research Collaborative e-mail lists, and electronic contact to local collaborative leads was cascaded locally. The collaborators were asked to circulate the survey locally to three consultants and thereafter return the completed questionnaires to the REDCap system.

Data analysis

All aspects of data management were provided by the CTRU in accordance with their standard operating procedures. The data emanating from this survey were captured and stored in the REDCap software.

Results

The link was followed by 200 surgeons and completed by 113 participants. Of these, 109 routinely cared for patients with PSD. These 109 were entered into the final analysis, giving a final response rate of 54.5%.

Respondent practice overview

Respondents reported a median caseload of 15 patients per year [interquartile range (IQR) 10–20 patients] and indicated that recurrent disease accounted for 20% of overall workload (IQR 10–30%). Of those estimating their recurrence rates (n = 97), 19 (19.5%) were unaware of their recurrence rate, 14 (14.4%) estimated their rate to be <5%, 36 (37.1%) to be in the 6–15% range, and 28 (28.8%) in the 16–30% range.

With regards to hair management, depilation was recommended by 54 (49.5%), laser hair removal by 32 (29.4%), salt baths by 14 (12.8%), shaving by 52 (47.7%) and waxing by 32 (29.4%).

Operative strategies employed

A wide range of treatment strategies were employed by responding surgeons, summarised in *Table 1*. Excision of disease with wound left open was the most frequently used strategy (71 responses; 65.1%), followed by Karydakis flap (62 responses; 58.1%). Curettage with phenol injection (1 response; 0.9%) and endoscopic pilonidal sinus treatment (EPSiT) (2 responses; 1.8%) were the least frequently performed interventions.

Participants were asked to provide a first-, second, and third-choice preference for their interventions. Karydakis was the first-preference treatment for 24/96 respondents (25.0%), followed by Bascom's II (*n* = 18; 18.7%), and curettage and glue (*n* = 15; 15.5%). For second-preference treatments, local excision with wound left open was the most popular with 21/85 participants (24.7%), followed by local excision with midline closure for 15 (17.6%) and Karydakis procedure for 14 (16.4%). The most popular third-preference treatment was local excision with wound left open (27/32; 84.4%), followed by local excision with midline closure for 12 (37.5%), and Bascom's II for 7 (21.9%) respondents (see *Appendix 3*, *Figure 12*).

Case vignettes

Case vignettes demonstrated heterogeneity across respondents. Case one (recurrent disease) showed a preference for rhomboid flap or 'other' procedures (22.6% and 25.5%, respectively). For case two (female with primary disease and cosmesis concerns), preferences turned to favour conservative management (21.6%), followed by excision and primary closure (16.0%) and cleaning/curettage of tracts (14.1%). Case three assessed recurrent disease and requirement for minimal time off work. For this scenario, most respondents opted for conservative management with hair removal (25.4%), followed by curettage of tracts (16.0%). Of note, 15.1% would offer a Karydakis procedure in this setting. Responses are summarised in *Table 2*.

Training

Surgical training programmes were the key training setting for commonly offered procedures. These included training in wide local excision with wound left open or closed for 59/71 (83.1%) and 36/48 (75.0%) of those offering the procedures, respectively. Similar numbers were seen for Bascom's I (21/27; 77.7%) and Karydakis procedure (49/62; 79.0%). For some procedures, no formal training was reported by 5–10% of respondents. Courses, observation of colleagues and reference material such as text or video was also variably used. A summary of training experiences is presented in *Table 3*.

TABLE 1 Summary of operations offered

Operation	Yes, *N* = 109
Excise and leave open	71 (65.1%)
Karydakis	62 (56.8%)
Excise and midline closure	48 (44.0%)
Bascom's cleft lift	47 (43.3%)
Rhomboid flap	30 (27.5%)
Bascom's I	27 (24.7%)
Curettage and glue	17 (15.5%)
Pit picking alone	10 (9.2%)
Other flap	7 (6.4%)
EPSiT	2 (1.8%)
Curettage and phenol	1 (0.9%)

TABLE 2 Case vignette responses

Operation	16-year-old male, six previous surgeries with other surgeons, has recurrent disease and partially open wound/sinus 1 cm long in natal cleft that has been like that for 9 months. Wants to play contact sport. Parents not happy (*N* = 107); *n* (%)	19-year-old female, fair skin, dark hair, previous abscess drainage, swelling and discomfort in natal cleft, very worried about cosmesis and what the scar will look like if you operate, *N* = 106; *n* (%)	30-year-old male plumber who has had previous surgery, no details available, and now present with recurrent disease. Single discharging pit around scar, and can't afford much time off work, *N* = 106; *n* (%)
Bascom's cleft lift procedure	12 (11.3%)	7 (6.6%)	12 (11.3%)
Bascom's I procedure	2 (1.9%)	13 (12.2%)	6 (5.7%)
Cleaning/ curettage tracts	7 (6.5%)	15 (14.1%)	17 (16.0%)
Conservative/ hair removal	14 (13.2%)	23 (21.6%)	27 (25.4%)
Excision and primary closure	0 (0%)	17 (16.0%)	9 (8.5%)
Karydakis procedure	11 (10.3%)	13 (12.2%)	16 (15.1%)
Lay open ± marsupialisation	9 (8.4%)	6 (5.7%)	8 (7.5%)
Other	27 (25.5%)	10 (9.4%)	10 (9.4%)
Rhomboid flap	24 (22.6%)	1 (0.9%)	1 (0.9%)
Z-Plasty flap	1 (0.9%)	1 (0.9%)	0 (0%)

Discussion

Overview

The key finding of this survey is the heterogeneity and variation in practice of consultant colorectal surgeons who treat PSD. It demonstrates a relatively low annual volume of operative procedures (around 15) when compared to other conditions such as colorectal cancer surgery. This number is slightly higher than the median of four cases per year identified through HES.

One in four surgeons perceived they had treatment failure rates approaching 30% in this study. This is somewhat at odds with the published literature, which often claims cure in > 90% of cases.[13] This demonstrates dissonance between published reports and real-world experience of clinicians. Conversely, 1 in 10 respondents reported recurrence rates of < 5%, which is concordant with the literature. This gap in outcomes may arise from issues with the quality of the literature, where the often surprisingly high quality of outcomes has been challenged.[14] Alternatively, it may reflect a small group of clinicians with a high volume of practice and associated good outcomes.[25] This poses three key questions. First, should complex pilonidal disease, or even all pilonidal disease, be managed by a group of high-volume surgeons? This is an approach that is advocated in other aspects of surgery such as rectal cancer and inflammatory bowel disease.[26,27] The second question to ask is whether we should improve training opportunities (highlighted here as limited) for colorectal surgeons to improve their skill set. Finally, we should ask whether there is a need for better monitoring of outcomes in PSD, as with

TABLE 3 Training in different procedures

	Number offering	No formal training	Course/ workshop	Observed colleagues	Training in registrar programme/ fellowship	Videos/text
Wide local excision, leave open	71	4/68 (5.8%)	1/68 (1.5%)	4/68 (5.9%)	59/68 (85.2%)	0/68 (0%)
Wide local excision with closure	48	4/48 (8.3%)	1/48 (2.1%)	6/48 (12.5%)	36/48 (75.0%)	1/48 (2.1%)
Bascom's cleft lift	42	2/26 (7.7%)	4/26 (15.3%)	9/26 (34.6%)	9/26 (34.6%)	2/26 (7.7%)
Pit picking/Bascom's I	27	3/34 (8.8%)	4/34 (11.7%)	4/34 (11.7%)	21/34 (61.8%)	2/34 (5.9%)
Karydakis	62	2/75 (2.7%)	4/75 (5.3%)	13/75 (17.3%)	49/75 (65.3%)	7/75 (9.3%)
Rhomboid flap	30	2/29 (6.9%)	2/29 (6.9%)	5/29 (17.2%)	15/29 (51.7%)	5/29 (17.2%)
Other flap	7	1/5 (20.0%)	0/5 (0%)	2/5 (40.0%)	2/5 (40.0%)	0/5 (0%)
Curettage and glue	17	0/22 (0%)	5/22 (22.7%)	7/22 (31.8%)	5/22 (22.7%)	7/22 (31.8%)
Curettage and phenol	1	0/1 (0%)	0/1 (0%)	0/1 (0%)	0/1 (0%)	0/1 (0%)
EPSiT	2	0/2 (0%)	0/2 (0%)	½ (50%)	0/2 (0%)	0/2 (0%)

registries established for other conditions. The findings of this study largely match those from a previous survey conducted in 2011,[28] suggesting little has changed in a decade, making these questions more important to improve care.

Surgeons expressed preferences for some treatments such as excise and leave open. The literature suggests these should be considered outdated as they are associated with significant wound morbidity.[15] Sixty-five per cent of surgeons used the leave open technique with healing occurring by secondary intention, and 44% used a midline closure technique. Surgeons expressed a stronger preference for asymmetric closure than when previously assessed,[28] in keeping with global trends.[29] In contrast, minimally invasive techniques such as EPSiT and pit picking appear to be relatively unpopular treatments. This suggests that surgeons are focused on cure rather than symptomatic relief. This may contrast with stated patient preferences where they are willing to trade a less major procedure in exchange for a higher risk of recurrence.[30]

The survey does have limitations. Surveys of this nature can present artificial choices and do not permit qualification of answers. The use of vignettes does, however, allow some direct comparison of preferences. The survey may have drawn in experts and enthusiastic practitioners in pilonidal disease. However, the heterogeneity of responses does not reflect unity of thought, and responses are in keeping with published surveys. The response rate of 54.5% should reassure us as to representativeness of the survey.

Implications for policy-makers

This study presents three key actions for policy-makers. First, there is a need to agree a general framework for interventions to standardise pathways of care. This should be supported by best available evidence. Where such evidence is not available, funding should be secured to inform such guidance. Secondly, this is an area with no clear registry or oversight. Policy-makers should consider whether collection of granular data on practice and outcomes might aid initiatives to improve care, or even justify specialisation or centralisation of practice. Finally, surgeons' practice is driven by their postgraduate training and persists into their independence. Therefore, it is important to offer opportunities for further

training in new techniques. This would allow implementation of new techniques and may support the decommissioning of outdated procedures.

Implications for researchers

The level of heterogeneity likely speaks in part to uncertainty. Researchers should consider whether the findings presented here would support the delivery of specific procedures in a head-to-head RCT, or whether a 'bucket' approach would be more pragmatic.

Conclusion

This survey demonstrates significant heterogeneity in surgeon practice preference. It suggests that limited access to training opportunities may impede efforts to improve practice in the area.

Chapter 3 Cohort study

Methods

Aims

A prospective, multicentre observational cohort study was conducted to:

- describe the disease characteristics of participants undergoing treatment for PSD
- describe procedures currently in use and quantify clinical outcomes and PROs associated with each
- identify patient-specific disease features that might predict poor outcome in each treatment group
- derive a case mix-adjusted estimate of the risks associated with common treatment options.

Participants

Eligible patients were undergoing definitive elective treatment for PSD at study recruiting centres.

Inclusion criteria

- Consenting patients aged 16 years or older and with PSD.

Exclusion criteria

- Asymptomatic disease.
- Currently pregnant.
- Unable to give consent.
- Acute abscess.
- Hypersensitivity to the sealants.

Study procedures

Recruitment and consent

Patients considered suitable for surgery were identified from general practitioner (GP) secondary care referrals, surgery waiting lists or clinics. Once identified, participants were given an approved participant information sheet detailing the study – sent in the post or provided in person. Patients were invited to attend a recruiting clinic. At the clinic, a member of the research team explained the study to the participant and offered them an opportunity for them to ask any questions. The principal investigator (PI) or delegated research team member confirmed eligibility and ensured written informed consent was obtained prior to any patient data being collected. Participants were advised that they were able to withdraw from the study at any point without any impact on their routine NHS care. As is standard practice, the surgeon would discuss the condition, possible interventions and their advantages and disadvantages. Patients were given a minimum of 24 hours between receiving the participant information sheet and consenting to the study.

In response to the COVID-19 pandemic, adjustments were made to allow the continuation of recruitment and consent procedures. Consent could be obtained by post. An invitation letter, participant information sheet and a postal consent form were posted to the patient. The research team were able to contact the patient to provide an overview of the study and answer questions. The patient was instructed to complete two consent forms: one to be returned to the research team, one for their own records. Once received, the research team contacted the patient to complete the postal consent review form.

Intervention

The study was observational, and surgeons were not asked to change their usual practice. Surveys suggested that around six procedures were in common use,[28,31–33] which can broadly be described as:

- major excision with asymmetric closure ('Bascom II cleft lift')
- major excision with lateral closure ('Karydakis')
- major excision with lateral closure with rhombic flap
- major excision with midline closure
- major excision and leave open
- minimal excision ('Bascom I' or 'pit picking').

Other approaches include curettage ('scraping out') or phenol injection with glue closure. For the purpose of analysis, procedures were classified as either 'minor' or 'major' procedures, with the latter further subdivided as 'asymmetric/lateral closure', 'midline closure' or 'leave open'.

Data collection

Data were collected by trained research personnel. All patient data were recorded on the case record form. Copies of the consent and patient information sheets were kept in the participant's hospital case notes. A copy of the consent form was uploaded onto the REDCap data capture system for monitoring purposes. All data were recorded on the REDCap data capture system.

Assessment schedule

Participants completed baseline questionnaires after eligibility and consent were confirmed. Details of the procedure were collected on the day of procedure. Outcome data were collected on days 1 and 7 after procedure and then at an in-person clinic visit and a further follow-up 6 months after the procedure. Participant data could also be collected opportunistically at a final 'study completion' visit. The outcomes collected are listed in *Recruitment and participant flow*.

Safety assessments

Participants were asked to report complications at days 1 and 7 post procedure, and again at the clinic visit and 6 months. Participants were prompted specifically for incidence of bleeding, seroma, haematoma, infection, dehiscence, maceration, flap necrosis or discharge. Other adverse events (AEs) were collected only if considered related to the study treatment.

Statistical methods

Sample size

The study aimed to recruit approximately 800 patients, with at least 100 within each of the front-running management strategies. Doing so allows proportions to be estimated within each management strategy to a standard error of ≤ 5% and pain numeric rating scale to within a standard error of 0.2 points, assuming that the standard deviation (SD) of a 10-point scale would not exceed two units.

Outcomes

The outcomes and their timing and description are listed in *Table 4*. All outcomes were self-reported aside from the clinician-assessed scarring question. No single primary outcome was prespecified in this study; methods to elicit outcomes of most importance to study participants are described in subsequent chapters.

Comparisons

The sample size precluded reliable comparisons between specific subtypes, with only one procedure (Karydakis) providing at least 100 participants with non-missing outcome data. Instead, we undertook the following risk-adjusted treatment comparisons on broader categories of procedure types:

TABLE 4 List of outcomes

Name/timing	Description
Pain [baseline, day 1 (current pain only), day 7, clinic visit and 6 months]	1. Rating of current pain related to PSD. 0 (no pain) to 10 (worst pain imaginable). 2. Rating of worst pain related to PSD in last week. 0 (no pain) to 10 (worst pain imaginable).
Health status (baseline, day 7, clinic visit and 6 months)	1. EQ-5D-5L health status. A score of 1 equates to perfect health, 0 is to a state comparable to death and negative scores to a state worse than death. 2. EQ-5D thermometer health status. Measure made by marking a point on a 'thermometer' scale. A score of 100 equates to the best health imaginable, and score of 0 represents the worst health imaginable.
Impression of shared decision-making (baseline)	1. CollaboRATE mean score. The average of three questions, each scored 0 (no effort was made) to 4 (every effort was made). 2. CollaboRATE top score, defined as 'every effort made' if all three questions are answered 4, and 'not every effort made' otherwise.
Return to normal activities (days 7, clinic visit and 6 months)	Time from procedure to return to normal activities (censored if not returned at last contact).
Length of time to healing (clinic visit and 6 months)	Time from procedure to wound healing (censored if not returned at last contact).
Recurrence[a] Treatment failure[b]	1. Recurrence of PSD (yes/no).
Wound impact (clinic visit and 6 months)	Cardiff Wound Impact Questionnaire (CWIQ): 1. QoL, integer-rated from 0 (worst possible) to 10 (best possible). 2. Satisfaction with QoL, integer-rated from 0 (worst possible) to 10 (best possible). 3. Impact on physical symptoms and everyday living, ranging from 0 (least impact) to 100 (greatest impact). 4. Impact on social life, ranging from 0 (least impact) to 100 (greatest impact). 5. Impact on well-being, ranging from 0 (least impact) to 100 (greatest impact).
Decision regret (6 months)	Decision regret scale based on five questions, ranging from 0 (least regret) to 100 (greatest regret).
Scarring (6 months)	1. Clinician-assessed presence of an undesirable scar. 2. Scar itching. 3. Scar pain.
Complications (days 1 and 7, clinic visit and 6 months)	Presence of complications (bleeding, seroma, infection, flap necrosis, haematoma, maceration, dehiscence, discharge other related to procedure).

EQ-5D-5L, EuroQol five dimensions five levels questionnaire; EQ-5D, EuroQol-five dimensions questionnaire.
a Patients were categorised as having experienced recurrence if:
- They had undergone further procedure for their PSD (excluding expected procedures of repacking and replacement/removal of dressings).
- They reported recurrence.
- They had a reported AE or other complications which were consistent with an unresolved PSD (as reviewed by the chief investigator).

b Patients were categorised as having treatment failure if:
- They were generalised as recurred based on the definition above.
- They reported not having returned to normal activity during the follow-up period.
- They reported the wound not having healed during the follow-up period.

- any major excisional procedure versus any minor procedure
- any major excisional procedure with asymmetric closure versus any minor procedure (minimal excision).

Outcomes were summarised descriptively in relation to the treatment received for the less broad treatment groupings whose number did not permit risk-adjusted modelling.

Methods for obtaining risk-adjusted comparisons

Procedures were compared using risk-adjusted methods to reduce bias due to treatment selection, since the extent of disease is likely associated with both the type of procedure and the response.

Statistical methods for risk-adjusted outcomes

Three broad approaches were taken to risk-adjust these comparisons.

1. Regression modelling.

Risk adjustment was undertaken separately to attempt to adjust for imbalance in prognostic features across the procedure groups. Each outcome was modelled separately since features do not affect all outcomes equally. For each outcome, three models were fitted:

i. all features
ii. features associated with the outcome
iii. the Wysocki disease classification alone.

Model 2 is a 'compromise' model which trades off missing potentially important features against model parsimony (overfitting) and the impact of missing covariate data. Covariates were removed on the basis of Akaike's information criteria (AIC) and the size of the *c*-statistic of the model. All models were discussed and agreed with a core study clinical team prior to revealing comparative data.

Continuous outcomes were modelled using linear regression, and differences with 95% confidence intervals (CIs) between treatment groups were estimated from the regression coefficient for the procedure group. Binary outcomes were modelled using logistic regression and absolute differences in proportions were assessed using the difference in marginal probabilities. Time to wound healing and time to return to normal activities were modelled using either Cox regression or parametric accelerated survival time, the choice of which depended on which fitted best to the distribution. Proportional hazards were assessed using scaled Schoenfeld residuals and the Grambsch–Therneau test, and the fit of parametric survival distributions was assessed using Q-Q plots.[34] The parametric model was chosen as the lowest AIC among four different approaches (Weibull, log-normal, log-logistic and generalised gamma).

2. Propensity score approaches.

The second approach used a different approach which attempts to balance treatment groups based on features affecting the choice of procedure they received, rather than features associated with outcome. Two approaches were taken.

i. inverse probability weighting (IPW)
ii. nearest neighbour matching.

Features were assessed using logistic regression in which treatment choice was the outcome. Covariates were identified analogously to model 1(ii) above. The same propensity score adjustments were used for

each outcome. The propensity score-adjusted models were then used to calculate predicted outcomes in both arms, following which their difference and 95% CI were estimated.

Linear and logistic regression models were used for continuous and binary outcomes, respectively. Time-to-event outcomes were fitted within the propensity score framework only if the assumption of accelerated failure time distributions was met based as outlined above.

3. Augmented IPW/IPW with regression adjustment.

The final approach is a combination of approaches 1 and 2 which simultaneously models both treatment selection and outcome using the same covariates used in 1(ii) and 2. The differences in predicted outcomes and their 95% Cis were estimated for binary or continuous outcomes. Time-to-event outcomes (wound healing and return to normal activities) were compared using IPW with regression adjustment. In this, treatment selection was balanced using IPW, and the outcome was modelled adjusting for covariates in 1(ii). The AIC was used to select the best-fitting distribution (Weibull, log-normal, log-logistic or generalised gamma).

Factors affecting outcome and choice of treatment

Previous publications have suggested several possible risk factors which may affect outcomes. The following factors were considered as potential risk factors for poor outcome:

- Demographic features: sex, body mass index (BMI), depth of natal cleft, presence and type of gluteal hair (none, mild, dense) and smoking status.
- Disease characteristics: pit density (number of pits divided by spread of pits), presence of unilateral or bilateral disease, distance from furthest lateral opening to the nearest pit, presence of pus and Wysocki disease classification.

The same features were also assessed for their association with choice of procedure.

Statistical assumptions

All modelling approaches make assumptions (some of which are untestable), and no single method is clearly optimal. Approaches 1 (regression) and 2 (propensity score) are unbiased only if the models incorporate all relevant features and are correctly specified. Approach 3 is termed a 'doubly robust' method and is unbiased if either of the two-component models is correctly specified, but is more complex and more susceptible to overfitting. In view of this, no single method was identified as the primary risk adjustment. Instead, the findings from models were assessed for their consistency and, where they provided conflicting estimates, the plausibility of each model was considered.

A preliminary assessment of overlap was undertaken prior to any modelling in order to ensure that different treatments had non-zero probability of uptake in different subgroups.[35]

All participants undergoing a procedure were included in the analyses.

Patient and public involvement

Two patient and public involvement (PPI) representatives joined the study. The PPI representatives reviewed all patient-facing documents to ensure readability, understanding and format. One PPI representative sat on the steering committee panel and provided an instrumental patient voice in the management of the study. This PPI representative suggested the inclusion of a supplementary participant information sheet to be available for patients who were waiting for their clinic appointment. The patient representatives were also consulted when writing the plain English summaries and dissemination materials.

Results

Recruitment and participant flow

Thirty-one UK sites recruited participants over a 46-month period from May 2019 to March 2022 (see *Appendix 2*). *Figure 1* shows the flow of participants through the cohort: in total, 729 participants consented to be part of the cohort study. Participants were excluded from analyses if they did not have a procedure during the study (*n* = 45), if they were ineligible due to an incorrect diagnosis (*n* = 7), or if there was not enough information provided in order to categorise their procedure (*n* = 10). A total of 667 participants were included in the analysis cohort, of whom 476 (71%) provided follow-up data at 6 months. The number of participants who consented was lower than our anticipated sample size of 800, with 100 patients in each of the front-running management strategies.

In response to COVID-19, participants were able to be followed up at the end of the study; this was completed for 574 participants and used to update complications and recurrence data.

Baseline characteristics

The characteristics of the cohort participants are included in *Tables 5* and *6*. There were more males (73%), and 85% of participants were white; the average age of participants was 29 years. Just over half (54%) of participants reported no previous procedures, and 22% reported a previous elective procedure for PSD.

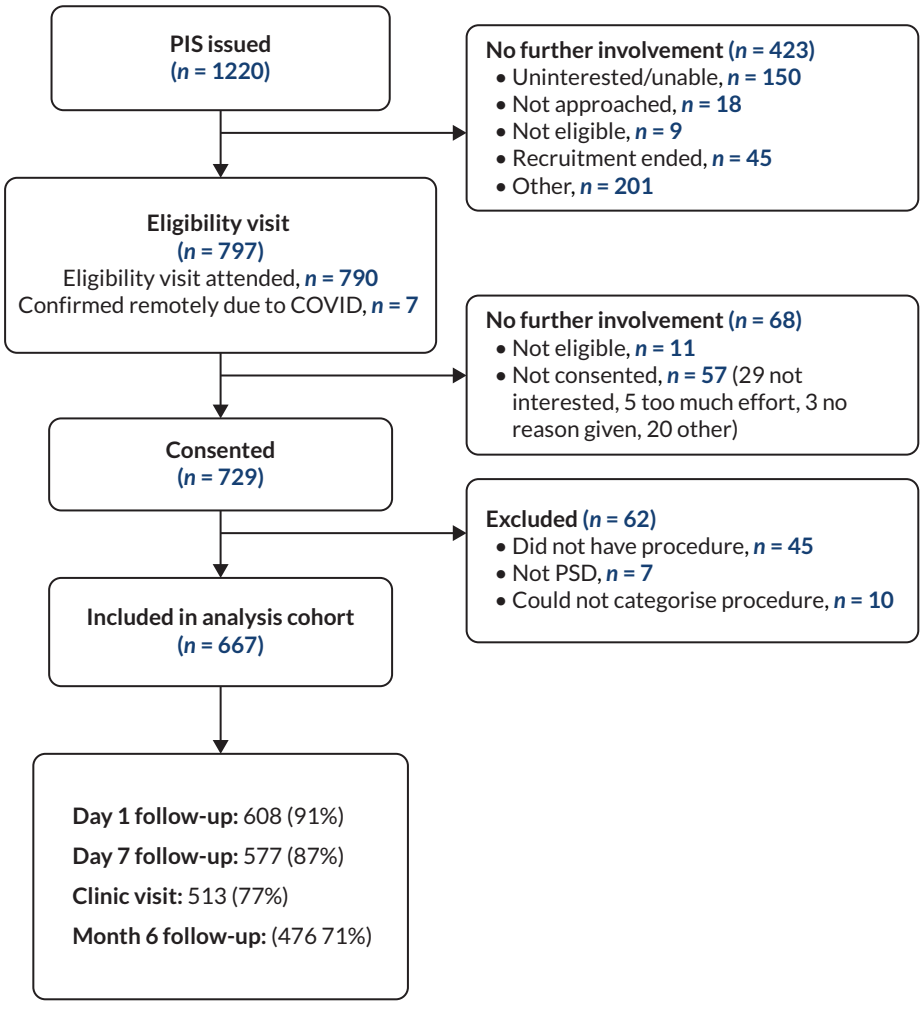


FIGURE 1 Participant flow for the PITSTOP study. PIS, patient information sheet.

TABLE 5 Participant characteristics

Characteristic	Asymmetric closure (*n* = 272)	Leave open (*n* = 49)	Midline closure (*n* = 76)	Minimal excision (*n* = 270)	All (*n* = 667)
Age					
N (%)	272 (100%)	49 (100%)	76 (100%)	270 (100%)	667 (100%)
Mean (SD)	28.5 (9.0)	28.1 (10.9)	28.1 (7.7)	29.7 (9.9)	28.9 (9.4)
Median (IQR)	27.0 (22.0–32.0)	25.0 (20.0, 33.0)	27.5 (22.5–31.5)	28.0 (23.0–34.0)	27.0 (22.0– 33.0)
Min, max	16.0, 73.0	16.0, 64.0	18.0, 58.0	16.0, 69.0	16.0, 73.0
BMI (kg/m^2)					
N (%)	253 (93%)	47 (96%)	71 (93%)	241 (89%)	612 (92%)
Mean (SD)	29.5 (5.5)	28.9 (6.9)	28.9 (5.0)	28.7 (6.1)	29.1 (5.8)
Median (IQR)	28.8 (25.5–32.8)	28.1 (23.0– 32.7)	28.0 (25.1–32.7)	27.8 (24.2–32.1)	28.3 (24.9– 32.7)
Min, max	17.6, 59.5	17.7, 49.2	17.0, 39.7	13.1, 47.6	13.1, 59.5
Number of baths and/or showers in a typical week					
N (%)	260 (96%)	49 (100%)	74 (97%)	261 (97%)	644 (97%)
Mean (SD)	6.9 (2.9)	7.8 (3.8)	7.6 (2.7)	6.8 (2.4)	7.0 (2.8)
Median (IQR)	7.0 (5.5–7.0)	7.0 (7.0–7.0)	7.0 (7.0–7.0)	7.0 (6.0–7.0)	7.0 (6.0– 7.0)
Min, max	1.0, 27.0	4.0, 21.0	3.0, 14.0	1.0, 14.0	1.0, 27.0
Sex					
Male	183 (67%)	36 (73%)	60 (79%)	206 (76%)	485 (73%)
Female	89 (33%)	13 (27%)	16 (21%)	64 (24%)	182 (27%)
Ethnicity					
White	228 (84%)	41 (84%)	72 (95%)	229 (85%)	570 (85%)
Asian/Asian British	23 (8%)	4 (8%)	4 (5%)	27 (10%)	58 (9%)
Mixed/multiple ethnic groups	9 (3%)	1 (2%)	0 (0%)	3 (1%)	13 (2%)
Black/African/ Caribbean/Black British	3 (1%)	2 (4%)	0 (0%)	3 (1%)	8 (1%)
Other ethnic group	4 (1%)	0 (0%)	0 (0%)	3 (1%)	7 (1%)
Prefer not to say	3 (1%)	1 (2%)	0 (0%)	1 (0%)	5 (1%)
Seated for more than 6 hours in a working day	142 (52%)	19 (39%)	36 (47%)	135 (50%)	332 (50%)
Smoking status					
Non-smoker	148 (54%)	31 (63%)	43 (57%)	152 (56%)	374 (56%)
Current smoker	86 (32%)	13 (27%)	26 (34%)	71 (26%)	196 (29%)
Current e-cigarette smoker	13 (5%)	0 (0%)	4 (5%)	20 (7%)	37 (6%)

continued

TABLE 5 Participant characteristics (*continued*)

Characteristic	Asymmetric closure (*n* = 272)	Leave open (*n* = 49)	Midline closure (*n* = 76)	Minimal excision (*n* = 270)	All (*n* = 667)
Employment status					
Employed	198 (73%)	35 (71%)	64 (84%)	201 (74%)	498 (75%)
House-partner or full-time parent/carer	4 (1%)	0 (0%)	0 (0%)	7 (3%)	11 (2%)
Volunteer or between jobs	12 (4%)	1 (2%)	1 (1%)	9 (3%)	23 (3%)
Student or trainee	36 (13%)	8 (16%)	9 (12%)	33 (12%)	86 (13%)
Retired	1 (0%)	0 (0%)	0 (0%)	3 (1%)	4 (1%)
Unemployed/not working	18 (7%)	4 (8%)	2 (3%)	13 (5%)	37 (6%)
Other	3 (1%)	0 (0%)	0 (0%)	3 (1%)	6 (1%)
Hair type					
0 Bald	5 (2%)	4 (8%)	2 (3%)	7 (3%)	18 (3%)
1a Straight (fine/thin)	57 (21%)	9 (18%)	20 (26%)	71 (26%)	157 (24%)
1b Straight (medium)	104 (38%)	12 (24%)	30 (39%)	87 (32%)	233 (35%)
1c Straight (coarse)	8 (3%)	2 (4%)	4 (5%)	14 (5%)	28 (4%)
2a Wavy (fine/thin)	24 (9%)	4 (8%)	4 (5%)	14 (5%)	46 (7%)
2b Wavy (medium)	31 (11%)	2 (4%)	10 (13%)	42 (16%)	85 (13%)
2c Wavy (coarse)	11 (4%)	8 (16%)	4 (5%)	17 (6%)	40 (6%)
3a Curly (loose)	20 (7%)	5 (10%)	1 (1%)	11 (4%)	37 (6%)
3b Curly (tight)	7 (3%)	2 (4%)	0 (0%)	3 (1%)	12 (2%)
4a Kinky (soft)	2 (1%)	0 (0%)	1 (1%)	2 (1%)	5 (1%)
4b Kinky (wiry)	1 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
4c Kinky (wiry)	2 (1%)	1 (2%)	0 (0%)	0 (0%)	3 (0%)
Hair cut frequency					
More than once every 4 weeks	80 (29%)	14 (29%)	20 (26%)	84 (31%)	198 (30%)
Once every 4–8 weeks	98 (36%)	20 (41%)	33 (43%)	109 (40%)	260 (39%)
Less than once every 8 weeks	94 (35%)	15 (31%)	23 (30%)	76 (28%)	208 (31%)

Data completion

Data completion rates for the cohort outcomes are presented in *Appendix 3*, *Table 25*. Return to normal activities, recurrence and wound healing were considered complete if a patient contributed those data at either clinic visit, 6-month follow-up or study close (or in the case of recurrence, it was apparent from AE reporting). Complication data were considered complete if the participant contributed data to at least one follow-up time point. Data for recurrence (94%), complication (96%) and return to normal activities (94%) were collected for most participants. The characteristics of participants that attended

TABLE 6 Participant disease characteristics

Characteristic	Asymmetric closure (*n* = 272)	Leave open (*n* = 49)	Midline closure (*n* = 76)	Minimal excision (*n* = 270)	All (*n* = 667)
Natal cleft depth (mm)					
N (%)	253 (93%)	40 (82%)	53 (70%)	236 (87%)	582 (87%)
Mean (SD)	19.1 (11.8)	23.6 (18.4)	20.8 (13.1)	19.9 (11.8)	19.9 (12.5)
Median (IQR)	16.0 (10.0–25.0)	18.5 (10.0– 33.5)	20.0 (10.0– 30.0)	20.0 (11.0– 25.0)	19.0 (10.0– 25.0)
Min, max	0.0, 80.0	5.0, 100.0	2.0, 63.0	0.0, 110.0	0.0, 110.0
Number of pits					
N (%)	267 (98%)	47 (96%)	62 (82%)	264 (98%)	640 (96%)
Mean (SD)	2.6 (2.0)	1.7 (1.2)	2.3 (2.6)	2.4 (1.7)	2.4 (1.9)
Median (IQR)	2.0 (1.0–3.0)	1.0 (1.0–2.0)	2.0 (1.0–3.0)	2.0 (1.0–3.0)	2.0 (1.0–3.0)
Min, max	0.0, 16.0	0.0, 6.0	0.0, 20.0	0.0, 17.0	0.0, 20.0
Length of pits (spread, mm)					
N (%)	176 (65%)	17 (35%)	34 (45%)	179 (66%)	406 (61%)
Mean (SD)	36.2 (37.1)	44.9 (36.9)	25.5 (20.3)	26.2 (21.9)	31.3 (30.4)
Median (IQR)	30.0 (10.0–50.0)	30.0 (21.0–70.0)	20.0 (10.0–35.0)	20.0 (10.0–40.0)	23.0 (10.0–41.0)
Min, max	0.0, 320.0	4.0, 150.0	0.0, 85.0	2.0, 140.0	0.0, 320.0
Pit density (pits per mm)					
N (%)	253 (93%)	42 (86%)	61 (80%)	256 (95%)	612 (92%)
Mean (SD)	0.1 (0.2)	0.0 (0.1)	0.2 (0.5)	0.1 (0.2)	0.1 (0.3)
Median (IQR)	0.1 (0.0–0.2)	0.0 (0.0–0.1)	0.0 (0.0–0.2)	0.1 (0.0–0.2)	0.1 (0.0–0.2)
Min, max	0.0, 2.0	0.0, 0.5	0.0, 3.0	0.0, 1.5	0.0, 3.0
Number of previous procedures					
0	129 (47%)	21 (43%)	48 (63%)	159 (59%)	357 (54%)
1	73 (27%)	14 (29%)	15 (20%)	72 (27%)	174 (26%)
2	37 (14%)	6 (12%)	9 (12%)	26 (10%)	78 (12%)
3 or more	33 (12%)	8 (16%)	4 (5%)	13 (5%)	58 (9%)
Previous procedures for PSD					
Elective procedure	57 (21%)	14 (29%)	9 (12%)	68 (25%)	148 (22%)
Acute drainage	101 (37%)	19 (39%)	20 (26%)	55 (20%)	195 (29%)
Emergency procedure	4 (1%)	0 (0%)	0 (0%)	1 (0%)	5 (1%)

continued

TABLE 6 Participant disease characteristics (*continued*)

Characteristic	Asymmetric closure (*n* = 272)	Leave open (*n* = 49)	Midline closure (*n* = 76)	Minimal excision (*n* = 270)	All (*n* = 667)
Months from last procedure to current procedure					
N (%)	56 (21%)	14 (29%)	9 (12%)	68 (25%)	147 (22%)
Mean (SD)	35.4 (43.3)	28.4 (29.1)	10.3 (8.5)	39.5 (61.3)	35.1 (50.6)
Median (IQR)	20.5 (8.0–47.0)	17.0 (9.0–41.0)	11.0 (2.0–16.0)	16.0 (8.0–41.0)	17.0 (8.0–41.0)
Min, max	0.0, 207.0	3.0, 100.0	0.0, 22.0	0.0, 329.0	0.0, 329.0
Wysocki classification					
Type 1	46 (17%)	7 (14%)	27 (36%)	102 (38%)	182 (27%)
Type 2	148 (54%)	19 (39%)	41 (54%)	116 (43%)	324 (49%)
Type 3	20 (7%)	8 (16%)	3 (4%)	19 (7%)	50 (7%)
Type 4	54 (20%)	13 (27%)	3 (4%)	31 (11%)	101 (15%)
None of the above	3 (1%)	1 (2%)	0 (0%)	0 (0%)	4 (1%)
Distribution of lateral openings					
No lateral openings	99 (36.4%)	20 (40.8%)	36 (47.4%)	140 (51.9%)	295 (44.2%)
Unilateral	150 (55.1%)	15 (30.6%)	17 (22.4%)	105 (38.9%)	287 (43.0%)
Bilateral	8 (2.9%)	4 (8.2%)	1 (1.3%)	7 (2.6%)	20 (3.0%)
Gluteal hair					
None	49 (18%)	7 (14%)	10 (13%)	31 (11%)	97 (15%)
Mild	137 (50%)	26 (53%)	32 (42%)	134 (50%)	329 (49%)
Dense	84 (31%)	13 (27%)	23 (30%)	99 (37%)	219 (33%)
Natal cleft skin					
Maceration	39 (14%)	8 (16%)	8 (11%)	23 (9%)	78 (12%)
Erosions	29 (11%)	4 (8%)	4 (5%)	14 (5%)	51 (8%)
Splits	15 (6%)	6 (12%)	8 (11%)	17 (6%)	46 (7%)
Wide pores	52 (19%)	16 (33%)	20 (26%)	53 (20%)	141 (21%)
First-degree relatives with history of PSD	51 (19%)	9 (18%)	16 (21%)	46 (17%)	122 (18%)
Relative with history of PSD					
Mother	10 (4%)	1 (2%)	4 (5%)	9 (3%)	24 (4%)
Father	19 (7%)	7 (14%)	6 (8%)	18 (7%)	50 (7%)
Sibling	14 (5%)	1 (2%)	3 (4%)	13 (5%)	31 (5%)
Child	2 (1%)	0 (0%)	0 (0%)	1 (0%)	3 (0%)
Multiple	6 (2%)	0 (0%)	3 (4%)	5 (2%)	14 (2%)

6-month follow-up were compared to those that did not attend 6-month follow-up (see *Appendix 3*, *Table 37*), and the distribution of characteristics was similar between the groups; there were marginally more participants that had lateral openings in the attenders (50%) compared to the non-attenders (44%). All analyses were conducted on available data.

Treatment decisions

The breakdown of treatments received is presented in *Table 7*. Recorded procedure details were categorised into four categories, which were further combined into major or minor procedure categories. Over half (60%) of the participants received a major treatment, most commonly asymmetric closure (41%). Of the participants that received minimal excision, the most common treatment options were glue (*n* = 106, 16% of the cohort) and pit picking (*n* = 60, 9% of the cohort).

Further treatment details are presented in *Appendix 3*, *Table 26*. Median length of surgery was 30 minutes, and 95% were performed as day cases. Procedures were typically performed by consultant surgeons (68%).

Figure 2 shows the treatment received by disease characteristics. Participants with recurrent disease (defined as reporting any previous procedure, including acute drainage) were more likely to be given asymmetric closure; participants that were not recurrent were more likely to receive minimal excision. Over half (56%) of Wysocki type 1 (only midline pit or sinuses) participants underwent minimal excision, whereas over half (53%) of Wysocki type 4 (disease after treatment with definitive intent) were given asymmetric closure. The extent of overlap is noteworthy; for all disease characteristic categories there were a number of participants that received each treatment type, suggesting there is variety in the types of procedures considered appropriate for patients with different disease characteristics. The distribution of the number of pits was similar across procedure types (see *Appendix 3*, *Figure 13*). The proportion of patients with a minor procedure varied substantially across the sites (see *Appendix 3*, *Figure 19*), although this may be due to differing case mix across centres.

TABLE 7 Treatment characteristics

N = 667					
Procedure type	***n* (%)**	**Procedure category**	***n* (%)**	**Procedure**	***n* (%)**
Major	397 (60%)	Asymmetric closure	272 (41%)	Bascom's cleft lift	86 (13%)
				Flap	22 (3%)
				Karydakis	164 (25%)
		Leave open	49 (7%)	Leave open	43 (6%)
				Leave open (marsupialisation)	6 (1%)
		Midline closure	76 (11%)	Midline closure	76 (11%)
Minor	270 (41%)	Minimal excision	270 (41%)	Bascom's I	39 (6%)
				EPSiT	44 (7%)
				Glue	106 (16%)
				Laser	11 (2%)
				Pit picking	60 (9%)
				Seton	10 (2%)

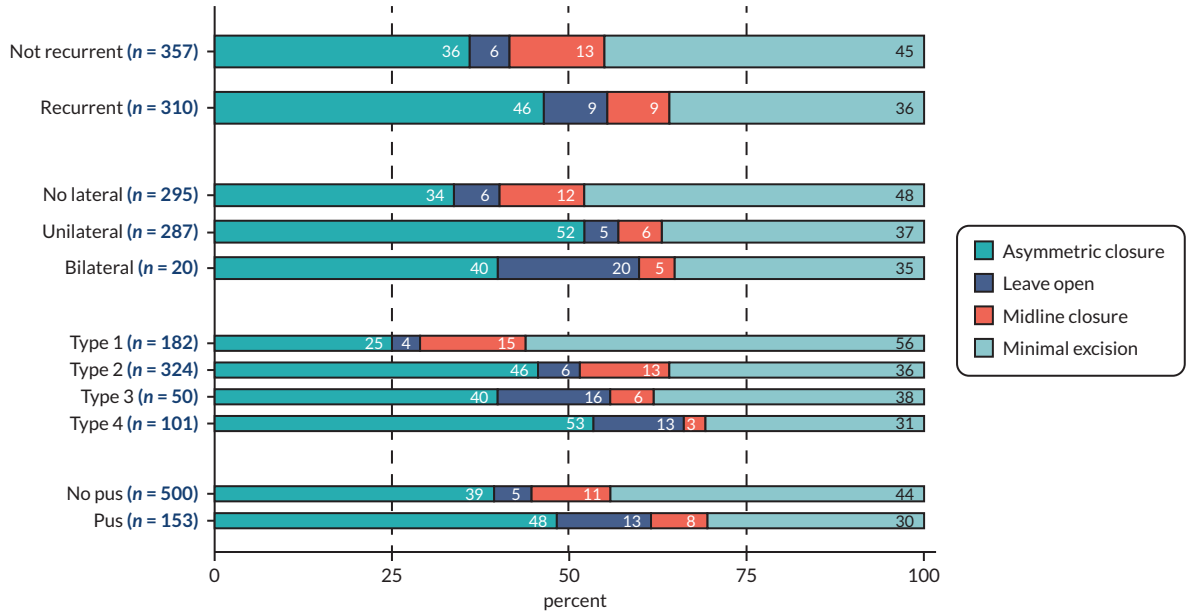


FIGURE 2 Treatment choice by disease characteristic. Note: 'recurrent' defined as any reported previous procedure; type 1: only midline pit or sinuses; type 2: any midline disease with secondary sinus/es or abscess scar/s; type 3: any midline or secondary disease extending below tip of coccyx; type 4, any disease after treatment with definitive intent.

Shared decision-making and decision regret

Participant ratings of their pre-op consultation were high (*Table 8*), with the median (IQR) of the CollaboRATE mean score response being 3 (3–4), where 3 represents 'a lot of effort was made'. The CollaboRATE top score was given in 36% of cases, reflecting that 'every effort' was made to help the patient understand their health issue, listen to the things that matter most and include what matters most to the patient in choosing what to do next. The decision regret (DR) scale, completed at month 6 follow-up, was low (median 8, IQR 0–20), and was broadly similar across the treatment categories.

The relationship between CollaboRATE mean score and DR is shown in *Appendix 3*, *Figure 14*. There is little clear evidence of a correlation between shared decision-making (SDM) and DR, with the majority of patients being in the top left corner of the graph (representing participants that were happy with their collaboration and had few regrets about their procedure).

Decision regret was low among patients [mean (SD) 14.5 (16.7)] and was broadly similar across the procedure types. The majority of patients reported being either satisfied or very satisfied (83%) with their procedure. Seven (21%) of the participants that received a leave-open procedure reported being either dissatisfied or very dissatisfied. The majority of patients returned to normal activity by the end of follow-up (n = 550, 88%) and 75% reported the wound as having healed during the study follow-up. Almost half (47%) of participants experienced a complication during follow-up.

Outcome summaries

The continuous outcome measures, recorded over time, are presented in *Appendix 3*, *Table 27*. Self-reported pain related to pilonidal sinus was at its highest on average on day 1 after procedure compared to baseline and reduced to its lowest at the 6-month visit. The highest average pain was reported by participants that received asymmetric closure. Patient-reported EQ-5D-5L health utility reduced from baseline to day 7 (overall means 0.80 and 0.69, respectively) but had recovered at both clinic and 6-month visits (overall means 0.83 and 0.89, respectively). Participants that received minimal excision reported the least change at day 7 and the highest health utility and QoL satisfaction at clinic visit and 6 months, although the scores were more similar among treatment groups at 6 months.

TABLE 8 Summary of pre-op consultation collaboration and post-op DR

	Asymmetric closure	Leave open	Midline closure	Minimal excision	All
	(*n* = 272)	(*n* = 49)	(*n* = 76)	(*n* = 270)	(*n* = 667)
CollaboRATE mean score[a]					
N (%)	270 (99%)	49 (100%)	75 (99%)	265 (98%)	659 (99%)
Median (IQR)	3 (3–4)	4 (3–4)	3 (3–4)	3 (3–4)	3 (3–4)
CollaboRATE top score given					
No	182 (67%)	28 (57%)	51 (67%)	155 (57%)	416 (62%)
Yes	88 (32%)	21 (43%)	24 (32%)	110 (41%)	243 (36%)
DR scale[b]					
N (%)	198 (73%)	35 (71%)	51 (67%)	173 (64%)	457 (69%)
Median (IQR)	10 (0–20)	8 (4–20)	8 (0–24)	8 (0–20)	8 (0–20)
Satisfaction with effect of treatment or care					
Very satisfied	113 (42%)	19 (39%)	21 (28%)	89 (33%)	242 (36%)
Satisfied	61 (22%)	9 (18%)	18 (24%)	54 (20%)	142 (21%)
Neither satisfied nor dissatisfied	15 (6%)	0 (0%)	6 (8%)	19 (7%)	40 (6%)
Dissatisfied	3 (1%)	6 (12%)	6 (8%)	9 (3%)	24 (4%)
Very dissatisfied	9 (3%)	1 (2%)	0 (0%)	6 (2%)	16 (2%)

a Recorded at baseline, score ranges from 0 to 4, higher scores represent more perceived effort made by professional in pre-op consultation.
b Recorded at 6-month follow-up, scored from 0 to 100, higher scores represent greater regret.

Other outcome measures are presented in *Appendix 3*, *Tables 28–30*. Repacking procedures were reported by 68 participants (12%) by day 7, while 87 participants (17%) reported repacking at clinic visit. At day 7, 226 (39%) patients reported a re-dressing procedure; re-dressing by day 7 was most common in asymmetric closure (49%) and midline closure (52%). Nearly half of participants experienced a complication during follow-up (*n* = 301, 45%), the most common of which were infection (26%) and discharge (18%). The numbers of complications were broadly similar across the three major surgery groups, and were lower for patients who received minimal excision, particularly for bleeding, dehiscence and infection.

Risk-adjusted treatment comparisons

The primary comparison between treatments was made between major procedures (asymmetric closure, leave open, midline closure, *n* = 396) and minor procedures (minimal excision, *n* = 270). No factors were found to be collinear and so all were included in the risk adjustment. Non-linearity of continuous features (BMI, natal cleft depth, pit spread and pit distance) was investigated and all features were deemed to be sufficiently modelled using linear terms. For the propensity score modelling, sufficient overlap in risk score was observed for all outcomes, and thus risk-adjusted analysis was deemed appropriate for major versus minor procedures.

Pain

The propensity score model identified sex, presence of pus and Wysocki classification as the most important features in treatment choice. Patients were more likely to undergo major procedure if they

were female, had pus, or were classified as Wysocki type 4 (disease after treatment with definitive intent) and least likely to have major procedure if type 1 (only midline pit or sinuses). These factors were used in the propensity-adjusted models for all outcome comparisons.

Pain on day 1, adjusted for factors predictive of treatment choice and outcome via augmented IPW, was higher for patients who received major procedures compared to minor procedures by 1.58 points (95% CI 1.14 to 2.01) (see *Appendix 3*, *Table 31*). This was very similar to the unadjusted difference (mean difference 1.62, 95% CI 1.23 to 2.02). The number of participants included in each analysis varied according to the factors included to adjust the models. A similar difference in pain at day 7 was observed (augmented IPW-adjusted mean difference 1.53, 95% CI 1.12 to 1.95). The difference in pain between treatment groups was similar regardless of the risk adjustment method. A post hoc sensitivity analysis that included baseline pain as a covariate yielded mean differences and 95% CIs that were within 0.1 points of these estimates.

Pain over time for major and minor procedures is shown in *Figure 3* (pain at clinic visit and month 6 were not prespecified as outcomes with formal comparisons); the raw difference in means was closer at clinic visit than at day 7, and there was no difference in mean pain reported at 6 months between the procedure types.

Complications

Just over half (54%) of participants receiving major procedures reported a complication, compared to 36% of participants who had a minor procedure. After augmented IPW risk adjustment, participants who received major procedures reported a 17.5% (95% CI 9.1 to 25.9%) higher absolute incidence of complications during follow-up (*Table 9*). The estimate of the difference was relatively consistent regardless of risk adjustment method.

Disease recurrence

Recurrence was less likely for major procedures compared to minor procedures (*Table 10*). Among participants who had a major procedure, 15% reported recurrence by 6 months after surgery, rising to 23% when the full follow-up period was included, compared to 27% and 34%, respectively, for participants who had minor procedures. The risk-adjusted absolute difference in 6-month recurrence was 10.1% (95% CI 2.1 to 18.1%) in favour of major procedure using augmented IPW and was of similar magnitude in other risk adjustments.

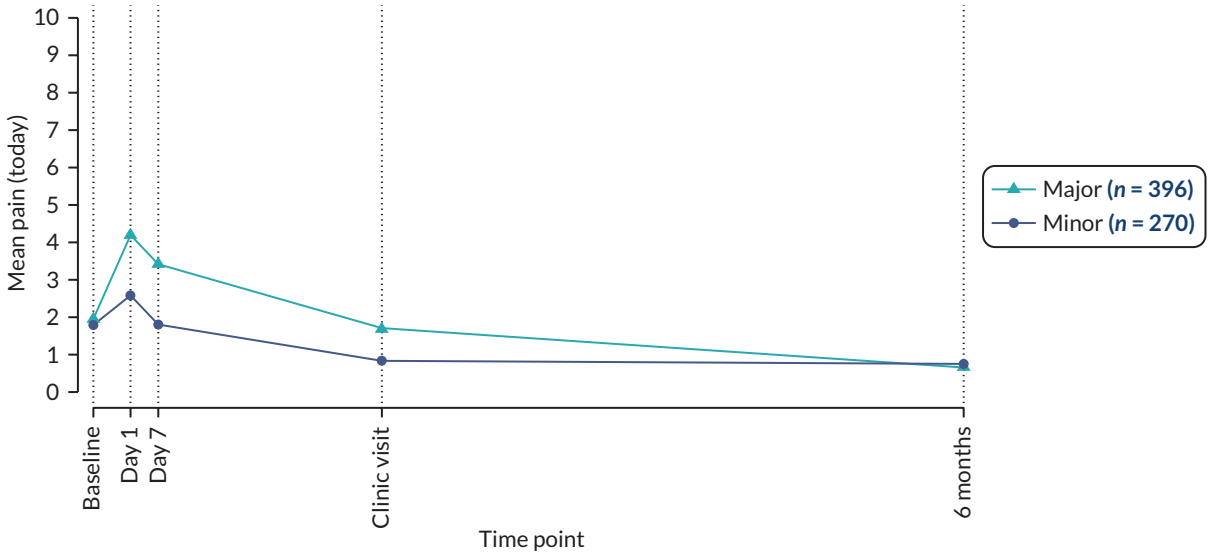


FIGURE 3 Pain with major and minor procedures, baseline to 6 months after surgery.

TABLE 9 Comparison of complications during follow-up between major and minor procedure

Complications	Major procedures	Minor procedures	*n*	Risk difference (95% CI)[a]
Raw difference	207/385 (54%)	94/258 (36%)	643	17.3% (9.6 to 25.0)
Risk-adjusted – Wysocki			638	16.7% (8.8 to 24.6)
Risk-adjusted – chosen model (BMI, Wysocki)			590	17.3% (9.1 to 25.6)
Risk-adjusted – full model			424	20.0% (10.4 to 29.6)
Propensity-adjusted – IPW			627	16.5% (8.1 to 24.8)
Propensity matching			627	15.9% (7.1 to 24.7)
Augmented IPW			579	17.5% (9.1 to 25.9)

a Reference group: minor procedures, risk-adjusted difference estimated using logistic regression with adjustment for covariates as listed, full model includes adjustment for sex, BMI, natal cleft depth, gluteal hair, smoking status, Wysocki classification, pit density, lateral distance, lateral distribution and presence of pus. Propensityadjusted and propensity matching adjusted for sex, Wysocki classification and presence of pus. Augmented IPW adjusted for sex, Wysocki classification, BMI and presence of pus.

Treatment failures (defined as the composite of recurred, not returned to normal activity, or not healed by the time of the last follow-up) were compared between treatment groups. The proportion of participants for whom treatment failed was more similar across treatment groups once healing and return to normal activities were introduced. In total, 47% of minor procedures failed at 6 months compared to 45% of major procedures (adjusted difference 2.3%, 95% CI −6.2 to +10.9%).

Return to normal activities

Time to return to normal activity was compared between major and minor procedures (*Table 11* and *Figure 4*). While nearly all participants had returned to normal activity by the end of follow-up, the time taken was far quicker among those undergoing minimal excision (median 7 days) than those who had major procedures (median 32 days). At 6 months, 12% of participants who had major procedures and 4% of participants who had minor procedures were yet to return to normal activity (see *Figure 4*). Several participants dropped out, providing either a censored time or no data at all; in the best-case scenario where these were assumed to have recovered, the proportion of participants who returned to normal activity at 6 months would be 96% for major procedures and 98% for minor procedures. Participants who received major procedures took on average 21 days longer to return to normal activity than those receiving minor procedures, and the difference was greater using risk adjustment models. The mean difference as estimated by augmented IPW was 25.9 days (95% CI 18.4 to 33.4 days), with regression adjustment approaches providing estimates closer to the unadjusted difference. Similar findings were present when comparing asymmetric closure procedures to the minimally invasive approaches (see *Appendix 3*, *Figure 19* and *Table 40*).

Wound healing

Participants having major procedures also took longer to heal than those who had a minor procedure. The median time to healing was 30 days among people undergoing minimal procedures, compared to 70 days among those undergoing a major procedure. However, as highlighted in *Figure 5*, around 25% of participants in both groups had wounds that had not healed. Some of the individuals lost to follow-up (LTFU) may have healed, but a best-case scenario where those censored prior to 6 months were assumed to have healed would still mean 10% of participants considered their wound unhealed at 6 months. Unadjusted and regression-based risk adjustments both estimated the difference in wound healing to be over a month greater following a major procedure, while propensity score methods estimated larger differences but with wider CIs (augmented IPW estimate 53.5 days, 95% CI 28.8 to 78.2 days; *Table 12*).

TABLE 10 Comparison of recurrence between major and minor procedures

	Recurrence				Recurrence (within 6 months)				Treatment failure			
Recurrence	Major procedures	Minor procedures	*n*	Risk difference (95% CI)[a]	Major procedures	Minor procedures	*n*	Risk difference (95% CI)[a]	Major procedure	Minor procedure	*n*	Risk difference (95% CI)[a]
Raw difference	86/373 (23%)	87/256 (34%)	629	−10.9% (−18.1 to −3.7%)	51/337 (15%)	61/229 (27%)	566	−11.5% (−18.4 to −4.6%)	169/373 (45%)	121/257 (47%)	630	−1.8% (−9.7 to 6.1%)
Risk-adjusted – Wysocki			624	−11.1% (−18.5 to −3.7%)			561	−11.3% (−18.4 to −4.2%)			625	−2.2% (−10.4 to 5.9%)
Risk-adjusted – chosen model (Wysocki, pit density)			577	−9.0% (−16.6 to −1.3%)			516	−9.4% (−16.7 to −2.0%)			578	−1.7% (−10.1 to 6.7%)
Risk-adjusted – full model			409	−8.4% (−17.7 to 0.8%)			366	−7.4% (−16.4 to 1.6%)			410	1.2% (−8.7 to 11.1%)
Propensity-adjusted – inverse weighting			613	−13.8% (−22.0 to −5.7%)			550	−12.9% (−20.7 to −5.1%)			614	−5.2% (−13.8 to 3.4%)
Propensity matching			613	−12.0% (−20.5 to −3.5%)			550	−12.5% (−21.0 to −4.1%)			614	−3.5% (−12.5 to 5.6%)
Augmented IPW			575	−10.1% (−18.1 to −2.1%)			514	−9.6% (−17.3 to −1.9%)			576	−2.3% (−10.9 to 6.2%)

a Reference group: minor procedures, risk-adjusted difference estimated using logistic regression with adjustment for covariates as listed, full model includes adjustment for sex, BMI, natal cleft depth, gluteal hair, smoking status, Wysocki classification, pit density, lateral distance, lateral distribution and presence of pus. Propensity-adjusted and propensity matching adjusted for sex, Wysocki classification and presence of pus. Augmented IPW adjusted for sex, Wysocki classification, BMI and presence of pus.

TABLE 11 Comparison of time to return to normal activity between major and minor procedures

Time to return to normal activity	*n*	Difference (days) (95% CI)[a]
Raw difference	607	21.0 (16.3 to 25.7)
Risk-adjusted – Wysocki	600	20.3 (15.6 to 24.9)
Risk-adjusted – chosen model (natal cleft depth, Wysocki, lateral distribution)	502	19.8 (14.7 to 24.9)
Risk-adjusted – full model	403	19.8 (14.0 to 25.6)
Propensity-adjusted – inverse weighting	589	35.6 (19.8 to 51.4)
IPW + regression	502	25.9 (18.4 to 33.4)

a Reference group: minor procedures risk-adjusted difference estimated using parametric accelerated survival time with adjustment for covariates as listed, full model includes adjustment for sex, BMI, natal cleft depth, gluteal hair, smoking status, Wysocki classification, pit density, lateral distance, lateral distribution and presence of pus. Propensity-adjusted for sex, Wysocki classification and presence of pus. IPW + regression adjusted for sex, Wysocki classification, natal cleft depth, lateral distribution and presence of pus.

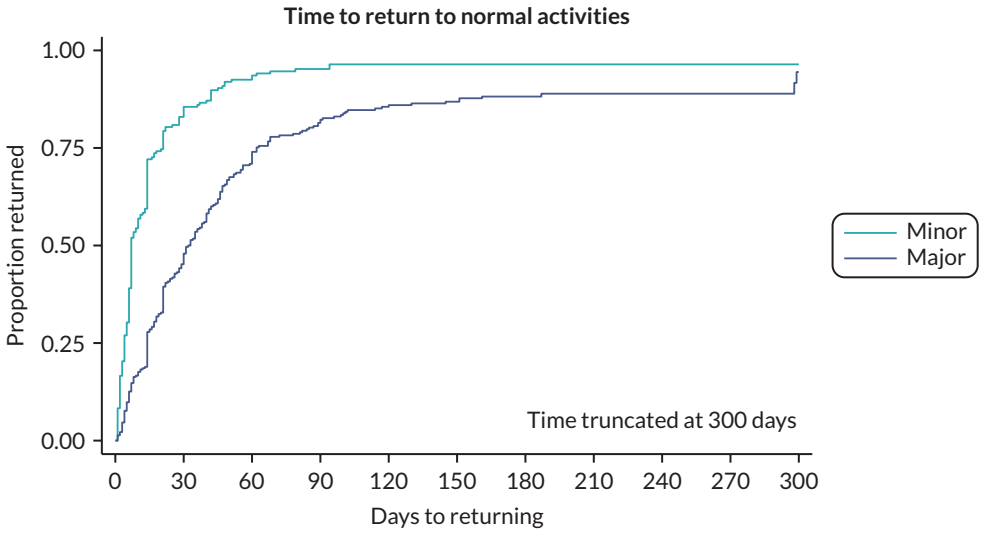


FIGURE 4 Time to return to normal activities by major or minor procedure type.

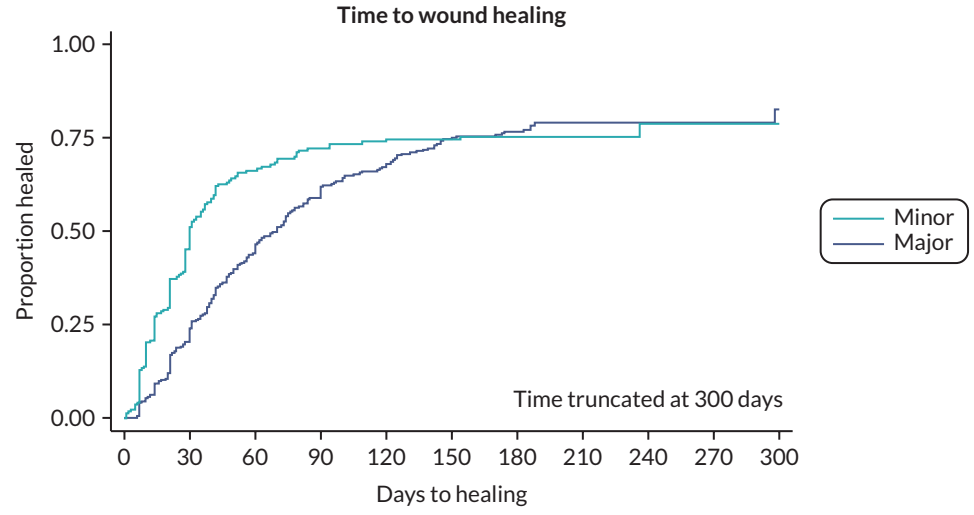


FIGURE 5 Time to wound healing by major or minor procedure type.

TABLE 12 Comparison of time to healing between major and minor procedures

Model	Major procedure		Minor procedure			
	n	Median (IQR) days	*n*	Median (IQR) days	*N*	Mean difference (95% CI)[a]
Raw difference	336	70 (31–52)	217	30 (14–54)	553	39.7 (27.0 to 52.4)
Risk-adjusted – Wysocki					546	36.7 (23.8 to 49.6)
Risk-adjusted – chosen model (Wysocki, BMI, smoking status, pus)					452	34.8 (19.9 to 49.6)
Risk-adjusted – full model					368	38.2 (22.3 to 54.1)
Propensity-adjusted –IPW					536	111.3 (–10.9 to 233.4)
Augmented IPW					452	53.5 (28.8 to 78.2)

a Reference group: minor procedures risk-adjusted difference estimated using parametric accelerated survival time with adjustment for covariates as listed, full model includes adjustment for sex, BMI, natal cleft depth, gluteal hair, smoking status, Wysocki classification, pit density, lateral distance, lateral distribution and presence of pus. Propensity-adjusted for sex, Wysocki classification and presence of pus. IPW + regression adjusted for sex, Wysocki classification and presence of pus.

Pairwise comparisons of asymmetric closure and minimal excision (removing participants that had the major procedures – leave open and midline closure) produced results in keeping with the comparison between major and minor procedures (see *Appendix 3*, *Tables 37–43*).

Surgeon variation

A post hoc analysis looked at recurrence rates among surgeons who operated on at least 10 participants. In total 13 surgeons undertook at least 10 procedures (range 10–55 procedures, median 14). In total 282/667 participants underwent procedure by one of these surgeons. While recurrence and treatment failure rates varied among the 13 surgeons, outcomes were more favourable among participants whose surgeons treated 10 or more cases (see *Appendix 3*, *Table 32*). Overall recurrence at 6 months was lower among participants treated by these surgeons (40%) compared with surgeons that treated fewer cases (60%), with similar differences seen for recurrence at any time (42 vs. 48%) and treatment failure (40 vs. 60%).

Adverse events

Adverse events and serious adverse events (SAEs) are presented in *Table 13*; 107 (16%) patients experienced at least one AE during follow-up, and the most common category of AE was wound infection (59%). Eleven participants experienced an SAE, nine of which were inpatient hospitalisation.

Discussion

The prospective cohort study was the main component of the PITSTOP study and consists of one of the largest data sets of real-world experience gathered on PSD to date. Although these are subject to the potential biases of non-randomised comparisons, the data suggest that there is more postoperative pain and failure of treatment after major excisional procedures compared to minimally invasive procedures, associated with an increased time to healing and longer time to return to normal activities. This is the case even after risk adjustment for patient demographics and severity of disease.

The demographics of this cohort are unsurprising: the disease tends to affect a young, predominantly male population who are overweight.[1] Interestingly, despite descriptions of patients as having coarse

TABLE 13 Adverse events and SAEs

	Asymmetric closure (*n* = 272)	Leave open (*n* = 49)	Midline closure (*n* = 76)	Minimal excision (*n* = 270)	All (*n* = 667)
Number (%) of participants who experienced ≥ 1 AE	64 (24%)	7 (14%)	19 (25%)	17 (6%)	107 (16%)
Number of all AEs (including repeated events)	94	8	24	19	145
Category					
Anaesthetic AE	2 (2%)	1 (13%)	0 (0%)	0 (0%)	3 (2%)
Bleeding/haematoma	7 (7%)	0 (0%)	4 (17%)	0 (0%)	11 (8%)
Dehiscence	22 (23%)	0 (0%)	6 (25%)	2 (11%)	30 (21%)
Discharge	2 (2%)	1 (13%)	3 (13%)	3 (16%)	9 (6%)
Medication AE	4 (4%)	0 (0%)	0 (0%)	0 (0%)	4 (3%)
Seroma	2 (2%)	0 (0%)	0 (0%)	0 (0%)	2 (1%)
Seton break	1 (1%)	0 (0%)	0 (0%)	0 (0%)	1 (1%)
Wound infection	54 (57%)	6 (75%)	11 (46%)	14 (74%)	85 (59%)
Number (%) of participants who experienced ≥ 1 SAE	6 (2%)	1 (2%)	2 (3%)	2 (1%)	11 (2%)
Number of all SAEs (including repeated events)	6	1	2	2	11
Seriousness					
Inpatient hospitalisation	4 (67%)	1 (100%)	2 (100%)	2 (100%)	9 (82%)
Considered medically significant by the investigator	1 (17%)	0 (0%)	0 (0%)	0 (0%)	1 (9%)
Category					
Bleeding/haematoma	1 (17%)	0 (0%)	0 (0%)	0 (0%)	1 (9%)
Seton break	1 (17%)	0 (0%)	0 (0%)	0 (0%)	1 (9%)
Wound infection	4 (67%)	1 (100%)	2 (100%)	2 (100%)	9 (82%)
Expected SAE	5 (83%)	1 (100%)	1 (50%)	2 (100%)	9 (82%)

hair,[36] we found the majority were assessed as having fine or medium hair. Around half of the group have had previous surgery for PSD, usually acute drainage of an abscess. Around one in five patients have had more than two procedures, with a significant minority having had three or more procedures. The disease varies in severity from simple midline disease to around half having some form of lateral extension. Around 10% have complex disease (bilateral disease or disease below the coccyx), making surgical intervention challenging and potentially limiting options for treatment. Recurrent disease – which may also, but not necessarily, be considered complex – was reported in 15%.

Twelve different types of surgical approach were utilised. This is more than in a previous survey on UK practice,[28] the increase being mainly due to the expanded repertoire of minimally invasive techniques including glue, laser and endoscopic treatment over the last 10 years. By far the commonest procedures were excision and asymmetric closure techniques (Karydakis and Bascom's II), which is consistent with previous data. While considered outdated due to the risk of failure and protracted recovery time,[15,37,38]

roughly one in six procedures involved excision-and-leave-open of the skin defect or primary closure in the midline. Reasons for the persistence of these procedures in UK practice have been discussed in *Chapter 2*. It is feasible that complex situations – for example, advanced bilateral disease, markedly infected wounds or other unusual variants of disease – meant that no other procedure was possible. However, it seems unlikely that these uncommon variants account for all such procedures carried out.

Despite the plethora of minimally invasive procedures currently practised in the UK, only 40% of patients had this approach. Given that this was among a group of surgeons interested in pilonidal disease, this approach could well be even less in the context of all UK surgeons. Minimally invasive procedures are certainly not suitable for all patients. For those with extensive disease or complex recurrence, minimally invasive approaches may not be effective. However, it would appear logical that, for most patients with non-recurrent disease confined to the midline or with simple lateral extensions, such a technique would have been feasible. One exception would be the patient with multiple pits within a small area or those with extensive underlying cavities where pit picking or Bascom's I would result in a large midline defect. Even considering these caveats, only around 60% of patients with disease confined to the midline and only 40% of patients with lateral extensions had a minimal procedure. Many more could have been treated less invasively.

Analysis of preoperative demographic and disease characteristics revealed some factors that made a major excisional technique more likely, including recurrent disease and the presence of pus. More surprising is the association of being female with an excisional procedure unrelated to disease extent and complexity.

The relatively low utilisation of minimally invasive procedures becomes very relevant when outcomes are considered. Minimally invasive techniques lead to less pain at all time points and especially in the first week after surgery. Patients undergoing minimal intervention reported pain (on a 0–10 scale) around 1.5–1.7 units lower than those undergoing major excision on day 1 and day 7. Complications were also more common with major excision, with 15–20% excess seen after major excision. Time to wound healing and return to normal activities were significantly shorter after minimally invasive procedures, allowing patients to return to work, study and socialisation much faster. In contrast, these data confirm that major excisional techniques are around 10% more likely to cure the disease. This draws into question whether patients prefer a higher chance of cure in preference to more pain, more complications and a more protracted recovery. Such trade-offs are explored in *Mixed-methods substudy* and *Discrete choice experiment*.

It is plausible that the differences seen when comparing these two intervention groups relate to the case mix. For example, more extensive disease would be more likely to require major excision, but patients would, regardless of intervention, be more likely to have a complicated and protracted recovery. We attempted to control for case mix by correcting for risk factors with statistical modelling. We prespecified demographic and disease characteristics that previous literature had identified as influencing outcome, and assessed sensitivity to this via alternative models that were developed after discussion with the core study clinical team. Regardless of the model used, the association with postoperative outcomes remained consistent. It is also worth noting the substantial overlap in procedure types even among ostensibly similar subgroups of disease, which suggests that the choice of procedure may be driven as much by patient choice and surgical familiarity as it is by the severity of the disease.

A noteworthy result was the time to return to normal activities and the time to healing. The literature is full of reports of spectacular efficacy for several procedures.[13,39,40] Yet regardless of intervention type, at least 10% and possibly up to 25% of patients had not healed by 6 months, and up to 12% of the major excision group had not returned to normal activities. This suggests that, in the real world as compared to specialist units, these procedures may not necessarily be as effective as the literature would suggest, and there may be a significant postoperative burden for patients of which many will not have been made aware during informed consent.

It is likely that patients have different interpretations of the terms 'recurrence' and 'wound healing' and may have considered these as interchangeable when telling us about their disease during follow-up. In the true sense, recurrence means disease that has healed after surgery but which then arises again, at the same site or at a different site to the original disease. This should be distinguished from disease that fails to heal at all after surgery or indeed when surgery achieves excision of the pits, but the patient is left with an unhealed wound, often in the midline. True recurrence is probably much less common and requires a protracted length of time to detect accurately.[41] Failure to heal or persistence of disease is probably easier to define, but there remains an issue as to the time point when the intervention should be considered to have healed and not recurred. Consensus from the European Society of Coloproctology working party on pilonidal guidelines has proposed that 6 months after surgery would seem an appropriate time point (Asha Senapati, September 2023, personal communication). If a 25% failure rate was observed, regardless of technique, this would certainly be inferior to most of the reported literature.[13,39,40] This has repercussions when it comes to SDM before surgery.

The difference between this study's observed failure rate and that reported may relate to the skill and experience of individual surgeons. It may be that experts in pilonidal surgery can achieve healing rates equivalent to those in the literature.[25,42–45] Our study partly triangulated this theory: surgeons responsible for more cases (defined here as treating 10 or more study participants) had better outcomes than surgeons who treated fewer participants, but recurrence was still higher than in previous literature. This could imply some surgeons are more skilled than others, although numbers were small, and the healing rates were not controlled for factors such as technique, case volume and disease severity. An important limitation is that this analysis was not risk-adjusted, and disease characteristics (and hence outcomes) may differ between experienced and less experienced surgeons. Nevertheless, it is unlikely that more experienced surgeons would be systematically allocated easier-to-treat patients, and so disease severity is unlikely to be the reason for these differences. This may justify referral to specialist centres for those patients with particularly complex disease.

The recognised issues with major excision in terms of protracted wound healing and the reported high failure rate with excision and midline closure mean that grouping such procedures along with asymmetric closure and flap procedures[15] may have skewed the results in favour of minimal intervention. Fortunately, the number of patients in the asymmetric closure group (Karydakis and Bascom cleft closure) was sufficient to allow us to carry out a more granular analysis. Despite excluding these other procedures, the outcomes were similar for cleft closure techniques compared to minimally invasive techniques. Patients who had an asymmetric closure had more early pain, a higher complication rate and longer time to healing and return to normal activity than those treated with minimally invasive procedures. Failure was 5–8% less likely but the rate was still much higher than in most of the reported literature.[13]

When one considers this potential high rate of failure, it is somewhat surprising that the DR was so low. Patients were mainly satisfied with their decision for surgery and in addition felt that the options and outcomes were explained well to them. The CollaboRATE scores were high, suggesting the SDM process was good. These findings somewhat contradict other data suggesting that many are not offered a range of operations, in particular minimally invasive procedures. Data should perhaps be interpreted with caution as they may reflect a social desirability bias.[46]

There are limitations to both these data and their interpretation. An obvious limitation is the incompleteness of the data: 1 in 10 patients had missing day 1 data, and 6-month data were only available in around three-quarters of patients. This is despite rigorous governance processes and dedicated research nurses assiduously following up the patients. The study period did fall during the COVID pandemic, and this will have influenced the ability to follow up rigorously in some cases. Otherwise, the incompleteness of the data is probably a reflection of the demographic of the group, which tends to consist of young active working people, predominantly male. Such a demographic may be less likely to respond to follow-up calls.[47,48] Interestingly, this is not the case in virtually all published

series with follow-up of greater than 12 months.[49–58] Most of these trials report complete data collection and all have more than 90% attending for follow-up. It is unclear how these other studies succeeded in such an astonishing rate of attendance. One study in particular reported follow-up data involving a clinic visit 5 years after surgery.[51] The authors countenance a follow-up of at least this long if recurrence rates are to be considered accurate.[41] While they may be correct, such a follow-up period is not practical and is likely to produce levels of incomplete data far exceeding 25% if carried out in the UK. While we did manage to obtain data for healing, complications and return to normal activities in at least 83% of participants, in some cases this was at only around the 6-week follow-up.

Other limitations relate to the multitude of interventions that were included in the cohort. While analysable data were available for 667 participants, there were 12 different interventions carried out. There were just not enough data to compare individual procedures in any meaningful way. If we had aimed to have 100 patients for each procedure rather than just the front-running procedures, a cohort of over 6000 patients would have been required. It was felt justified to group similar techniques by invasiveness. Minimally invasive procedures have a distinct commonality in that they do not involve major skin/subcutaneous tissue excision and instead focus on destruction/removal of the pit and underlying cavity. Major procedures involve excision of the disease and surrounding skin with or without closure of the wound created. There are clearly subtleties for the way both minimal and major interventions are carried out and advocates will proclaim the benefits and harms of each. Indeed, the literature is full of such comparisons.[40] As such, a broad categorisation may be criticised, particularly for the major excisional group where excision-and-leave-open or closure in the midline techniques is considered by many to be obsolete compared with an asymmetric closure technique.[15,37,38] Nevertheless, even when a more granular analysis of asymmetric closure versus minimally invasive techniques was carried out, similar differences in terms of pain, recovery, failure to heal and complications were seen.

The final limitation relates to the risk adjustment. Demographic and disease characteristics that may influence outcome are currently unproven. We decided on such parameters by consensus among the core study clinical team. We accommodated the uncertainty by including multiple permutations of the risk adjustment model. Overall, these permutations led to similar results, strengthening the justification for the model and the overall conclusions. Nevertheless, all risk adjustment is predicated on being able to fully quantify ‘risk’, which is both a strong and untestable assumption.

Conclusions

The real-world experience of surgery for pilonidal disease is not as good as the literature would suggest. Many patients have a protracted recovery regardless of intervention, and failure is common. The utilisation of minimally invasive techniques could be increased and would reduce the burden of postoperative recovery substantially while accepting a small reduction in cure rate. Patient QoL and health economics studies, including investigation on the cost to society of longer absences from work and education, may inform shared decisions on first-line treatment of PSD in the future.

Chapter 4 Mixed-methods substudy

Methods

Design

A multiple-case design was employed to compare more than one data type between and within more than one person.[59] The case study was nested in the observational cohort, with two embedded units of analysis at baseline and 6 months: qualitative longitudinal semistructured interviews and quantitative cohort data.

Participants

Sampling was purposive and sought to recruit PITSTOP cohort participants with symptomatic PSD referred for elective surgical treatment. This sampling method aimed for maximum variation based on the following: Wysocki classification (an indicator of disease severity), surgical procedure (excision and closure techniques) and NHS Foundation Trust. Initial contact was made by telephone. All potential participants were e-mailed the participant information sheet and provided verbal consent prior to participation in the interview.

Data collection

At baseline, participants completed a patient-reported experience measure (PREM) of SDM. SDM is scored 0 to 9 (0 indicating poor SDM; 9 indicating good SDM).[60] At 6 months, participants completed a PREM of DR. Using a 5-point scale, healthcare DR is scored 0 (low DR) to 100 (high DR).[61] In addition, the following outcome measures were collected: pain, length of time to healing and post-surgery complications. Semistructured telephone interviews were conducted between June 2019 and September 2020. A minimum of 20 interviews was considered adequate to understand common perceptions and experiences of treatment choices, thereby achieving data saturation.[62,63] A topic guide was designed using the coping in deliberation (CODE) and Sekhon's Acceptability framework.[64,65] Baseline interviews adapted key 'choice' (e.g. 'did you let the surgeon choose your treatment?') and 'options' (e.g. 'did the surgeon talk you through the risks and benefits?') questions from the CODE framework.[64] At 6 months, the interview guide asked CODE questions related to decision 'consolidation' (e.g. 'was this the right decision?'). Throughout, probing questions covered dimensions of Sekhon's acceptability framework,[65] as well as intervention attributes, to inform the discrete choice experiment (DCE). The interviews were recorded using an encrypted digital recorder and transcribed verbatim.

Patient and public involvement

One patient representative participated in a pilot interview. The aim of the pilot interview was to assess the apprehension of the topic guide questions and review interview cohesion. No amendments were made to the topic guide.

Data analysis

Transcripts were analysed using the National Centre for Social Research 'Framework' analysis approach. This approach was chosen because it allows for coding of a priori and de novo themes.[66] After familiarising ourselves with the data set, we independently coded a sample of transcripts using the CODE and Acceptability frameworks[64,65] on NVivo (QSR International, Warrington, UK) version 11 before conferring. During the analysis and interpretation, integration of qualitative and quantitative data

occurred to understand: (1) how disease characteristics and surgeon preferences interacted with patient values in treatment choices; and (2) how participants appraised treatments given particular outcomes. We used joint display tables to look for convergences and divergences between cohort data (disease features/treatment choices/outcomes) with experiences, views and values.[67,68] Patient experts were invited to provide feedback on a lay summary of triangulated results to assess acceptability.

Results

Sample

An initial expression of interest was made by 266 cohort participants. The final sample comprised 20 patients (median age 28; range 20–64) from 13 NHS Foundation Trusts in the UK (*Table 14*). Only 13 participants completed baseline (median 16 minutes, range 6–47 minutes) and follow-up interviews (median 18 minutes, range 11–37 minutes).

Coping in deliberation framework: health threat

Newly diagnosed participants were unfamiliar with, and expressed confusion about, the cause and prognosis of PSD. They detailed their experience of soreness, inflammation, discharge and odour which disrupted employment, physical activity and relationships. Participants discussed the impact this had on psychosocial well-being.

> *I was told initially, 'Oh that could be it, and then it might go away' … but once you get it once, that's it: it's coming back … If I was a bit more aware of that I would have probably started to look into the surgeries quicker.*
>
> *18: no previous pilonidal disease*

Some participants were reluctant to address their condition, deciding to tolerate distress and delay treatment. Often, an exacerbation of symptoms would cause participants to present to emergency services for treatment.

> *I said to [my girlfriend], 'Look, I can't really see it properly. Is it still getting bigger?' And, she said, 'Oh bloody hell … get in the car.' So, we went straight to [hospital].*
>
> *03: one previous episode of PSD*

Other participants discussed barriers to secondary care treatment referrals due to their GP not taking their condition seriously.

> *… just gave me some antibiotics … it just kept getting more painful and worse … I went back three times … then she put me on sort of the path to go back to surgery, but she didn't send me [as] an urgent patient … so I had to wait for maybe like five months.*
>
> *14: no previous PS*

Another saw their GP numerous times over 25 years and was repeatedly dissuaded from surgery.

> *he basically sort of said to me that it's a very precarious operation … that the success rate wasn't very high … that it was something that if I could live with.*
>
> *15: no previous pilonidal disease*

Coping in deliberation framework: choice

Once referred to secondary care, 9/20 participants were offered a choice of treatment. Absence of choice was rarely viewed negatively. Due to their own limited knowledge of the condition, some participants viewed healthcare professionals as best placed to make decisions about their care, especially where emergency surgery was concerned. If offered a choice, participant preference was

TABLE 14 Participant characteristics: baseline, surgery and 6-month outcomes

			Baseline			Surgery	6 month		
ID	Male/female	Age	Prior operations	Severity	SDM	Excision/closure	DR[a]	Recurrence[b]	Follow-up interview?
			Complete data set						
1	Male	31	0	1	2.6	LE/MC	10	No	Y
3	Male	29	1	1	4	PP/LO	10	Yes	Y
5	Male	28	0	1	3	LE/LC	5	No	Y
6	Female	20	3	1	4	LE/LO	30	No	Y
8	Male	27	2	4	3	LE/LO	20	No	Y
9	Male	25	0	1	4	LE/LC	5	No	Y
10	Male	64	0	3	3.6	Se only	5	No	Y
11	Male	27	1	4	3	Cu/LO	5	No	Y
14	Male	23	0	2	2	EPSiT/LO	15	No	Y
16	Female	26	2	4	4	LE/LC(K)	0	No	Y
17	Male	22	2	3	4	LE/LC(K)	0	No	Y
18	Female	27	0	1	4	PP/LO	50	No	Y
19	Male	40	0	4	2.66	LE/MC,M,LC	0	No	Y
			Incomplete data set						
2	Female	31	2	4	4	LE/LC(K)	LTFU	LTFU	Refused
7	Female	33	1	2	4	Cu,PP/FG	0	No	LTFU
12	Male	49	0	1	4	Cu/FG	LTFU	LTFU	LTFU
13	Male	28	0	1	4	Cu/FG	LTFU	LTFU	LTFU
15	Male	44	0	2	2.33	LE/Se,Fl	LTFU	LTFU	Refused

continued

TABLE 14 Participant characteristics: baseline, surgery and 6-month outcomes (*continued*)

			Baseline			Surgery	6 month		
ID	Male/female	Age	Prior operations	Severity	SDM	Excision/closure	DR[a]	Recurrence[b]	Follow-up interview?
20	Male	34	0	4	4	LE/Fl	LTFU	Yes	LTFU
21	Female	25	1	2	2	LE/MC	0	No	LTFU

Excision types: Cu, curettage; LE, local excision; PP, pit picking. Closure types: Fl, flap; K, Karydakis; LC, lateral closure; LO, leave open; M, marsupialisation; MC, midline closure; Se, seton.

a Score of 0–100 where 0 = no regret and 100 = maximum regret

b Definition of recurrence as per Table 4: further procedure for PSD (excluding expected procedures of repacking and replacement/removal of dressings).

- Patient-reported recurrence.
- AE or other complications which were consistent with an unresolved PSD.

Severity (Wysocki classification): high scores more severe. SDM: self-reported quality of SDM using the CollaboRATE instrument – high score (highest 5) denotes best SDM. DR: highest score 100 = high level of regret.

based on one or more of the following factors: previous experience of surgery (*n* = 3); surgeon's guidance (*n* = 3); invasiveness of the treatment (*n* = 3); or anticipated recovery time (*n* = 2). One participant rejected their consultant's treatment advice to undergo a 'leave open' procedure due to the employment opportunity costs – they would require additional time off work (*Table 15*). Other participants utilised social networks (significant others, friends and relatives) or the internet to support decision-making, with some acquiring a sense of control from independently researching the condition and treatment options.

Coping in deliberation framework: key outcomes at the time of decision-making

Not all participants specified a single important outcome when undergoing surgery. The following were discussed: avoiding recurrence (*n* = 8), return to normal activities (*n* = 6) and/or the elimination of symptoms (*n* = 7). Six participants were not aware of procedural risks; others expressed awareness of risks presented by anaesthesia (*n* = 2), infection or bleeding (*n* = 4), the wound not healing (*n* = 5) and recurrence (*n* = 8).

Coping in deliberation framework: consolidation

After surgery, most participants were anxious of aggravating the wound and/or delaying wound healing. Many implemented adaptations, including physical (altering seating positions) and behavioural (reducing exercise duration and changing the type of exercise) changes, which negatively impacted their well-being.

> *It has made me reticent to engage in some activities … exercise and things like that … through the pain and discomfort, and also the chance of sort of popping the cyst…*
>
> *01: no previous pilonidal disease*

Many participants required daily or weekly primary care wound management support via GP appointments or district nurse visits. Due to its location, most experienced difficulties managing the wound independently. They used mirrors for physical inspection or were dependent on others to help examine and manage the wound (including cleaning, dressing and packing). Some participants cited the reliance on others as a loss of independence, while others acknowledged how emotional support alleviated distress.

> *I think the worst part of it is that you always have to rely on someone else to do, like, a dressing for you … you can't drive cos you can't sit down … you basically you can't do anything.*
>
> *06: sinus excised and left open*

Acceptability framework: key outcomes at 6-month follow-up

Six months post surgery, some participants reported hoping surgery would address: pain (*n* = 3); recurrence (*n* = 5); wound healing (*n* = 1); the smell (*n* = 1); the inconvenience (*n* = 1); impaired ability to perform ADL (*n* = 1). In five cases (01, 05, 11, 14, 17; see *Appendix 3*, *Table 33*), these priorities had changed since baseline. During the recovery period, some participants accepted the recurrent nature of the condition. Others managed their own treatment expectations by considering any symptomatic improvement as an indicator of effectiveness.

> *I've still got some kind of stuff going down on there that is just a recurring thing … if I've had four operations, it probably won't get rid of [it].*
>
> *06*

Some participants specified a single important outcome 6 months after surgery. The following were discussed: the wound healing in the expected time (*n* = 3); avoidance of recurrence (*n* = 4); return to ADL (*n* = 1). Three participants who had undergone PSD surgery for the first time found their treatment to be effective. When asked, they did not feel they would have done anything differently (see *Appendix 3*, *Table 33*).

TABLE 15 Decision-making – cases ordered by self-reported quality of SDM

Participant information			Decision-making	
ID	Number of prior procedures	Key outcome	CollaboRATE score	Sample quote (coding)
21	1	Recurrence	2	'[The surgeon] said they'd cut like a flap out, get everything out and sort of stitch it back up … that was the only option … that or managing with medication … I was like, yeah, do what you have to do.' (Presentation of choice)
14	0	ADL	2	'I only really got a say in it this time … cos it was a new surgery coming though …. They offered me to do the other one if I wanted' (Presentation of choice)
15	0	Recovery time	2.33	'If you're asking me how it felt like, it felt like I didn't have a choice.' (Presentation of choice) 'at first I, [the consultant] sort of said, oh you might be back in a … couple of weeks and then when my friend said oh, 12 weeks for this open wound to heal, I thought … I can't take that long off work. I can't afford it.' (Preference construction)
1	0	Recurrence	2.6	'[The surgeon] said you either don't have the surgery and hope that it maybe sorts itself out … I took the decision that the chance of the surgery resolving the matter was worth the risk that it might still reoccur … with no other sort of major health issues that seemed like an easy enough choice' (Presentation of choice)
19	0	Recurrence	2.66	'No, [the surgeon] did not give me any option. He just said, just, just he only mentioned the surgery. As I say, I wasn't given any other options' (Presentation of choice)
5	0	Pain	3	'[The doctor] said that they'll operate and that was pretty much it … just leave it, or you could have the operation and I thought, well, best to try and get it sorted before it keeps getting infected, and gets worse' (Presentation and interpretation of options)
8	2	Smell	3	'[The surgeon] give me options of what I wanted and I just wanted one, like obviously cos I had it packed last time, it healed better that way, so I asked for it that way.' (Preference construction)
11	1	Pain and ADL	3	'[The consultant] explained to me that you know, we could try medication first and then if that doesn't work, we could try surgery … it was a scraping out I think … that was something [the consultant] recommended' (Presentation of choice)
10	0	(Not specified)	3.6	'I didn't decide any treatment. The treatment was decided for me by the consultant … I'm not medically qualified, you know … I'm told what the problem is and how it can be rectified. We go along with that.' (Decision)

TABLE 15 Decision-making – cases ordered by self-reported quality of SDM (*continued*)

Participant information			Decision-making	
ID	Number of prior procedures	Key outcome	CollaboRATE score	Sample quote (coding)
2	2	ADL	4	'No, there was only one procedure left' (Presentation and interpretation of options)
3	1	Recurrence and pain	4	'It's not me fighting this battle … I'm just a battlefield. You guys are fighting it … by the time I got to A&E, they may have given me options, I can't remember …. I'm quite happy to accept that I don't know what I'm talking about, so even if I'm given options I will say to the man giving me options, what would you do' (Presentation and interpretation of options)
6	3	Reducing anxiety of knocking the sinus (reduce symptoms)	4	'I saw my consultant and he said … depending on the MRI, I'll give you a few options … one is that we do the same but obviously different in theatre and then the, the other option is to have it, like, lasered, removed' (Health threat)
7	1	ADL	4	'They gave me two options but obviously because I have to get a mastectomy in September … I wouldn't have been healed in time … my immune's so low as well, we said that the glue one'd be more beneficial for me.' (Preference construction)
9	0	Recurrence	4	'The wording was this is the best thing to go for … either don't have the surgery and hope that it maybe it sorts itself out … or sort of cutting it out … I wasn't really exploring every single option available' (Preference construction)
12	0	Recurrence	4	'[The surgeon] gave me two or three different options that we could take i.e. stitching, gluing, leaving alone etc. and I thought the gluing one sounded the best and of course she agreed that she would like to do the gluing one anyway but she wanted me to make the choice really.' (Presentation of choice)
13	0	Closing the wound	4	'It was either an option of having it packed, which the doctor said can take up to a month for it to be fully healed … obviously being self-employed, I need to be back in work … I just plumped for the one that sounded like the one that I thought would work the best and I think it was a newer procedure' (Presentation of choice/presentation and interpretation of options)
16	2	ADL	4	'I could leave it and just live with it, which obviously for me wasn't an option! … my other option was to get a cosmetic surgeon in … So, I was just kind of worried that I would always kind of be left with some sort of wound' (Health threat)

continued

TABLE 15 Decision-making – cases ordered by self-reported quality of SDM (*continued*)

Participant information			Decision-making	
ID	Number of prior procedures	Key outcome	CollaboRATE score	Sample quote (coding)
17	2	Recurrence	4	'[The nurse] just told me I'd be having emergency surgery … someone looked at me that following morning and decided that I definitely had to have the incision and drainage. They didn't go through the details of why that was, I'll be honest … I didn't know the ins and outs of what I had, and I didn't know if there was any other options available.' (Presentation of choice)
18	0	Solve the problem	4	'I wasn't given the choice as such of which ones to do but when [the consultant] said that this is what she recommends, I completely took that on board from somebody with her kind of experience and knowledge of it' (Presentation of choice)
20	0	ADL and recurrence	4	'[The GP] said you've got two options, I either give you some antibiotics and pain relief now or I recommend you go to hospital … I wanted to maintain as much quality of life as possible while listening to the consultant's guidance' (Preference construction)

A&E, accident and emergency; ADL, activities of daily living.
Key outcome: primary desired outcome for each participant expressed at baseline interview. CollaboRATE score: mean CollaboRATE score regarding SDM of treatment recorded at baseline; high score = more SDM, low score = less SDM (0–9); table reports CollaboRATE scores low–high. Sample quotes taken from interview at baseline.

Discussion

This substudy explored how patients make, and sometimes regret, treatment decisions for PSD. Patients are often reluctant to address the condition due to inadequate knowledge and embarrassment. GPs are often hesitant to refer patients to secondary care services as surgical approaches are perceived to be poorly evidenced. Once referred, patients may not be involved in the choice of surgical treatment. Typically, they are unconcerned with and uninformed about the burden post procedure (wound management, practical support, and risk of infection and/or recurrence). Therefore, they may experience unanticipated difficulties when trying to cope with these matters. Patients undergoing their first surgery are often overly optimistic about the effectiveness of treatment. In contrast, those with recurrent disease may experience higher treatment regret and psychosocial distress due to poorly informed decisions. Irrespective of prior experience of PSD, making decisions about surgical treatment is challenging.[64] Digesting new and complex treatment information can be difficult. Insufficiently informed, patients are unable to articulate what they would have done differently, with many citing a change in outcome priorities after 6 months. Patients with a previous history of PSD who had undergone an excise-and-leave-open procedure – which is associated with high levels of pain, wound management and extended healing times – demonstrated the highest levels of DR.

This substudy is limited as postsurgical wound healing can take over 6 months and recurrence of PSD may take place over many years. Therefore, the attitudes of participants may be affected due to the short follow-up period. The COVID-19 pandemic significantly impacted our ability to follow up participants.[69,70] Although remote data collection was possible, engaging patients in research was difficult as many were experiencing COVID-related barriers.[71] In PSD studies, attrition rates are poorly reported and this is thought to be due to the young, mobile, mainly male population.[15,72–74] High attrition rates may also result from an unwillingness to disclose negative experiences,[75] or a loss of interest in research after the wound has healed and they have returned to work.[76]

The mixed-methods approach identified divergences and inconsistencies between different data sets – particularly, how patients reflect on their treatment decisions after surgery. Even if they are involved in SDM, if patients are not fully informed about treatment and post-surgery pathways, their expectations may not be met.[77] Levels of self-reported DR in this study are in line with the 1 in 7 rate reported across 73 surgical studies, in which regret was mainly associated with type of surgery, health outcomes and absence of SDM.[78] Another systematic review has flagged decisional conflict and anxiety as predictive of DR.[79] Surgeons[80,81] and patients[82] may have reasons for avoiding SDM, and our findings complicate the common assumption that SDM leads to increased decisional satisfaction.[83] Systematic reviews in other contexts suggest that unmet information needs are common and distressing.[84–87] There are growing concerns that self-report measures of SDM may not capture the quality of the interaction or the multistaged nature of the process.[88,89] PREMs may be compromised by social desirability or acquiescence bias,[90–93] and open-ended questions may reveal significant problems from patients who report high levels of satisfaction on survey instruments.[94,95] Triangulation of research methods is useful to identify such problems.[96,97]

Clinical teams should ensure patients are provided with sufficient information about available surgical procedures. Expectations associated with postprocedural aftercare and the uncertainties surrounding clinical outcomes should be managed. Surgeons may not actively engage in wound care discussions as this is perceived as the responsibility of primary care services.[98] However, insufficient information may impact a patient's ability to self-manage post surgery. Therefore, patients should receive tailored verbal and written information regarding treatment outcomes during relevant consultation appointments.[99] Discussing expectations with patients may provide an opportunity to address false optimism.[100,101] In other settings, patient expectations predict treatment satisfaction and functional outcomes post surgery.[102]

Awareness-raising among GPs and surgeons is needed to avoid delays in treatment where PSD is poorly recognised. Where pilonidal surgery is seen as unglamorous,[103] or surgeons only specialise in one

technique,[104] patients with recurrent disease should be referred rapidly onward to relevant specialists. Both SDM and the consent process may be compromised if patients are poorly informed about their condition, available treatments and the probability of various outcomes.[80] This is challenging when there are many available treatments supported by variable evidence.[14] There are around 20 systematic reviews and meta-analyses on surgical techniques alone, and around 15 more on medical, wound care and other topics. An overview of these reviews should be an urgent research priority to adequately inform SDM and the development of decision support tools. Until then, the review by Stauffer and colleagues remains one of the most comprehensive overviews focusing on time to recurrence with different surgical techniques.[13] Finally, discharge planning should begin at pre-assessment visits, involving the patient, day surgery nurses and district nurses.[105–107] Postoperative wound care is enhanced by continuity of care from a limited number of community-based health professionals.[108]

In conclusion, ensuring people with PSD are provided with sufficient information regarding wound care management and risks of recurrence associated with different surgical approaches may facilitate decision-making and minimise treatment regret. An overview of systematic reviews is needed to inform decision support tools. Healthcare professionals should communicate the uncertainties about treatment effects in addition to the time frames, adaptations and psychosocial impact associated with recovery.

Chapter 5 Discrete choice experiment

Methods

Design and theoretical/conceptual framework

Discrete choice experiments are an attribute-based measure of benefit; they assume that healthcare interventions, services or polices can be described by their attributes. In a DCE, respondents are required to make trade-off decisions regarding the quantity or quality of a good or service. The resulting choices are analysed to estimate the overall utility (value) and willingness to trade between services. The use of DCEs to identify patient preferences in health and healthcare has increased. Where no clear treatment decision exists, accurate quantification of patient preferences for risks and benefit is crucial.

Development of survey

The survey contained four separate sections. These were:

1. patient characteristics and disease history
2. a treatment ranking exercise
3. DCE-specific tasks
4. survey feedback questions.

The ranking exercise and DCE tasks were developed by conducting qualitative interviews with 20 patients. The patients were asked to identify the key attributes and levels that they considered when choosing a treatment (see *Chapter 4* for further information). A thematic analysis of the interview data was conducted and identified a list of factors that patients considered important when making treatment decisions. PPI members and clinicians reviewed an initial list of themes and selected attributes considered the most important for inclusion in the DCE and ranking task.

Nine attributes were included. These were:

1. type of excision and closure
2. type of anaesthetic
3. length of hospital stay
4. wound care
5. pain medication requirement
6. infection risk
7. healing time
8. risk of recurrence
9. scarring.

Currently, there are 18 theoretically possible surgical options for PSD.[14,110] Clinicians developed a treatment classification of five treatment groups. This was considered important due to the potential cognitive burden of ranking 18 treatment categories. The treatment categories, related descriptions and attributes were informed by the literature, clinical input and PPI piloting (see *Appendix 3*, *Box 1*).

Two attributes were included in the DCE: risk of infection/persistence and recovery time. This selection was based on two reasons. Firstly, these two attributes were assessed as most important by patients and

clinicians. Secondly, DCE attributes must be independent to avoid presenting implausible combinations of attribute and level profiles. The levels for the DCE attributes were selected based on plausible values published in the literature and additional input from clinicians.

The choice tasks were constructed based on an orthogonal design using a design catalogue.[111] The survey contained 16 hypothetical DCE tasks. Participants were asked to choose between two combinations of outcomes with varying levels (see *Appendix 3*, *Box 1*). Forced unlabelled choices were presented – 'treatment A' or 'treatment B'. An 'opt-out' alternative was not provided for the purposes of realism. A dominant task was included – where one treatment option is logistically superior – to test participant understanding of the task (see *Appendix 3*, *Box 2*).

Patient and public involvement

Patient and public involvement representatives were heavily involved in the design and implementation of the DCE. They assisted with the following tasks:

- Reviewed the initial list of themes to finalise the attributes.
- Reviewed the acceptability of the treatment category descriptions.
- Prior to the dissemination of the survey, three PPI representatives reviewed the survey to assess comprehensibility, interpretation and complexity of tasks.

Sampling

Discrete choice experiment sample sizes can vary from 100 to 1000 plus.[112] An Orme[113] rule of thumb formula was adopted to estimate the minimum sample size. Using this formula – {N = 500 × [4 (maximum number of levels)]/[2 (number of alternatives) × 16 (number of tasks)]} – the estimated sample size was 63.

Recruitment

All participants aged 16 years and above with symptomatic PSD, referred for elective surgery and participating in the PITSTOP cohort study, were invited to take part in the survey. Interested participants were e-mailed a Qualtrics® (Qualtrics, Provo, UT, USA) link which included a participant information sheet, consent form and the questionnaire. Participants with symptomatic PSD and not participating in the PITSTOP cohort could also take part in the DCE by accessing a digital QR code advertised on a study leaflet. The study leaflet was displayed in NHS Foundation Trust colorectal outpatient clinics and disseminated via the PITSTOP Twitter website.

Data analysis

Descriptive statistics were calculated for patient characteristics, disease history and survey feedback variables. DCE responses were modelled using conditional logistic regression where the dependent variable was the preferred treatment choice, and the independent variables were risk of infection/persistence and recovery time. Linearity of the attributes was assessed before deciding to treat risk as a linear variable (*Figure 6*). Regression coefficients were used to estimate the relative importance of attributes. Maximum acceptable risk (MAR) is the rate at which patients are willing to sacrifice a benefit of one attribute in exchange for an improvement in another. MARs were calculated by dividing the ratio of recovery time coefficients by the infection/persistence coefficient and 95% CIs calculated using the delta method.[114] Latent class models were used to analyse individual heterogeneity and to identify subsets of patients with varying preferences. The optimal number of classes was selected using the Bayesian information criterion (BIC) and consistent Akaike information criterion (CAIC) and model parsimony. Data were analysed using Stata® 17 (StataCorp LP, College Station, TX, USA).

Results

Participants

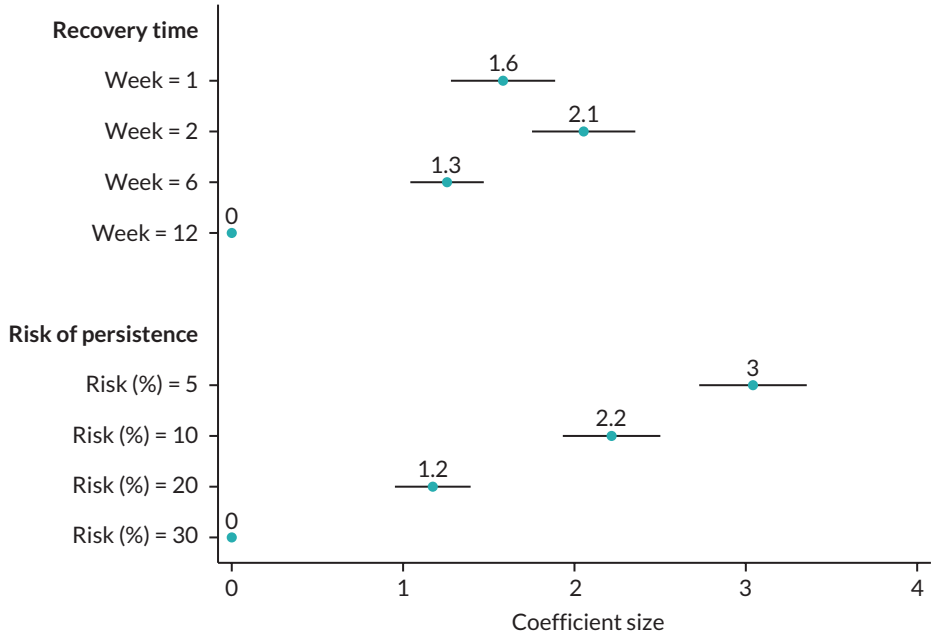
One hundred and eleven participants were included in the DCE survey. The completion rate was 74% (423 unique visitors entered the survey, 150 participants consented to take part and 3 participants declined). *Table 16* reports the characteristics of the 111 included participants. Of these, 75 (68%) respondents were male and 73 (66%) were between the ages of 17 and 29 years; 89 respondents (80%) were employed and 97 (87%) were white. Except for six respondents, the rest of the sample reported having had at least one surgery for PSD. The respondents had various types of surgeries, including excision of the skin and closure of the wound with stitches (26%); excision of the sinuses only and leave the wound open to heal (23%); excision of skin and leave the wound open (23%); excision of the skin and closure of the wound with a skin flap and stitches (9%); and excision of the sinuses and closure of the wound with glue (19%).

Patient preferences

Appendix 3, *Table 34* presents the regression modelling results of the DCE. Patients preferred treatments with lower risk of infection/persistence and this attribute was modelled linearly (see model 2 in *Appendix 3*, *Table 34*). Risk of infection/persistence was the most important attribute when patients were choosing a treatment, with an attribute importance score of 70%. Patients also preferred treatments with shorter recovery time; for example, compared to a treatment that takes 12 weeks to recover, a treatment with a 1-, 2- or 6-week recovery period was preferred. However, their preferences were not linear, so this attribute was modelled as a categorical variable. Treatments with shorter recovery time had an attribute importance score of 30%.

Maximum acceptable risk

When choosing a treatment, patients were willing to make trade-offs between risk of infection/persistence and recovery time (*Table 17*). These trade-offs were measured using MAR, which is the maximum risk of infection/persistence participants are willing to accept to have a treatment with faster recovery times. The highest-risk patients were willing to accept was a 17.08 risk of infection/persistence in return for a treatment with 2-week recovery period compared to a treatment with 12 weeks recovery



Note: Positive coefficients show attribute levels that increase the likelihood of patients choosing a treatment

FIGURE 6 Regression model 1 results reproduced in a diagram.

TABLE 16 Participants' sociodemographic and clinical characteristics

N = 111	No.	%
Sex		
Male	75	68
Female	36	32
Age (years)		
17–29	73	66
30–39	24	22
40–49	8	7
50–59	4	4
60–69	2	2
Median age, years (range)	28 (17–65)	
Which of the following best describes your main activity?		
Employed	89	80
Retired	1	1
Homemaker	3	3
Carer	1	1
Student	17	15
Ethnicity		
White	97	87
Black	2	2
Asian	2	2
Mixed	10	9
Education		
Primary	4	4
GCSE	16	14
A-level	48	43
Degree	42	38
Prefer not to say	1	1
Previous pilonidal sinus surgeries (including both emergency drainage and previous 'definitive' elective repair)		
0	6	6
1	67	64
2	17	16
3	6	6
4	5	5
5	2	2
6	1	1
10	1	1

TABLE 16 Participants' sociodemographic and clinical characteristics (*continued*)

N = 111	No.	%
Type of previous pilonidal sinus surgeries		
Excision of skin and leave the wound open (e.g. leave open/marsupialisation)	29	23
Excision of the skin and closure of the wound with stitches (e.g. midline closure, Bascom's cleft closure, Karydakis)	33	26
Excision of the skin and closure of the wound with a skin flap and stitches (e.g. rhomboid, Limberg, Dufourmental)	11	9
Excision of the sinuses and closure of the wound with glue	24	19
Excision of the sinuses only and leave the wound open to heal (e.g. pit picking, EPSiT, laser)	29	23

TABLE 17 Estimated MAR

Treatment benefit (for selected level changes)	MAR of infection/ persistence	95% CI calculated using the delta method
Recovery time reduction from 12 weeks to 2 weeks	17.08[a]	14.83 to 19.33
Recovery time reduction from 12 weeks to 6 weeks	10.49	8.76 to 12.22
Recovery time reduction from 6 weeks to 2 weeks	6.59	4.88 to 8.30

a MAR of 17.08 = patients were willing to accept a 17.08-percentage-point increase in risk of infection/persistence to have a treatment with a faster recovery period (2 weeks compared to 12 weeks).

period. Patients were willing to accept a 10.49 increase in risk of infection/persistence to have a treatment with 6-week recovery period compared to 12-week recovery period. Patients were willing to accept a 6.59 increase in risk of infection/persistence to have a treatment with a faster recovery period (2 weeks compared to 6 weeks).

Preference heterogeneity

Differences in preferences between patients were explored using the latent class modelling approach. This identified two groups of respondents with different preferences (*Table 18*). The first subgroup of respondents (class 1) were risk-averse and so they were only willing to accept a small risk (1.51–2.15) in exchange for a treatment with faster recovery time (*Table 19*). The second subgroup of respondents (class 2) showed stronger preferences for treatments with shorter recovery time: they were willing to accept higher risks of infection/persistence (22.35–34.67) to receive treatments with quicker recovery time. Of all the demographic variables used to predict whether a respondent belonged to class 1 or 2, only age of the respondents was statistically significant (see *Table 18*). The results show that respondents in the 17–29 age group were more likely to belong to class 1 and respondents above the age of 30 were more likely to belong to class 2.

Ranking of treatments

Patients ranked the treatments they preferred in order of importance, and *Figure 7* shows the results of this ranking task. The best preferred treatment was complex flap (e.g. Limberg, Dufourmental)

TABLE 18 Preference heterogeneity – latent class model results

	Class1	Class2
Week = 12 (reference level)		
Week = 1	0.159 (0.405)	2.192*** (0.209)
Week = 2	0.929 (0.488)	2.861*** (0.214)
Week = 6	0.652 (0.363)	1.844*** (0.171)
Risk (%)	−0.432*** (0.049)	−0.083*** (0.008)
Class membership		
Sex = female	−0.280 (0.469)	Reference
Age = 17–29 years	1.365** (0.486)	
Activity = employed	0.375 (0.564)	
Number of non-emergency surgeries patients had for pilonidal sinus	−0.008 (0.163)	
Education = degree	0.450 (0.456)	
Constant	−1.485* (0.692)	
Observations	3328	
Log-likelihood	−569.55	
BIC	1206.51	
CAIC	1220.51	

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Note
Standard errors in parentheses.

TABLE 19 Estimated MAR based on subgroups

Treatment benefit (for selected level changes)	Subgroup 1 (class 1) MAR of infection/persistence (95% CI)	Subgroup 2 (class 2) MAR of infection/persistence
Recovery time reduction from 12 weeks to 2 weeks	2.15 (−0.03 to 4.34)	34.67 (28.24 to 41.10)
Recovery time reduction from 12 weeks to 6 weeks	1.51 (−0.11 to 3.13)	22.35 (17.35 to 27.35)

procedures (27%), followed by excision of the sinuses only (22%), glue (19%), excision of the skin and closure of the wound with stitches (18%) and lastly leave open (14%). The least preferred treatment was leave open (35%), followed by glue (23%), complex flap procedures (18%), excision of the sinuses only (17%) and excision of the skin and closure of the wound with stitches (7%).

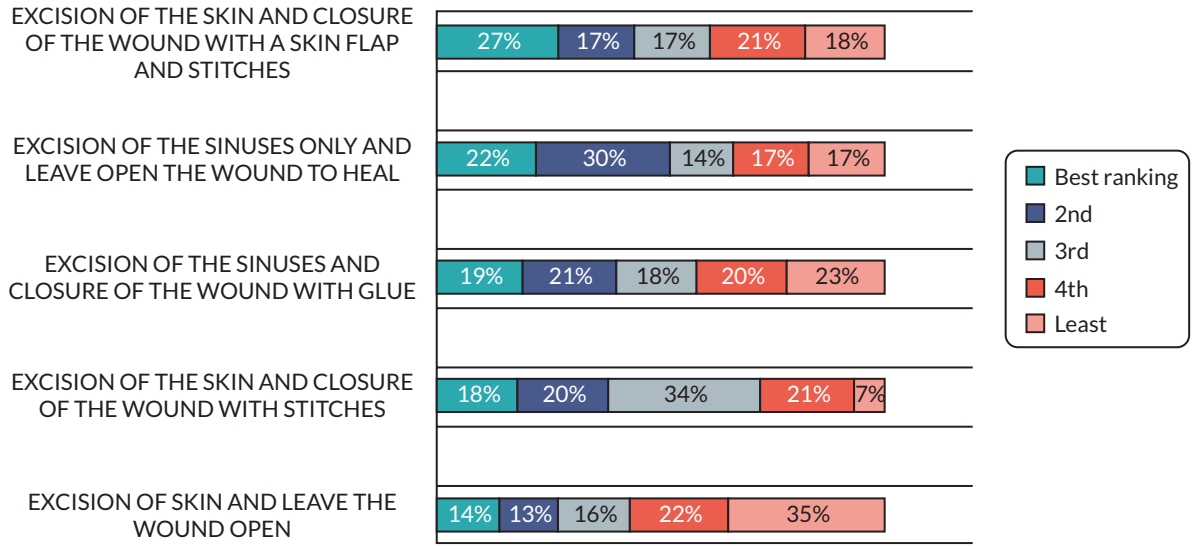


FIGURE 7 Ranking of treatments.

Participants' understanding of and engagement with the survey

Table 20 reports the results of the questions used to test the internal validity of the DCE.

Most of the respondents said that they understood the DCE tasks (91%) and ranking task (86%). Ninety-three (84%) of the respondents correctly answered the DCE question with a logically correct answer (dominance test). None of the respondents were always choosing the same side (left or right) profiles of the DCE tasks. Fewer than 25% said that the DCE task was confusing and that they needed more information.

Discussion

This substudy assessed patient treatment preferences for PSD. When choosing a surgical treatment, patients prioritised risk of infection/persistence relative to recovery time. However, patients were willing to compromise and accept treatments associated with varying degrees of greater risk of infection/persistence in favour of treatments that were associated with quicker recovery. The results provide insight into the type of treatments patients accept and value. In the overall group, patients were willing to accept up to a 17-percentage-point increase in risk of infection/persistence for treatments with a shorter recovery period. This suggests that some patients are willing to accept less invasive treatments with shorter recovery periods and greater risk of infection/persistence (e.g. glue and/or pit picking) over more invasive treatments with longer recovery periods but reduced risk of infection/persistence (e.g. an excise-and-leave-open procedure). In the ranking task, similar results were found: open surgery was ranked as the least favoured treatment option. This is understandable given the impact of prolonged wound care management on psychosocial well-being.[115]

The results demonstrated preference heterogeneity, which indicates the importance of providing treatments tailored to subgroups of patients with distinct preferences. Patients aged 30 years and over were willing to accept up to a 35-percentage-point increase in risk of infection/persistence for treatments with a shorter recovery period. This suggests that patients within this age bracket would be likely to reject an excise-and-leave-open procedure in favour of a treatment associated with a faster recovery period. In our sample, this age demographic reported that they were either homemakers, retired or had caring responsibilities. Therefore, it is plausible that their personal circumstances may have influenced their preference for a treatment associated with a shorter recovery time. In contrast, patients aged 17–29 years were more risk averse and were only willing to accept a

two-percentage-point increase in risk of infection/persistence for treatments with a shorter recovery period. The differences in preference heterogeneity further support the tenet that patients should be involved in making decisions about their surgical care to avoid treatment DR.[30] Literature exploring SDM in PSD is growing.[30,74,116] Such studies have highlighted the importance of providing patients with sufficient information for each available surgical procedure (including wound care management) to aid treatment decision-making.[30,74,116]

TABLE 20 Survey comprehension and internal validity

Internal validity (dominance questions) passed or failed?	*N*	%
Failed	18	16
Passed	93	84
Always choose the same side (e.g. left profile) in all the DCE questions?		
Yes	0	0
No	111	100
I found the ranking question made sense – please tell us how strongly you agree		
Strongly disagree	3	3
Disagree	2	2
Uncertain	10	9
Agree	60	54
Strongly agree	36	32
I understood the questions about making choices between different treatment options		
Strongly disagree	1	1
Disagree	1	1
Uncertain	8	7
Agree	59	53
Strongly agree	42	38
When choosing options, I needed more information than was provided		
Strongly disagree	16	14
Disagree	43	39
Uncertain	24	22
Agree	21	19
Strongly agree	7	6
I found making a choice between different treatments confusing		
Strongly disagree	21	19
Disagree	56	50
Uncertain	15	14
Agree	15	14
Strongly agree	4	4
Median time to complete survey, minutes (range)	12 (4–5388)	

This substudy is the first to conduct a DCE to assess PSD patient treatment preferences. A robust methodology was employed; qualitative interviews were conducted to inform the development of the survey; experienced clinicians and PPI representatives contributed throughout the design and implementation phases. Relevant interval validity checks were made and identified task comprehension, supporting confidence in the results.

However, this substudy has limitations. The sample size was sufficient to estimate overall modelled preferences, but a larger sample size would have allowed greater confidence in the analyses classifying members to different latent classes. During recruitment, several methods were employed to increase response rate. As PSD typically affects a young, working-age population, engagement barriers can be incurred.[117] Currently, there is no consensus on how PSD should be classified.[110] The clinicians developed five treatment categories for the ranking task based on their own clinical experience and the literature. A different team of clinicians may have made different classification decisions. In addition, the DCE included two key attributes to avoid presenting implausible combinations of attribute and level profiles. However, in a real-life context, patients may consider other factors not included in the DCE to support treatment decision-making. Finally, 16% of participants failed the internal validity test, demonstrating that the risk information presented was not well understood. In future, including a numeracy test or presenting risk information using pictorial icon ranges may support internal validity.

Chapter 6 Consensus exercise

Methods

This process was conducted over three phases:

Phase 1: item generation using a 'So what, now what' focus group
Phase 2: online modified Delphi over three rounds of iterative voting
Phase 3: consensus meeting to confirm prioritisation of items.

Stakeholders

Two stakeholder groups were defined: patients with previous experience of PSD and clinicians with an interest in PSD. Clinicians included those with certificates of completion of training in general surgery and nurse specialists with wound care practice. Participants were recruited via e-mail, national organisations and social media. Snowball sampling was also used for clinician recruitment. Participants were invited to participate in one, two or all three phases of the consensus exercise.

Patient and public involvement

Fifteen patient representatives with relevant experience were recruited to the patient stakeholder group following substudy conception, and contributed to the delivery and analysis. Of these, 6/15 PPI representatives attended the initial workshop and supported the generation of the longlist of recommendations; 4/15 PPI representatives attended the virtual consensus meeting and highlighted the importance of ensuring the final set of recommendation statements were conceivable to a patient audience.

Generation of longlist

In accordance with Delphi methodology, the study consisted of three phases. In phase 1, an online workshop was conducted. This was based on Rolfe's critical reflection model, 'What? So What? Now What?'.[118] In the 'What?' phase, data or information is presented. In this case, researchers presented findings from the cohort study, mixed-methods study and survey-based work. In the 'So What?' phase, participants are encouraged to reflect and discuss the information. Participants were asked to consider how the presented data reflected their experiences, and how this matched wider experiences. In the 'Now What?' phase, participants discuss how the information should be used to influence the next stage. Participants were asked to frame their ideas as statements related to policy or research ideas. In the workshop, the following data were presented: preliminary PITSTOP cohort data, a systematic review of classification systems,[110] a mapping reviewing of PSD,[119] the PITSTOP DCE survey[109] and the PITSTOP mixed-methods study.[30] Participants were asked to consider two questions: 'How can we use this data to improve and/or inform clinical practice?' and 'What are the key research questions generated by this data?'. A longlist of potential practice and research recommendation statements was generated. The steering committee assessed the readability of these statements. Prior to attending the workshop, all participants received an information sheet and completed an online electronic consent form.

e-Delphi consensus

In phase 2, a three-round e-Delphi consensus was conducted. The Delphi surveys were delivered using Qualtrics. In round 1, all participants were e-mailed a participant information sheet and a link to the survey. Upon accessing the survey, participants were asked to complete an online consent form. The following information was captured: age, gender, demographics, ethnicity, e-mail address and stakeholder respondent group (patient, surgeon and specialist nurse). The longlist of recommendation statements was presented in a random order, and each statement was supplemented with a written summary to aid understanding. At the end of the survey, respondents were encouraged to propose any additional statements. Additional items were reviewed at the end of round 1.

In rounds 2 and 3, the remaining longlisted items were presented in random order. Ratings of items were reviewed after the close of each round. Respondents received an e-mail copy of results which included their vote and how that compared to each stakeholder group's votes.

During each round, participants voted on the importance of each recommendation using a 9-point Likert scale (one being 'not important' and nine being 'very important'). Recommendations were shortlisted if the following was satisfied: (1) > 70% participants within both stakeholder groups rate the recommendation as 7–9; or (2) 90% participants within a single stakeholder group rate the recommendation as 7–9. Recommendations that reached consensus after three rounds were considered at the consensus meeting. All items had to be rated to complete the surveys. Only those who completed a survey round were eligible to participate in the subsequent round. At the end of round 3, all participants were asked if they were interested in participating in the virtual consensus meeting.

Virtual consensus meeting

In phase 3, an online consensus meeting was held to finalise the set of recommendations. Invitations were issued to interested participants, with efforts made to encourage participation from members of the public/patients. A target of > 15 participants with at least three patient representatives was felt to be reasonable for this prioritisation exercise as it reflected proportions recruited to round 1 of the consensus. The meeting was held using the Google Meet™ videoconferencing platform (Google Inc., Mountain View, CA, USA). Electronic consent was taken prior to the meeting. Participants were presented with a total of 34 statements: 15 policy and 19 research. After the presentation, participants were instructed to complete two separate Qualtrics surveys. A constant sum question type was used to calculate the sum of scores for each statement. In the first survey, participants were asked to distribute 100 points between the 15 policy statements dependent on priority. This could be all points to a single item, an even division, or however the respondent felt appropriate, as long as 100 points were distributed. The total points allocated to each item were then calculated, and the five highest-scoring items were considered as top priorities. In the second survey, participants distributed 100 points between the 19 policy statements dependent on priority. The same approach to summing points was conducted.

Results

Longlisting of potential outcomes

Following the initial 'So What, Now What' workshop, 33 items were generated for the longlist by clinicians and patients. The flow of items is presented in *Figure 8*.

Delphi round 1

Consent forms were completed by 57 potential candidates from both stakeholder groups, and 57 completed round 1. Characteristics of respondents' participation among stakeholder groups are presented in *Table 21*. This included 15 patients, 40 surgeons and 1 nurse specialist. A total of 33 items were considered for level of priority; 15 items reached a priori consensus for inclusion for the final consensus meeting. Respondents proposed a further 12 items for review. Outcomes voted on in each round, along with the proportion of each panel rating them 7–9, are presented in *Appendix 3*, *Tables 35* and *36*.

Delphi round 2

In round 2, 53 participants completed the survey. This included 14 patients, 38 surgeons and 1 nurse specialist, 95% of those who completed round 1. Respondents were sent a summary of voting patterns from the first round and were asked to vote on 30 items: 18 for reconsideration and an additional 12 statements proposed. Of these, 18 items met the a priori criteria for inclusion in the consensus meeting.

Delphi round 3

In round 3, 51 participants completed the survey. This included 14 patients, 36 surgeons and 1 nurse specialist. This reflected 96% of those who completed round 2 and 91% who completed round 1. One further item was carried to the consensus meeting.

Consensus meeting

The consensus meeting was attended by three patient representatives and 15 clinicians. One clinician withdrew during the meeting due to work commitments. The top five policy statements and research recommendations are presented in *Table 22*.

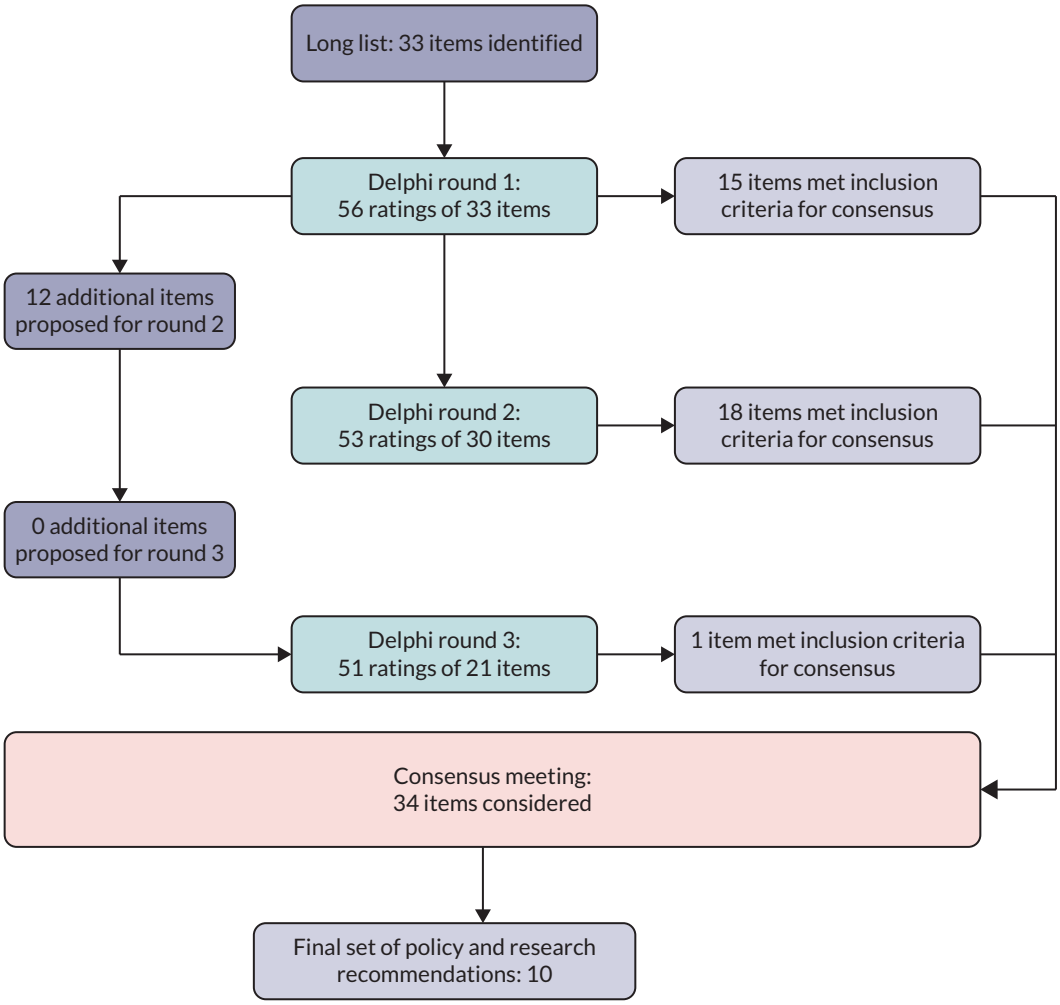


FIGURE 8 Flow of items through the e-Delphi.

TABLE 21 Summary of Delphi consensus participant characteristics

		Round 1	Round 2	Round 3
Participant type	Patient	15	14	14
	Surgeon	40	38	36
	Nurse specialist	1	1	1
Retention rate	–	–	95%	91%

TABLE 22 Final set of statements presented in order of sum

Statement number	Policy statement	Sum
1	Any treatment of pilonidal disease should aim to be less disruptive than the disease itself.	270
2	Minimally invasive techniques should be considered as the first-line intervention, as these are associated with low operative morbidity and comparable recurrence and healing rates to more extensive interventions.	174
3	Surgeons should have access to opportunities to learn new techniques for the treatment of pilonidal sinus disease.	140
4	A classification tool for pilonidal sinus should help to inform treatment options.	140
5	Delayed return to work is an important outcome following treatment.	134
Statement number	**Research statement**	**Sum**
1	A future randomised trial (RCT) should include two broad groups of interventions – major (i.e. asymmetric closure, leave open and midline closure) versus minor (i.e. minimal excision).	189
2	A core outcome set for pilonidal disease might help us understand what outcomes are important to clinicians and patients following treatment of pilonidal disease. It may also improve future evaluations of treatments.	179
3	Future research should compare major procedures (e.g. flaps) against minor procedures (e.g. pit picking, glue) stratified by disease severity.	148
4	There is a need for a patient-reported outcome to be used in future pilonidal sinus research.	119
5	Future research should aim to define an algorithm or decision tree to aid surgeon decision-making.	100

Discussion

Overview

Research in surgery has been much maligned over the years,[120] and pilonidal disease is no exception.[14] This is reflected in relatively weak or vague guidance to support practice. The Delphi we conducted has identified high-priority practice and research topics to guide the further development of the field. Consensus statements on practice topics strongly reflect the findings from previous sections. The top practice recommendation reflects the need to avoid harm related to interventions. This is supported by findings from the mixed-methods study which showed high levels of regret associated with poor wound healing. The cohort study supports the top two policy recommendations as it demonstrates the ongoing morbidity from poor wound healing after major procedures. In contrast, the third policy recommendation does not relate to clinical outcomes, but to the need to train surgeons in new techniques as highlighted by the clinician survey. This is particularly important if the top two priorities are to be achieved, as upskilling of surgeons will be required to facilitate techniques beyond 'excise and leave open'. The need for a classification tool with clinical reference is clear. In other areas of colorectal surgery, classification tools and systems facilitate decision-making around treatment pathways.[121,122] This inconsistency in PSD means that it can be difficult to compare outcomes between patients and surgeons due to an inconsistent baseline description and treatment selection. Finally, mixed-methods and cohort studies highlighted the importance of return to work in this typically young and economically active patient group.

The recommended research priorities provide direction on next steps. Priorities one and three discuss the potential design of a future RCT. These suggest that a pragmatic approach to design, using an umbrella-type approach with interventions grouped into severity or grade of procedure, would be a appropriate.[123] We have seen similar approaches in other proctology studies.[124,125] The fifth priority

also demonstrates the need to understand interventions which work across the treatment pathway. These are not limited to surgical interventions, as adjuvant treatment such as hair removal, use of antibiotics and wound dressings may also play a role in this pathway.[119] With this in mind, it may be more appropriate and efficient for a funder to commission a multiarm, multistage trial, with rerandomisation of patients who develop recurrence of PSD. The importance of measuring relevant outcomes in a consistent manner is emphasised, with the need for a core outcome set identified as a priority.[126] In addition, participants highlighted the need for the development of patient-reported outcome measures (PROMs). PROMs are important in any core outcome set.[127] A multidimensional PROM may include an assessment of return to work, wound healing and recurrent symptoms, all of which have been identified as essential in prior work packages.

Patient and public involvement

The consensus was based on the engagement of patients and members of the public. Initial findings were discussed with patient representatives, and these ideas were used when delivering the initial workshop. This was also attended by patients who were able to express their priorities. This engagement continued at all stages. During the final consensus meeting, patient representatives were invited to comment regularly to ensure their voice was considered in the final ranking.

Impact for policy-makers

This exercise sets out five clear policy statements which could form the basis for the development of future guidelines by informing PICO (population, intervention, control/comparison, outcome) development. The findings of limited training opportunities following qualification as a consultant surgeon may not be limited to this area, and policy-makers should be cognisant of this when commissioning services.

Impact for researchers

Researchers can use the findings of this study to direct future research. The directives here are for pragmatic trials established to address questions along the treatment pathway. Studies should also have a clear patient focus.

Conclusion

This consensus exercise has involved patients and clinicians to identify five key policy and five key research priorities. These should form the foundation of future work in the field.

Chapter 7 Wysocki classification validation exercise

Background

Clinically, classification systems may have a prognostic function and could ultimately allow stratified treatment. Such systems also ensure that more precise research comparisons can be carried out. While there are some existing classification systems for PSD,[128–135] they are not used in routine practice or for research comparisons. Few studies evaluate their use to inform choice of treatment[130,133–135] and no study has analysed the reliability or predictive validity of such a tool.[110] Given the huge variation in treatment that we have shown both with the PITSTOP surgeon survey and the cohort study, and the existing issues with the current comparative literature, there is a need for a classification tool that is both reliable and valid. A suitably pragmatic classification system could be integrated into clinical practice to support treatment decisions and the counselling of patients on likely outcomes. Such a system should be simple to use, reflect clinical practice and be meaningful in terms of prognostication.

Rationale

There is no commonly used classification tool to characterise PSD. The absence of such a tool represents an important knowledge gap, since surgeons faced with PSD typically have little exposure to the disease and there exists little guidance on how to treat it.

At the inaugural meeting of the Pilonidal Society in Berlin in 2017, a panel of 13 surgeons gathered to establish a simple classification system for PSD.[136] Each member of the panel had a special interest in PSD and had either written other classifications[132,134] or had published widely on PSD. Components of a classification system were longlisted and refined by online consensus involving a large group of PSD surgeons[137] to form a final configuration that was felt to be easy to use, clinically meaningful and had potential statistical validity.

The tool classifies PSD into one of four categories:

- type 1: only midline pit or sinuses
- type 2: any midline disease with secondary sinus/es or abscess scar/s
- type 3: any midline or secondary disease extending below tip of coccyx
- type 4: any disease after treatment with definitive intent.

While there may be agreement among the panel for the eventual classification (referred to as the 'Wysocki classification'), none of the parameters of ease of use, reliability and validity have been tested. We aimed to do so by assessing the level of agreement between surgeons when used in clinical practice.

Aims

The aims of this substudy were to:

- Quantify how well different assessors agree in their classification.
- Quantify how classification relates to surgeon's experience.

Additional exploratory aims were to:

- Identify specific patients with low agreement, which may in turn help clarify or even modify the classification system.
- To present, within each subtype, the frequency with which each treatment option is chosen.

Methods

Participating surgeons

Sampling was purposive and sought to recruit 15 colorectal surgeons. This sampling method aimed for maximum variation based on experience: five expert surgeons who registered an interest in pilonidal disease and who offered a specialised service, five surgeons who carried out pilonidal sinus surgery as part of a general surgical service and five final-year colorectal trainees. Initial contact with surgeons was made by e-mail, and interested surgeons were e-mailed the participant information sheet.

Participating patients

Stimuli required for the validation exercise were obtained from the main cohort study.

All participants referred for elective surgery and participating in the PITSTOP cohort study were asked if they agreed to the surgical site being photographed before surgery (an optional item on the consent form).

Participant photographs of the PSD surgical site were usable for the exercise if they satisfied the following criteria:

- The participant was eligible for the main cohort study (aged 16 years and above with symptomatic PSD, referred for elective surgery).
- The participant consented to a preoperative photograph to be taken of the surgical site.
- The photograph was usable (i.e. unblurred).
- The participant associated with the photograph had been classified using the Wysocki classification system at the time of procedure.

Assessment schedule

Each participant was asked to independently rate 36 cases using the Wysocki classification, with allocation of surgeons to cases selected to ensure overlap with other assessors (*Figure 9*). A total of 90 photographs were each assessed by two specialist surgeons, two general surgeons and two trainee

Patient ID	Assessor ID					Surgical assessor
	S1	S2	S3	S4	S5	
	G6	G7	G8	G9	G10	
	T11	T12	T13	T14	T15	
1 to 9	■	■				1&2, 6&7, 11&12
10 to 18	■		■			1&3, 6&8, 11&13
19 to 27	■			■		1&4, 6&9, 11&14
28 to 36	■				■	1&5, 6&10, 11&15
37 to 45		■	■			2&3, 7&8, 12&13
46 to 54		■		■		2&4, 7&9, 12&14
55 to 63		■			■	2&5, 7&10, 12&15
64 to 72			■	■		3&4, 8&9, 13&14
73 to 81			■		■	3&5, 8&10, 13&15
82 to 90				■	■	4&5, 9&10, 14&15

FIGURE 9 Allocation of patients with PSD to surgical assessors.

surgeons. Surgeons were sent photographs accompanied by the medical history (previous PSD history, including number of elective procedures and emergency drains) electronically. Surgeons were not told the classification as recorded by the original surgeon at the time of procedure.

The substudy was also used to provide an exploratory assessment of surgical opinion. Participating surgeons were asked to record their preferred treatment for each patient they assessed, with the aim being to quantify variation in practice among practitioners. Surgeons recorded their assessments in Microsoft Excel® (Microsoft Corporation, Redmond, WA, USA) spreadsheets which were returned to CTRU and combined into an analysis data set.

Statistical methods

Agreement was quantified as both raw and chance-corrected agreement. Raw agreement is the percentage of patients for whom the assessments agree, while chance-corrected agreement is the ratio of observed to expected agreement. Both raw and chance-corrected agreement are essentially proportions in which one means complete agreement while zero reflects complete disagreement. As the four categories are not ordinal, agreement is a simple yes/no construct in which any difference is considered 'disagreement'.

Raw agreement was defined as (100 × number in which all raters agree / number rated) and was accompanied by a 95% Wilson score interval. Chance-corrected agreement was defined using the unweighted kappa statistic and the unweighted Gwet AC1 statistics.[138] Agreement among surgeons was reported overall and within for expert, general and trainee surgeons. Finally, the agreement was calculated for each surgeon in relation to the original assessment made at the time of procedure. Analyses were conducted using Stata version 17.[139]

Sample size

The sample size was based on: (1) a hypothesis test to rule out a minimal kappa statistic;[140] (2) the standard error of raw agreement; and (3) the number of patients expected to consent and provide usable photographs. For (1), an internal pilot was undertaken in which study surgeons were asked to assess photographs obtained either online or via published articles. A total of 41 pictures were assessed by seven surgeons (five specialists and two trainees) and yielded an overall kappa statistic of 0.42 (0.55 if trainees were excluded). Since the assessments were based on pictures alone and did not include prior history, these may be an underestimate. Based on this, a target of $\kappa = 0.45$ was used. The expected lowest limit for chance-corrected agreement was set at $\kappa = 0.3$, which represents the lowest acceptable agreement if this classification were to be introduced into practice. The sample size calculation also depends on the expected prevalence of each class, which was estimated from the cohort study (approximately 25% type 1, 50% type 2, 10% type 3 and 15% type 4 at the point of data review). A sample size of 90 was adequate to rule out differences of 15% between expected and minimum kappa with 90% power and 5% significance (1); to estimate raw agreement to within a CI half-width of ±10% (2); and to be accommodated by the number of photographs available (expected around 150) (3).

Results

Participants

Fifteen surgeons participated in the classification exercise as described in *Chapter 7*, *Assessment schedule*. A total of 166 patients consented to and provided a photograph of the diseased area, of which three were too unclear to use and were removed. Ninety participants were randomly selected from these for assessment, all of whom had been assigned a classification of 1–4 by the original treating clinician. The majority were classified as type 2 disease at the time of procedure, and 60% underwent a major procedure, with the most common treatments being a Karydakis asymmetric closure ($n = 42$) or glue ($n = 29$) (*Table 23*).

Agreement among participating surgeons

The agreement among surgeons is summarised in *Table 24*. Of the 540 assessments (90 patient photographs and case histories each having 6 assessments), 14 (3%) of assessments were classified 'none of the above', affecting 12 (13%) of the patients.

Overall, the six assessors all reached the same consensus in 38% of participants, with a chance-corrected kappa statistic of 0.52 (95% CI 0.42 to 0.61) and a Gwet AC1 statistic of 0.63 (95% CI 0.56 to 0.69). Agreement between pairs of surgeons was higher, with specialist surgeons agreeing in 72% of patients, general surgeons agreeing in 70% of patients and trainee surgeons agreeing in 71% of patients. The overall agreement is lower since this measure required all six assessors to agree. All six surgeons agreed in 34 (38%) of patients, and five of the six surgeons agreed for 19 (21%). The chance-corrected

TABLE 23 Patients participating in the classification exercise

Characteristic	No (%)
Classification at procedure	
Type 1	14 (16%)
Type 2	52 (58%)
Type 3	10 (11%)
Type 4	14 (16%)
Procedure	
Major excision	54 (60%)
Karydakis	42 (47%)
Bascom's cleft lift	2 (2%)
Leave open	5 (6%)
Leave open (marsupialisation)	1 (1%)
Midline closure	4 (4%)
Minimal excision	36 (40%)
Glue	29 (32%)
EPSiT	6 (7%)
Pit picking	1 (1%)

TABLE 24 Summary of agreement between assessors

Surgeon	No. agree[a]	Percentage (95% CI)	Kappa (95% CI)	Gwet AC1 (95% CI)
Expert	65	72% (62 to 80%)	0.54 (0.38 to 0.69)	0.67 (0.56 to 0.79)
General	63	70% (60 to 78%)	0.54 (0.40 to 0.69)	0.64 (0.52 to 0.76)
Trainee	64	71% (61 to 79%)	0.58 (0.43 to 0.72)	0.65 (0.54 to 0.77)
Overall	34	38% (28 to 48%)	0.52 (0.42 to 0.61)	0.63 (0.56 to 0.69)

a Number in agreement is the number where both assessors agree (or all six assessors for overall agreement).

Note
95% CIs are used throughout.

kappa agreement was above 0.5 (conventionally considered 'moderate') and the Gwet AC1 measure over 0.6 for all subgroups of surgical expertise.

Agreement between participating surgeons and original classification

Each assessor's agreement with the original classification is described in *Figure 10* and *11*. Surgeons were less likely to agree with the original classification than with other surgeons given the same photograph. Raw agreement ranged between 47% (17/36) and 75% (27/36) with chance-corrected kappa statistics ranging from 0.11 to 0.59, and chance-corrected Gwet AC1 agreement statistics from 0.35 to 0.71.

Treatment choice

Surgeons differed markedly when asked how they would treat the individual participants. Overall, surgeons in the substudy recommended a minimally invasive procedure in 46% of cases, but the figure ranged from 0% to 94% (34/36) of patients. There was substantial variation in practice among all levels

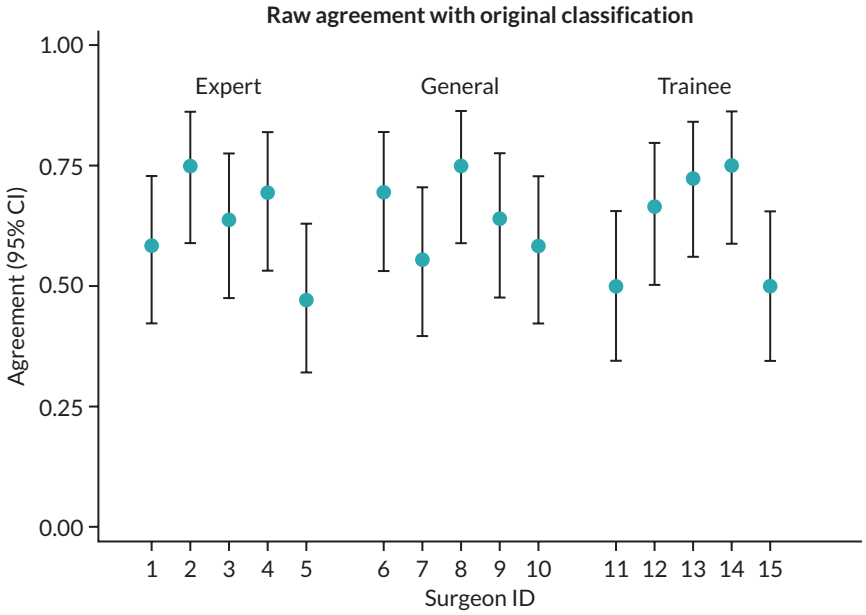


FIGURE 10 Raw agreement between assessment and original.

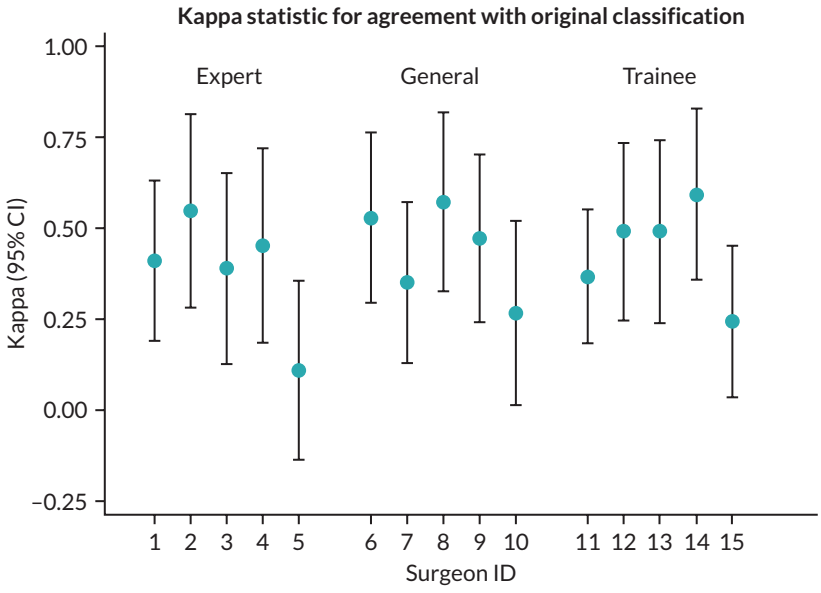


FIGURE 11 Kappa chance-corrected agreement between assessment and original.

of expertise, and the variation in recommendation did not appear related to the actual patients assessed (*Appendix 3*, *Figure 16*).

The surgeons surveyed were slightly more likely to recommend minimally invasive surgery (46%) than those that actually received this approach (40%). Although the percentage favouring asymmetric closure (45%) was similar to the treatment actually received (49%), the specific procedure types differed: the surgeons surveyed were more likely to use a Bascom's cleft lift (22%) than a Karydakis (16%). The use of midline closure (5%) and leave open (4%) approaches were uncommon, most notably among the specialist surgeons surveyed.

Discussion

There are three potential roles of a classification system. Two are clinical – predicting prognosis and guiding treatment – and the third is primarily for research purposes, allowing more precise comparative studies to be carried out and reducing the potential for selection bias. While there have been eight previously proposed classification systems for PSD,[110] each used judgemental methodology to develop their systems and identified homogeneous categories based on the experience of the investigators. The classifications were mainly used to select patients for different procedures. However, none provided analyses to demonstrate ease of use, reliability or predictive criterion validity. We have shown that the Wysocki classification, developed from components of these other systems, demonstrates moderate but acceptable agreement among the surgeons participating in this classification exercise. Agreement was similar among specialist, general and trainee colorectal surgeons, which offers reassurance that this system could be used across a range of surgeons.

While the level of agreement between substudy assessors and the original classification was lower, this is likely to be attributable to a mixture of picture quality and other features that were not available to the assessor in this exercise. A minority of cases (2.5%) were considered not to fall into any of the four categories, which was similar to the incidence seen in the cohort study (*Chapter 3*). This itself is relevant and indicates that this classification system incorporates reliably definable disease characteristics for almost all presentations of disease.

The findings from this substudy inform the main study, and vice versa. The general agreement seen in this substudy among surgeons lends weight to the Wysocki classification as an objective measure when conducting the risk adjustment. In turn, the cohort study found this classification to be an important feature in predicting response to treatment, with class 1 disease, in particular, being associated with more favourable outcomes. If the Wysocki classification is used to prognosticate and to inform the appropriate treatment, it is therefore important to demonstrate its reproducibility. While this substudy demonstrated only moderate agreement, its magnitude exceeded previously reported inter-rater reliability of other surgical classifications such as grading of haemorrhoids ($\kappa = 0.38$[141]) and dysplastic colorectal adenomas ($\kappa = 0.38$[142]). These findings support the use of the Wysocki classification in accurately defining subtypes of PSD.

Finally, the opportunistic question 'what treatment would you recommend' – while not central to the validation exercise – triangulates the findings of the consultant surgeon survey (see *Chapter 2*) in demonstrating that the choice of procedure is highly surgeon-dependent rather than evidence-based. This is, perhaps surprisingly, even the case in the specialist group where, for instance, the variation in minimally invasive procedures is immense. This reiterates that the disease characteristics are not the sole driver of treatment choice. This matters for the risk-adjusted comparisons in the cohort study, since non-randomised comparisons are known to be biased in situations when (1) treatment is defined by severity and (2) severity cannot easily be quantified and modelled.

Chapter 8 Discussion

Overview

The Idea, Development, Exploration, Assessment, Long-term follow-up (IDEAL) framework[143] was designed to improve the quality of surgical research. Several large, multicentre, prospective longitudinal cohorts have followed this framework, allowing some understanding of variations in practice and their effects on outcomes of other surgical techniques.[144–146] PITSTOP also followed this framework, with the aim of answering some key questions in the surgical field of pilonidal disease. The need for this approach to pilonidal disease is clear. While the literature on this subject is vast, the quality of this literature is poor. An initial mapping exercise found that only 12% of the 983 identified primary research articles were randomised trials, and our current understanding of treatment relies mainly on poorly designed cohort studies which cannot provide us with reliable and reproducible estimates of treatment effects.[119] There is an absence of clear front-running surgical interventions.[13] Interventions are numerous and there are issues with heterogeneity of definitions and measurements of outcome.[13]

Given the multitude of management options available, the initial work stream focused on surgeon preferences. A survey of over 100 UK surgeons was considered representative of real-world UK practice. Even with evidence to the contrary, a substantial proportion of surgeons who answered the survey perceived very high failure rates regardless of intervention. Again, regardless of the evidence, many practised non-surgical interventions and carried out procedures that experts in the field would consider obsolete (namely excision and leave open or midline closure). Even when minimally invasive procedures were perhaps appropriate, they were not considered as options by a substantial proportion.

This apparent disregard for the evidence could relate to the recognition of a low-quality evidence base and dismissal of the literature.[13] It could also relate to the unglamorous nature of pilonidal surgery. The specialist colorectal surgeon who is often tasked with managing the condition in the UK may tend to focus on more challenging conditions. ‘Pilonidal sinus cases are often just fillers on all day lists’ is a quote from an involved member of the core clinical team. The apprentice style of UK surgical training for pilonidal disease may also offer an explanation for the perpetuation of outdated techniques or even newer techniques done incorrectly. A mentor surgeon who has ‘always done it this way’ and a training syllabus that fails to keep pace with modern developments may be contributors.

Although a survey hints at the real-world experience of pilonidal sinus surgery in the UK, a more robust confirmation of this experience is required. The main work package of the PITSTOP trial was therefore a prospective cohort study. Involving over 30 centres throughout the UK, this again was considered reflective of current practice. The multitude of procedures utilised, and the most performed procedures, were consistent with the survey findings. Indeed, the continued use of potentially obsolete procedures (excise and leave open and midline closure) was also confirmed.

While 40% of procedures were classed as minimally invasive, the disease characteristics of the cohort suggest that more patients would have been eligible for such techniques. This is pertinent because such procedures were shown to result in less pain, less complication risk and more rapid return to normal activities, even after accounting for case mix. While the chance of cure was increased with the more major excisions, it may be that some patients would prefer to trade more rapid and less complex recovery for a moderate decrease in efficacy.

Perhaps the most surprising result was the protracted time a significant proportion of patients took to heal and return to normal activities. It is possible that one-quarter of patients had not healed or had persistent disease at 6 months regardless of type of intervention. In an essentially active working population, possibly one in eight patients had not resumed normal activities. This has important

implications in terms of the impact on health resource as well as the ability to counsel patients accurately before surgery. It contradicts most of the literature on the subject, with many studies reporting extremely low failure rates.[39,40,147]

One explanation for this difference between our findings and the reported literature is the definition of 'failure' compared with 'recurrence'. This is a major issue with the existing evidence. Very few studies attempt to define what is meant by 'recurrence'. A review of the relevant literature taken from the most recent guidelines[148] indicates that over 85% of studies that investigate recurrence as an outcome fail to define it at all.

In the true sense of the word, 'recurrence' refers to disease that has resolved after an intervention but then recurs after a sufficient period to indicate it is not simply the original disease. It should be differentiated from disease that never resolves, or symptoms that remain unresolved due to an unhealed wound. Therefore, if previous studies report on true recurrence, the incidence may be very low as this is likely to be rare. The most clinically valid outcome is a combination of true recurrence and failure of healing after a reasonable time point. We reported on this combination as being relevant clinically. However, the incomplete data on 'recurrence' specifically after at least 6 months and the fact that patients themselves reported on 'recurrence', adding an element of subjectivity, may limit our interpretation of the results, and further in-depth analysis is required to allow robust comparisons with other studies.

Another explanation of the difference between our data and the literature is the skill of the surgeon. The cohort study included multiple surgeons of varied expertise. It may be that certain experts can achieve the success portrayed in the literature.[11] Analysis of individual surgeon data did show a difference in treatment failure rates between surgeons. Although numbers were small, this may justify that for optimal care patients, particularly those with complex disease, should be referred to specialist units. Alternatively, the skills of the more general surgeon should be enhanced.

Data from the consultant survey and the cohort study revealed a preference in favour of more aggressive interventions rather than minimally invasive procedures. This suggests that some surgeons may focus on cure rather than symptomatic improvement and believe that more major procedures result in a higher chance of cure even if minimally invasive procedures are possible. Our cohort data confirm a higher cure rate. However, patients may wish to trade this 10–15% increased chance of cure for significantly less pain, fewer complications and a more rapid return to normal activities. We explored this hypothesis utilising two qualitative methodologies: a mixed-methods substudy and a DCE.

The mixed-methods study suggested a lack of SDM for some patients, with many not being given a choice of procedure or informed fully about postoperative burden of care. This led to high levels of DR when procedures were not completely successful and protracted periods of recovery became necessary. Even when patients were involved in the decision-making, if not fully informed about postsurgical pathways, their expectations were often not met. Sufficient and accurate information about the risks of protracted postprocedural aftercare should be highlighted to address the false optimism many patients may have.[99,101] If given a choice, some patients may elect for alternative procedures where outcomes may differ.

The element of patient choice was explored further in the DCE. While cure of the disease was considered a priority by many, some were prepared to trade the chance of complete cure for a less protracted recovery. This was particularly the case for older patients (> 30 years) where an acceptance of up to 35% increased risk of persistent disease was tolerated in exchange for a shorter recovery period. This again emphasises the need for SDM and tailoring treatment according to the individual patient and their treatment goals.

While these two workstreams suggested a need for improved decision-making and the potential for DR after surgery, the data from the cohort study looking at these parameters revealed conflicting results.

The median CollaboRATE score, a tool for assessing the quality of SDM, was very high. In addition, 84% of those who were assessed for DR were either very satisfied or satisfied with the surgery. This is despite around 45% having complications of surgery and 25% having persistent symptoms 6 months after surgery. These contrasting data could be explained by social desirability bias – the tendency to report higher scores out of gratitude or deference.[46]

Of course, SDM becomes difficult if the surgeon only specialises in one technique.[30] Such surgeons should consider expanding their armamentarium to provide an individualised recommendation and choice for the patient, or consider referring to a specialist who may be able to offer such a service.

The literature on pilonidal sinus surgery is confusing and misleading due not only to multiple interventions and no obvious gold standard comparator, the lack of definitions (particularly of recurrence), but also to the heterogeneity of disease severity. Many researchers make no attempt to classify or stratify disease. As such, it is difficult to draw meaningful conclusions about comparative studies. Attempts to classify disease for the purposes of improving the quality of research have been made, and these have been reviewed as part of the PITSTOP study.[110] The Wysocki classification demonstrated moderate but acceptable agreement among the surgeons participating in this classification exercise. While there was only moderate agreement, the kappa exceeded previously reported inter-rater reliability of other surgical classifications such as grading of haemorrhoids[120] and dysplastic colorectal adenomas.[142] Agreement was similar among specialist, general and trainee colorectal surgeons, which offers reassurance that this system could be used across a range of surgeons. Only 2.5% of cases were considered not to fall in the four categories, indicating that the classification system incorporates reliably definable disease characteristics for almost all presentations of disease. Finally, there was a suggestion that the classification could be prognostically valid, with class 1 disease, in particular, being associated with more favourable outcomes.

We concluded the PITSTOP study with a consensus exercise. This utilised a 'so what, now what' workshop incorporating data from the cohort study, an e-Delphi exercise and Qualtrics survey technology to consolidate patient and survey views as to the front-running policy and research statements. The policy statements highlighted some key outcomes from the other work packages and are included in the implications for practice and research discussed below.

Implications for practice

While minimally invasive procedures may not be suitable for all, they should form part of the armamentarium of each pilonidal sinus surgeon. Such interventions fit with the philosophy of not making the surgery worse than the disease itself.

The perceived high failure rate for pilonidal disease by many UK surgeons is a concern. Perpetuation of obsolete techniques by a substantial proportion, combined with newer techniques potentially done badly, emphasises the need for better guidance and training. National associations should take on this challenge by providing up-to-date guidance and influencing training through workshops and mentorship programmes.

Shared decision-making is essential, with patients offered an array of interventions allowing them to choose based on preferred outcomes. Many patients would be happy to trade a shorter recovery period for less chance of cure. They value the time to return to normal activities as an outcome, and this should be included in the decision-making process, aiding selection of interventions. If surgeons practise a 'one fit for all' intervention, they should consider learning a broader range of techniques or referring patients to a surgeon who can offer this service. An individualised approach based on the severity of disease and the wishes of the patient, combined with detailed information about interventions and potential aftercare, will improve patient expectations and reduce DR.

The Wysocki classification seems to provide a reliable tool and has some validity when it comes to prognostication. Further work is required to develop the tool to include some form of stratification of disease. This will help the surgeon in the choice of which interventions are appropriate.

Implications for research

The grouping of procedures into those that involve major excision and those that are minimally invasive could simplify both the process and the interpretation of future comparative trials. The impression from the core clinical team involved in PITSTOP was that this grouping was fair, although perhaps with the exclusion of excise-and-leave-open and midline closure techniques in the major excision group.

A classification system involving relevant disease characteristics is essential if future comparative trials are to be interpreted and meta-analysed in a meaningful way. Such a system should strive to include some form of stratification of disease severity. The Wysocki classification goes some way to meeting these requirements. Further application may allow development of a treatment algorithm or decision tree to aid surgical decision-making.

Future trials should include a robustly developed core outcome set which includes important PROs.

Equality, diversity and inclusion

No active steps were taken to make participation representative. Participants were representative of the disease population, with people of different disease severity included. White people were marginally overrepresented (see *Table 5*): 85% in PITSTOP versus 82% in the UK. We achieved a representative sample of Asian/Asian British people (9% in PITSTOP vs. 9% in the UK). We somewhat under-represented mixed/multiple ethnic groups (2% in PITSTOP vs. 3% in the UK) and black/African/Caribbean/Black British (1% in PITSTOP vs. 4% in the UK). This deficit could be addressed in future studies by opening more sites in London and the West Midlands and developing materials that are inclusive, accessible and encouraging to under-represented groups. Our core research team includes non-white members and a range of clinical and methodological expertise. Development opportunities were provided for entry-level researchers to present at conferences[149] and act as first/corresponding authors[30,110] on papers in peer-reviewed journals. The Associate PI scheme gave five junior clinicians, two of them non-white, the opportunity to contribute to the study.

Patient and public involvement

Patients informed the design of this study, ensured the methods selected were appropriate for patients, and reviewed and commented on questionnaires and other data collection methods. They advised on the appropriateness of the plain English summary and were named as co-applicants. Patient representatives steered the project through the research process, attending management group meetings. They assisted in the design of the protocol, patient information and consent forms. In the mixed-methods substudy, they assisted the analysts in developing themes from the data and contributed to the interpretation of data, with one person with lived experience acting as a co-author on the resulting publication. Expert patients were participants in the Delphi survey. Patients have helped us to design plain English summaries of findings.

Additional information

Contributions of authors

Steven Brown (https://orcid.org/0000-0002-0980-2793) (Professor of Surgery, Chief Investigator) conceived of or designed the work, was involved in the acquisition of data for the work, was involved in the interpretation of data for the work, drafted the monograph, revised the work critically for important intellectual content, was involved in the final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Daniel Hind (https://orcid.org/0000-0002-6409-4793) (Professor of Evaluation, Assistant Director) conceived of or designed the work, was involved in the acquisition of data for the work, was involved in the interpretation of data for the work, drafted the monograph, revised the work critically for important intellectual content, was involved in the final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Emily Strong (https://orcid.org/0000-0002-2381-4088) (Study Manager) was involved in the interpretation of data for the work, drafted the monograph, revised the work critically for important intellectual content, was involved in the final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Mike Bradburn (https://orcid.org/0000-0002-3783-9761) (Senior Statistician) conceived of or designed the work, was involved in the acquisition of data for the work, was involved in the interpretation of data for the work, drafted the monograph, revised the work critically for important intellectual content, was involved in the final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Farhat Vanessa Nasim Din (https://orcid.org/0000-0001-5466-8282) (Professor of Surgery, Co-applicant) conceived of or designed the work, was involved in the acquisition of data for the work, revised the work critically for important intellectual content, was involved in the final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Ellen Lee (https://orcid.org/0000-0003-4529-7410) (Statistician Research Fellow) conceived of or designed the work, was involved in the acquisition of data for the work, was involved in the interpretation of data for the work, drafted the monograph, revised the work critically for important intellectual content, was involved in the final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Matthew J Lee (https://orcid.org/0000-0001-9971-1635) (Clinical Research Fellow, Co-applicant) conceived of or designed the work, was involved in the acquisition of data for the work, was involved in the interpretation of data for the work, drafted the monograph, revised the work critically for important intellectual content, was involved in the final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Jonathan Lund (https://orcid.org/0000-0001-5195-2181) (Professor of Surgery, Co-applicant) conceived of or designed the work, was involved in the acquisition of data for the work, revised the work critically for important intellectual content, was involved in the final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Christine Moffatt (https://orcid.org/0000-0002-2436-0129) (Wound Management Nurse, Co-applicant) conceived of or designed the work, was involved in the acquisition of data for the work, revised the work critically for important intellectual content, was involved in the final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Jonathan Morton (https://orcid.org/0000-0003-0544-5474) (Consultant Surgeon, Co-applicant) conceived of or designed the work, was involved in the acquisition of data for the work, revised the work critically for important intellectual content, was involved in the final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Asha Senapati (https://orcid.org/0000-0003-2597-9967) (Consultant Surgeon, Co-applicant) conceived of or designed the work, was involved in the acquisition of data for the work, revised the work critically for important intellectual content, was involved in the final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Philip Shackley (https://orcid.org/0000-0002-1862-0596) (Reader in Health Economics, Co-applicant) conceived of or designed the work, was involved in the acquisition of data for the work, revised the work critically for important intellectual content, was involved in the final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Peter Vaughan-Shaw (https://orcid.org/0000-0002-9790-6882) (Surgical Trainee, Co-applicant) conceived of or designed the work, was involved in the acquisition of data for the work, revised the work critically for important intellectual content, was involved in the final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Arkadiusz Peter Wysocki (https://orcid.org/0000-0001-6880-9285) (Consultant Surgeon, Co-applicant) conceived of or designed the work, revised the work critically for important intellectual content, was involved in the final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Tia Callaghan (https://orcid.org/0000-0003-3255-2849) (Study Manager) was involved in the acquisition of data for the work, was involved in the interpretation of data for the work, produced the first draft of the report (with Nyantara Wickramasekera), revised the work critically for important intellectual content, was involved in the final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Helen Jones (https://orcid.org/0000-0001-8238-7779) (Consultant Surgeon) was involved in the acquisition of data for the work, revised the work critically for important intellectual content, was involved in the final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Nyantara Wickramasekera (https://orcid.org/0000-0002-6552-5153) (Health Economist Research Fellow) was involved in the interpretation of data for the work, produced the first draft of the report (with Tia Callaghan), revised the work critically for important intellectual content, was involved in the final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acknowledgements

We gratefully acknowledge the hard work, support and advice from the following: research nurses and support staff in the 33 participating NHS Trusts for participant screening, consent, data collection and patient follow-up. We thank all participants who took part in the study.

We acknowledge advice and oversight from members of the project steering committee: Dr Dermot Burke (Chair, Consultant Colorectal Surgeon), Mr Stephen Chapman (NIHR Academic Clinical Fellow), Professor Catrin Tudor-Smith (Statistician) and Mr Ryan Edridge (Patient representative).

The following were also involved in the acquisition of data for the work:

Dale Vimalachandran (consultant surgeon), Jennie Grainger (consultant surgeon), Jeremy Wilson (consultant surgeon), Jared Torkington (consultant surgeon), Sandeep Kapur (consultant surgeon), Richard Stevenson (consultant surgeon), Najam Husain (consultant surgeon), Karim Muhammad (consultant surgeon), Richard Brady (consultant surgeon), David Donnelly (consultant surgeon), Michael Lim (consultant surgeon), Raj Rajaganeshan (consultant surgeon), Martyn Evans (consultant surgeon), Lyndsay Pearce (consultant surgeon), Godwin Dennison (consultant surgeon), Paul Mackay (consultant surgeon), Kenneth Keogh (consultant surgeon), Sudhaker Mangam (consultant surgeon), Mohan Harilingham (consultant surgeon), James Pitt (consultant surgeon), Sanjay Chaudri (consultant surgeon), Franceso De Fabio (consultant surgeon), Graham Branagan (consultant surgeon), Alex Hardy (consultant surgeon), Ali Khalafalla (consultant surgeon), Feliz Mazarelo (consultant surgeon), Yasuko Maeda (consultant surgeon), Nikhill Pawa (consultant surgeon).

Special thanks to Jennie Grainger (consultant surgeon), Khalafalla Ali (colorectal surgeon), Sunanda Mahapatra (general surgeon) and Peter Vaughan-Shaw for proofreading draft chapters of the monograph.

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 anonymised data may be granted following review.

Ethics statement

The study received approval from East of England – Cambridge South Research Ethics Committee (REC reference 18/EE/0370) on 26 November 2018.

Information governance statement

NIHR and The University of Sheffield 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 Data Protection legislation The University of Sheffield is the Data Processor; Sheffield Teaching Hospitals NHS Foundation Trust is the Data Controller, and we process personal data in accordance with their instructions. You can find out more about how we handle personal data, including how to exercise your individual rights and the contact details for University of Sheffield's Data Protection Officer, at (www.sheffieldclinicalresearch.org/).

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/KFDQ2017.

Primary conflicts of interest: All authors have completed the unified competing interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare (1) no financial support for the submitted work from anyone other than their employer; (2) no financial relationships with commercial entities that might have an interest in the submitted work; (3) no spouses, partners, or children with relationships with commercial entities that might have an interest in the submitted work; and (4) no non-financial interests that may be relevant to the submitted work. Mike Bradburn is a current member of the HTA Commissioning Committee. Steven Brown was a member of the HTA Commissioning Committee from October 2017 to September 2019. Daniel Hind was a member of the HTA Clinical Evaluation and Trials Committee and HTA Fast Track Committee – June 2021.

Publications

Beal EM, Lee MJ, Hind D, Wysocki AP, Yang F, Brown SR. A systematic review of classification systems for pilonidal sinus. *Tech Coloproctol* 2019;**23**(5):435–43.

Kumar M, Clay WH, Lee MJ, Brown SR, Hind D. A mapping review of sacrococcygeal pilonidal sinus disease. *Tech Coloproctol* 2021;**25**(6):675–82.

Strong E, Callaghan T, Beal E, Moffatt C, Wickramasekera N, Brown S, *et al.*, PITSTOP Project Management Group, PITSTOP Collaborators. Patient decision-making and regret in pilonidal sinus surgery: a mixed-methods study. *Colorectal Dis* 2021;**23**(6):1487–98.

Wickramasekera N, Strong E, Shackley P, Callaghan T, Lee M, Hind D, Brown S, PITSTOP Project Management Group and PITSTOP Collaborators. Patient preferences for pilonidal sinus treatments: a discrete choice experiment survey. *Colorectal Dis* 2023;**25**:984–994. https://doi.org/10.1111/codi.16482

Brown SR, Hind D, Strong E, Bradburn M, Din F, Lee E, *et al.* PITSTOP Management Group. Real-world practice and outcomes in pilonidal surgery: Pilonidal Sinus Treatment Studying The Options (PITSTOP) cohort. *Br J Surg* 2024;**111**(3):znae009. https://doi.org/10.1093/bjs/znae009

Lee MJ, Lee E, Bradburn M, Hind D, Strong EB, Din F, *et al.* PITSTOP Management Group and the PITSTOP Validators. Classification and stratification in pilonidal sinus disease: findings from the PITSTOP cohort [published online ahead of print 21 April 2024]. *Colorectal Dis* 2024. https://doi.org/10.1111/codi.16989.

References

1. Søndenaa K, Andersen E, Nesvik I, Søreide JA, Sondenaa K, Andersen E, *et al.* Patient characteristics and symptoms in chronic pilonidal sinus disease. *Int J Colorectal Dis* 1995;**10**(1):39–42.
2. Hull TL, Wu J. Pilonidal disease. *Surg Clin North Am* 2002;**82**(6):1169–85.
3. Karydakis GE. Easy and successful treatment of pilonidal sinus after explanation of its causative process. *Aust N Z J Surg* 1992;**62**(5):385–9.
4. Ertan T, Koc M, Gocmen E, Aslar AK, Keskek M, Kilic M, *et al.* Does technique alter quality of life after pilonidal sinus surgery? *Am J Surg* 2005;**190**(3):388–92.
5. Jensen SL, Harling H. Prognosis after simple incision and drainage for a first-episode acute pilonidal abscess. *Br J Surg* 1988;**75**(1):60–1.
6. Elsey E, Lund JN. Fibrin glue in the treatment for pilonidal sinus: high patient satisfaction and rapid return to normal activities. *Tech Coloproctol* 2013;**17**(1):101–4.
7. Patti R, Angileri M, Migliore G, Sparancello M, Termine S, Crivello F, *et al.* Use of fibrin glue in the treatment of pilonidal sinus disease: a pilot study. *G Chir* 2006;**27**(8–9):331–4.
8. Thompson MR, Senapati A, Kitchen P. Simple day-case surgery for pilonidal sinus disease. *Br J Surg* 2011;**98**(2):198–209.
9. Enriquez-Navascues JM, Emparanza JI, Alkorta M, Placer C. Meta-analysis of randomized controlled trials comparing different techniques with primary closure for chronic pilonidal sinus. *Tech Coloproctol* 2014;**18**(10):863–72.
10. Hospital Episode Statistics. *Admitted Patient Care – England 2012-13*; 2013.
11. Vermeulen H, Ubbink D, Goossens A, de Vos R, Legemate D. Dressings and topical agents for surgical wounds healing by secondary intention. *Cochrane Database Syst Rev* 2004;**2004**(2):CD003554.
12. Price PE, Butterworth RJ, Bale S, Harding KG. Measuring quality of life in patients with granulating wounds. *J Wound Care* 1994;**3**(1):49–50.
13. Stauffer VK, Luedi MM, Kauf P, Schmid M, Diekmann M, Wieferich K, *et al.* Common surgical procedures in pilonidal sinus disease: a meta-analysis, merged data analysis, and comprehensive study on recurrence. *Sci Rep* 2018;**8**(1):3058.
14. Brown SR, Lund JN. The evidence base for pilonidal sinus surgery is the pits. *Tech Coloproctol* 2019;**23**(12):1173–5.
15. Al-Khamis A, McCallum I, King PM, Bruce J. Healing by primary versus secondary intention after surgical treatment for pilonidal sinus. *Cochrane Database Syst Rev* 2010;**2010**(1):CD006213.
16. McCallum IJD, King PM, Bruce J. Healing by primary closure versus open healing after surgery for pilonidal sinus: systematic review and meta-analysis. *BMJ* 2008;**336**(7649):868–71.
17. Harris CL, Holloway S. Development of an evidence-based protocol for care of pilonidal sinus wounds healing by secondary intent using a modified Reactive Delphi procedure Part 2: methodology, analysis and results. *Int Wound J* 2012;**9**(2):173–88.
18. Lund J, Tou S, Doleman B, Williams JP. Fibrin glue for pilonidal sinus disease. *Cochrane Database Syst Rev* 2017;**1**(1):CD011923.
19. Handmer M. Sticking to the facts: a systematic review of fibrin glue for pilonidal disease. *ANZ J Surg* 2012;**82**(4):221–4.

20. Lund JN, Leveson SH. Fibrin glue in the treatment of pilonidal sinus: results of a pilot study. *Dis Colon Rectum* 2005;**48**(5):1094–6.

21. Isik A, Eryılmaz R, Okan I, Dasiran F, Firat D, Idiz O, *et al.* The use of fibrin glue without surgery in the treatment of pilonidal sinus disease. *Int J Clin Exp Med* 2014;**7**(4):1047–51.

22. Eysenbach G. Improving the quality of web surveys: the checklist for reporting results of internet e-surveys (CHERRIES). *J Med Internet Res* 2004;**6**(3):e34.

23. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, CondeJG. Research electronic data capture (REDCap) – a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2008;**42**(2):377–81. https://doi.org/10.1016/j.jbi.2008.08.010

24. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, *et al.* The REDCap consortium: building an international community of software platform partners. *J Biomed Inform* 2019; **95**:103208. https://doi.org/10.1016/j.jbi.2019

25. Ojo D, Flashman K, Thomas G, Tozer P, Senapati A. Cleft closure (Bascom's cleft lift) for 714 patients – treatment of choice for complex and recurrent pilonidal disease (a cohort study). *Colorect Dis* 2023;**25**(9):1839–43.

26. Fearnhead NS, Lee MJ, Acheson AG, Worley G, Faiz OD, Brown SR. Variation in practice of pouch surgery in England – using SWORD data to cut to the chase and justify centralization. *Colorect Dis* 2018;**20**(7):597–605.

27. O'Connell E, McDevitt J, Hill ADK, McNamara DA, Burke JP. Centralisation of rectal cancer care has improved patient survival in the republic of Ireland. *Eur J Surg Oncol* 2022;**48**(4):890–5.

28. Shabbir J, Chaudhary BN, Britton DC. Management of sacrococcygeal pilonidal sinus disease: a snapshot of current practice. *Int J Colorectal Dis* 2011;**26**(12):1619–20.

29. Burnett D, Smith SR, Young CJ. The surgical management of pilonidal disease is uncertain because of high recurrence rates. *Cureus* 2018;**10**(5):e2625.

30. Strong E, Callaghan T, Beal E, Moffatt C, Wickramasekera N, Brown S, *et al.* Patient decision-making and regret in pilonidal sinus surgery: a mixed-methods study. *Color Dis* 2021;**23**(6):1487–98.

31. Murphy S, Wysocki AP. Pilonidal sinus disease surveys. *Pilonidal Sinus J* 2017;**3**(1):19–27.

32. Schein M. Treating pilonidal disease: you do not need to detonate a naval mine to catch a fish. *World J Surg* 2017;**41**(5):1303–4.

33. Fabricius R, Petersen LW, Bertelsen CA. Treatment of pilonidal sinuses in Denmark is not optimal. *Dan Med Bull* 2010;**57**(12):A4200.

34. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival analysis part III: multivariate data analysis – choosing a model and assessing its adequacy and fit. *Br J Cancer* 2003;**89**(4):605–11.

35. Busso M, DiNardo J, McCrary J. New evidence on the finite sample properties of propensity score reweighting and matching estimators. *Rev Econ Stat* 2014;**96**(5):885–97.

36. Harlak A, Mentes O, Kilic S, Coskun K, Duman K, Yilmaz F. Sacrococcygeal pilonidal disease: analysis of previously proposed risk factors. *Clinics* 2010;**65**(2):125–31.

37. Iesalnieks I, Ommer A, Petersen S, Doll D, Herold A. German national guideline on the management of pilonidal disease. *Langenbecks Arch Surg* 2016;**401**(5):599–609.

38. Steele SR, Perry WB, Mills S, Buie WD; Surgeons SPTF of the AS of C and R. Practice parameters for the management of pilonidal disease. *Dis Colon Rectum* 2013;**56**(9):1021–7.

39. Baur T, Stauffer VK, Vogt AP, Kauf P, Schmid M, Luedi MM, *et al.* Recurrence rates after uncommon surgical procedures for pilonidal sinus disease. *Coloproctology* 2019;**41**(2):96–100.

40. Bi S, Sun K, Chen S, Gu J. Surgical procedures in the pilonidal sinus disease: a systematic review and network meta-analysis. *Sci Rep* 2020;**10**(1):13720.

41. Milone M, Velotti N, Manigrasso M, Anoldo P, Milone F, De Palma GD. Long-term follow-up for pilonidal sinus surgery: a review of literature with metanalysis. *Surgeon* 2018;**16**(5):315–20.

42. Gallo G, Carpino A, De Paola G, Fulginiti S, Novelli E, Ferrari F, *et al.* Endoscopic pilonidal sinus treatment: a tertiary care academic center experience. *Front Surg* 2021;**8**:723050.

43. Yildiz MK, Ozkan E, Odabasi HM, Kaya B, Eris C, Abuoglu HH, *et al.* Karydakis flap procedure in patients with sacrococcygeal pilonidal sinus disease: experience of a single centre in Istanbul. *ScientificWorldJournal* 2013;**2013**:807027.

44. Sinnott CJ, Glickman LT. Limberg flap reconstruction for sacrococcygeal pilonidal sinus disease with and without acute abscess: our experience and a review of the literature. *Arch Plast Surg* 2019;**46**(03):235–40.

45. Immerman SC. The Bascom cleft lift as a solution for all presentations of pilonidal disease. *Cureus* 2021;**13**(2):e13053.

46. Badejo MA, Ramtin S, Rossano A, Ring D, Koenig K, Crijns TJ. Does adjusting for social desirability reduce ceiling effects and increase variation of patient-reported experience measures? *J Patient Exp* 2022;**9**:237437352210791.

47. Tsai T, Shih L-C, Lee IT, Ng T-Y, Wang J-Y, Hsu C-L, *et al.* Older age is associated with better compliance with follow-up in Taiwan after functional endoscopic sinus surgery. *In Vivo* 2020;**34**(5):2571–6.

48. DiMatteo MR. Variations in patients' adherence to medical recommendations. *Med Care* 2004;**42**(3):200–9.

49. Dag A, Colak T, Turkmenoglu O, Sozutek A, Gundogdu R. Phenol procedure for pilonidal sinus disease and risk factors for treatment failure. *Surgery* 2012;**151**(1):113–7.

50. Boztug CY, Karaagac Akyol T, Benlice C, Koc MA, Doganay Erdogan B, Ozcebe OI, *et al.* Platelet-rich plasma treatment improves postoperative recovery in patients with pilonidal sinus disease: a randomized controlled clinical trial. *BMC Surg* 2021;**21**(1):373.

51. Milone M, Velotti N, Manigrasso M, Vertaldi S, Di Lauro K, De Simone G, *et al.* Long-term results of a randomized clinical trial comparing endoscopic versus conventional treatment of pilonidal sinus. *Int J Surg* 2020;**74**:81–5.

52. Elbanna HG, Emile SH, Youssef M, Thabet W, El-Hamed TMA, Ghnnam WM. Novel approach of treatment of pilonidal sinus disease with thrombin gelatin matrix as a sealant. *Dis Colon Rectum* 2016;**59**(8):775–80.

53. Milone M, Fernandez LMS, Musella M, Milone F. Safety and efficacy of minimally invasive video-assisted ablation of pilonidal sinus: a randomized clinical trial. *JAMA Surg* 2016;**151**(6):547–53.

54. Yang Y, Yu L, Wang Y, Shi J, Li J, Shang F, *et al.* Comparative analysis on the effect of Z-plasty versus conventional simple excision for the treatment of sacrococcygeal pilonidal sinus: A retrospective randomised clinical study. *Int Wound J* 2020;**17**(3):555–61.

55. Yildiz A. Partial primary closure in sacrococcygeal pilonidal sinus: Modified with suture technique. *Asian J Surg* 2022;**45**(1):381–5.

56. Shabbir F, Ayyaz M, Farooka MW, Toor AA, Sarwar H, Malik AA. Modified Limberg's flap versus primary closure for treatment of pilonidal sinus disease: a comparative study. *J Pak Med Assoc* 2014;**64**(11):1270–3.

57. Arslan K, Said Kokcam S, Koksal H, Turan E, Atay A, Dogru O. Which flap method should be preferred for the treatment of pilonidal sinus? A prospective randomized study. *Tech Coloproctol* 2014;**18**(1):29–37.

58. Saydam M, Ozturk B, Sinan H, Balta AZ, Demir P, Ozer MT, *et al.* Comparison of modified Limberg flap transposition and lateral advancement flap transposition with Burow's triangle in the treatment of pilonidal sinus disease. *Am J Surg* 2015;**210**(4):772–7.

59. Yin R. *Case Study Research: Identifying Your Case(s) and Establishing the Logic of Your Case Study*. London: Sage Publications; 2014.

60. Elwyn G, James P, Grande SW, Thompson R, Walsh T, Ozanne EM. Patient Education and Counseling Developing CollaboRATE: a fast and frugal patient-reported measure of shared decision making in clinical encounters. *Patient Educ Couns* 2013;**93**(1):102–7.

61. Brehaut JC, O'Connor AM, Wood TJ, Hack TF, Siminoff L, Gordon E, *et al.* Validation of a decision regret scale. *Med Decis Making* 2003;**23**:281–92.

62. Guest G, Bunce A, Johnson L. How many interviews are enough?: An experiment with data saturation and variability. *Field Methods* 2006;**18**(1):59–82.

63. Morse JM. Data were saturated. *Qual Health Res* 2015;**25**(5):587–8.

64. Witt J, Elwyn G, Wood F, Brain K. Decision making and coping in healthcare: the coping in deliberation (CODE) framework. *Patient Educ Couns* 2012;**88**(2):256–61.

65. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. *BMC Health Serv Res* 2017;**17**(1):88.

66. Ritchie J, Cartwright MFJ. Chapter 11. In: Bryman A, Burgess RG, editors. *Analyzing Qualitative Data*. London: Routledge; 1994. pp. 173–94.

67. Yin RK. Mixed methods research: are the methods genuinely integrated or merely parallel? *Res Sch* 2006;**13**(1):41–8.

68. Guetterman TC, Fetters MD, Creswell JW. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. *Ann Fam Med* 2015;**13**(6):554–61.

69. Upadhaya S, Yu JX, Oliva C, Hooton M, Hodge J, Hubbard-Lucey VM. Impact of COVID-19 on oncology clinical trials. *Nat Rev Drug Discov* 2020;**19**(6):376–7.

70. Patel SS, Webster RK, Greenberg N, Weston D, Brooks SK. Research fatigue in COVID-19 pandemic and post-disaster research: causes, consequences and recommendations. *Disaster Prev Manag* 2020;**29**(4):445–5.

71. Bradt J. Impact of COVID-19 on clinical research. *Nord J Music Ther* 2020;**29**(4):297–9.

72. Slauson-Blevins K, Johnson KM. Doing gender, doing surveys? Women's gatekeeping and men's non-participation in multi-actor reproductive surveys. *Sociol Inq* 2016;**86**(3):427–49.

73. Rourke DO, Lakner E. Gender bias: analysis of factors causing male underrepresentation in surveys. *Int J Public Opin Res* 1989;**1**:164–76.

74. Segre D. What you should remember in managing pilonidal disease. *Front Surg* 2021;**8**:792121.

75. Graham K, Grewal I, Lewis J. *Ethics in Social Research: The Views of Research Participants*. London; 2007.

76. Carson P, Hong CJ, Otero-Vinas M, Arsenault EF, Falanga V. Liver enzymes and lipid levels in patients with lipodermatosclerosis and venous ulcers treated with a prototypic anabolic steroid (stanozolol). *Int J Low Extrem Wounds* 2015;**14**(1):11–8.

77. Yamauchi K, Nakao M, Nakashima M. Correlates of regret with treatment decision-making among Japanese women with breast cancer: results of an internet-based cross-sectional survey. *BMC Womens Health* 2019;**19**(1):86.

78. Wilson A, Ronnekleiv-Kelly SM, Pawlik TM. Regret in surgical decision making: a systematic review of patient and physician perspectives. *World J Surg* 2017;**41**(6):1454–65.

79. Becerra Pérez MM, Menear M, Brehaut JC, Légaré F. Extent and predictors of decision regret about health care decisions. *Med Decis Mak* 2016;**36**(6):777–90.

80. Kannan S, Seo J, Riggs KR, Geller G, Boss ZDB EF. Surgeons' views on shared decision-making. *J Patient-Centered Res Rev* 2020;**7**:8–18.

81. Shelton RC, Brotzman LE, Crookes DM, Robles P, Neugut AI. Decision-making under clinical uncertainty: An in-depth examination of provider perspectives on adjuvant chemotherapy for stage II colon cancer. *Patient Educ Couns* 2019;**102**(2):284–90.

82. Gaston CM, Mitchell G. Information giving and decision-making in patients with advanced cancer: A systematic review. *Soc Sci Med* 2005;**61**(10):2252–64.

83. Niburski K, Guadagno E, Abbasgholizadeh-Rahimi S, Poenaru D. Shared decision making in surgery: a meta-analysis of existing literature. *Patient* 2020;**13**(6):667–81.

84. Harrison JD, Young JM, Price MA, Butow PN, Solomon MJ. What are the unmet supportive care needs of people with cancer? A systematic review. *Support Care Cancer* 2009;**17**(8):1117–28.

85. Raybould G, Babatunde O, Evans AL, Jordan JL, Paskins Z. Expressed information needs of patients with osteoporosis and/or fragility fractures: a systematic review. *Arch Osteoporos* 2018;**13**(1):55.

86. Adams E, Boulton M, Watson E. The information needs of partners and family members of cancer patients: a systematic literature review. *Patient Educ Couns* 2009;**77**(2):179–86.

87. Wang T, Molassiotis A, Chung BPM, Tan J-Y. Unmet care needs of advanced cancer patients and their informal caregivers: a systematic review. *BMC Palliat Care* 2018;**17**(1):96.

88. Williams D, Edwards A, Wood F, Lloyd A, Brain K, Thomas N, *et al.* Ability of observer and self-report measures to capture shared decision-making in clinical practice in the UK: a mixed-methods study. *BMJ Open* 2019;**9**(8):e029485.

89. Heen AF, Vandvik PO, Brandt L, Montori VM, Lytvyn L, Guyatt G, *et al.* A framework for practical issues was developed to inform shared decision-making tools and clinical guidelines. *J Clin Epidemiol* 2021;**129**:104–13.

90. Lagha E, Noble A, Smith A, Denvir M, Leslie S. Patient reported experience measures (PREMs) in chronic heart failure. *J R Coll Physicians Edinb* 2012;**42**(4):301–5.

91. Saunders CL, Elliott MN, Lyratzopoulos G, Abel GA. Do differential response rates to patient surveys between organizations lead to unfair performance comparisons? *Med Care* 2016;**54**(1):45–54.

92. Ahmed F, Burt J, Roland M, Ahmed F, Burt J, Roland AM, *et al.* Measuring patient experience: concepts and methods. *Patient* 2014;**7**:235–41.

93. Cabitza F, Dui LG, Banfi G. PROs in the wild: assessing the validity of patient reported outcomes in an electronic registry. *Comput Methods Programs Biomed* 2019;**181**:104837.

94. Jenkinson C. The picker patient experience questionnaire: development and validation using data from in-patient surveys in five countries. *Int J Qual Heal Care* 2002;**14**(5):353–8.

95. Trujols J, Iraurgi I, Oviedo-Joekes E, Guàrdia J. A critical analysis of user satisfaction surveys in addiction services: opioid maintenance treatment as a representative case study. *Patient Prefer Adherence* 2014;**8**:107–17.

96. Moffatt S, White M, Mackintosh J, Howel D. Using quantitative and qualitative data in health services research – what happens when mixed method findings conflict? *BMC Health Serv Res* 2006;**6**(1):28.

97. Munafò MR, Davey Smith G. Robust research needs many lines of evidence. *Nature* 2018;**553**(7689):399–401.

98. McCaughan D, Sheard L, Cullum N, Dumville J, Chetter I. Nurses' and surgeons' views and experiences of surgical wounds healing by secondary intention: a qualitative study. *J Clin Nurs* 2020;**29**(13–14):2557–71.

99. Bekker CL, Mohsenian Naghani S, Natsch S, Wartenberg NS, van den Bemt BJF. Information needs and patient perceptions of the quality of medication information available in hospitals: a mixed method study. *Int J Clin Pharm* 2020;**42**(6):1396–404.

100. McHugh GA, Luker KA. Individuals' expectations and challenges following total hip replacement: a qualitative study. *Disabil Rehabil* 2012;**34**(16):1351–7.

101. Baumeister RF, Vohs KD, Oettingen G. Pragmatic prospection: how and why people think about the future. *Rev Gen Psychol* 2016;**20**(1):3–16.

102. Mahomed NN, Liang MH, Cook EF, Daltroy LH, Fortin PR, Fossel AH, Katz JN. The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. *J Rheumatol* 2002;**29**(6):1273–9.

103. Purkiss SF. Decision making in surgery: a pilonidal sinus. *Br J Hosp Med* 1993;**50**(9):554–6.

104. Devereaux PJ, Bhandari M, Clarke M, Montori VM, Cook DJ, Yusuf S, *et al.* Need for expertise based randomised controlled trials. *BMJ* 2005;**330**(7482):88.

105. Mottram A. 'They are marvellous with you whilst you are in but the aftercare is rubbish': a grounded theory study of patients' and their carers' experiences after discharge following day surgery. *J Clin Nurs* 2011;**20**(21–22):3143–51.

106. Aune E, Struksnes S. Home care nurses' experience of providing health-care to patients with hard-to-heal wounds. *J Wound Care* 2019;**28**(3):178–87.

107. Friman A, Klang B, Ebbeskog B. Wound care by district nurses at primary healthcare centres: a challenging task without authority or resources. *Scand J Caring Sci* 2011;**25**(3):426–34.

108. Maybin J, Charles A, Honeyman M. *Understanding Quality in District Nursing Services: Learning from Patients, Carers and Staff*. The Kings Fund; 2016. URL: https://www.kingsfund.org.uk/insight-and-analysis/reports/understanding-quality-district-nursing-services (accessed 23 September 2023).

109. Wickramasekera N, Strong E, Shackley P, Callaghan T, Lee M, Hind D, *et al.* Patient preferences for pilonidal sinus treatments: a discrete choice experiment survey. *Colorect Dis* 2023;**25**: 984–994. https://doi.org/10.1111/codi.16482

110. Beal EM, Lee MJ, Hind D, Wysocki AP, Yang F, Brown SR. A systematic review of classification systems for pilonidal sinus. *Tech Coloproctol* 2019;**23**(5):435–43.

111. Hahn G, Shapiro S. *A Catalogue and Computer Program for the Design and Analysis of Orthogonal Symmetric and Asymmetric Fractional Factorial Designs*. Schenectady, New York: General Electric Research and Development Center; 1966.

112. de Bekker-Grob EW, Donkers B, Jonker MF, Stolk EA. Sample size requirements for discrete-choice experiments in healthcare: a practical guide. *Patient* 2015;**8**(5):373–84.

113. Orme B. *Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research*. 2nd ed. Madison: Research Publishers; 2010.

114. Hole AR. A comparison of approaches to estimating confidence intervals for willingness to pay measures. *Health Econ* 2007;**16**(8):827–40.

115. Stewart AM, Baker JD, Elliott D. The effects of a sacrococcygeal pilonidal sinus wound on activities of living: thematic analysis of participant interviews. *J Clin Nurs* 2011;**20**(21–22):3174–82.

116. Harris C, Sibbald RG, Mufti A, Somayaji R. Pilonidal sinus disease: 10 steps to optimize care. *Adv Skin Wound Care* 2016;**29**(10):469–78.

117. Goodson N, Wicks P, Morgan J, Hashem L, Callinan S, Reites J. Opportunities and counterintuitive challenges for decentralized clinical trials to broaden participant inclusion. *Digit Med* 2022;**5**(1):58.

118. Rolfe G, Freshwater D, Jasper M. *Critical Reflection for Nursing and the Helping Professions: A User's Guide*. London: Palgrave; 2001.

119. Kumar M, Clay WH, Lee MJ, Brown SR, Hind D. A mapping review of sacrococcygeal pilonidal sinus disease. *Tech Coloproctol* 2021;**25**(6):675–82.

120. Horton R. Surgical research or comic opera: questions, but few answers. *Lancet* 1996;**347**(9007):984–5.

121. Hinchey EJ, Schaal PG, Richards GK. Treatment of perforated diverticular disease of the colon. *Adv Surg* 1978;**12**:85–109.

122. Kozar RA, Crandall M, Shanmuganathan K, Zarzaur BL, Coburn M, Cribari C, *et al.* Organ injury scaling 2018 update: spleen, liver, and kidney. *J Trauma Acute Care Surg* 2018;**85**(6):1119–22.

123. Park JJH, Siden E, Zoratti MJ, Dron L, Harari O, Singer J, *et al.* Systematic review of basket trials, umbrella trials, and platform trials: a landscape analysis of master protocols. *Trials* 2019;**20**(1):572.

124. Jayne DG, Scholefield J, Tolan D, Gray R, Senapati A, Hulme CT, *et al.* A multicenter randomized controlled trial comparing safety, efficacy, and cost-effectiveness of the Surgisis anal fistula plug versus surgeon's preference for transsphincteric fistula-in-ano: the FIAT trial. *Ann Surg* 2021;**273**(3):433–41.

125. Girling C, Lee MJ, Vimalchandran D, Jayne DJ, Stancliffe S, Wailoo A, *et al.* Protocol for the ORION trial (RadiO fRequency ablatION for haemorrhoids): a randomised controlled trial. *Tech Coloproctol* 2023;**27**(2):117–24.

126. Webbe J, Sinha I, Gale C. Core outcome sets. *Arch Dis Child Educ Pract Ed* 2018;**103**(3):163–6.

127. Tripartite Gastrointestinal Recovery SBO Group. A core outcome set for clinical studies of adhesive small bowel obstruction. *Colorectal Dis* 2022;**24**(10):1204–10.

128. Lapsekill E, Coskun M, Oztas M, Urkan M, Can MF. A classification proposal for the sacrococcygeal pilonidal sinus disease (SPSD). *Eur Surg Res* 2013;**50**(Suppl 1):144.

129. Irkörücü O. Management for pilonidal disease: before you compare, use a classification system. *Asian J Surg* 2016;**39**(4):260–1.

130. Irkorucu O, Erdem H, Reyhan E. The best therapy for pilonidal disease: which management for which type? *World J Surg* 2012;**36**(3):691–2.

131. Quinodoz PD, Chilcott M, Grolleau JL, Chavoin JP, Costagliola M. Surgical treatment of sacrococcygeal pilonidal sinus disease by excision and skin flaps: the Toulouse experience. *Eur J Surg* 1999;**165**(11):1061–5.

132. Doll D, Vassiliu P. Another pilonidal classification—PLLATIN. *Pilonidal Sinus J* 2018;**4**(1):1–3.

133. Guner A, Cekic AB, Boz A, Turkyilmaz S, Kucuktulu U. A proposed staging system for chronic symptomatic pilonidal sinus disease and results in patients treated with stage-based approach. *BMC Surg* 2016;**16**(1):18.

134. Tezel E. A new classification according to navicular area concept for sacrococcygeal pilonidal disease. *Color Dis* 2007;**9**(6):575–6.

135. Awad M, Elbaset A, Ebraheem S, Tantawy E, Elhafez M, Elsayed A. A scoring system as a method to evaluate pilonidal sinus disease to make an easy decision for its management. *Indian J Plast Surg* 2009;**42**(1):43.

136. Wysocki AP, Andersson RE, Gips M, Girgin M, Guner A, Immerman S, *et al.* Towards a classification for sacrococcygeal pilonidal disease – Berlin 2017. *Pilonidal Sinus J* 2018;**4**(1):5–12.

137. International Pilonidal Society. Survey towards pilonidal classification. *Pilonidal Sinus J* 2018;**4**(1):5–12.

138. Flight L, Julious SA. The disagreeable behaviour of the kappa statistic. *Pharm Stat* 2015;**14**(1):74–8.

139. StataCorp. *Stata Statistical Software: Release 17*. College Station, TX: StataCorp LLC; 2021.

140. Rotondi MA, Donner A. A confidence interval approach to sample size estimation for interobserver agreement studies with multiple raters and outcomes. *J Clin Epidemiol* 2012;**65**(7):778–84.

141. Dekker L, Han-Geurts IJM, Grossi U, Gallo G, Veldkamp R. Is the Goligher classification a valid tool in clinical practice and research for hemorrhoidal disease? *Tech Coloproctol* 2022;**26**(5):387–92.

142. Jensen P, Krogsgaard MR, Christiansen J, Braendstrup O, Johansen A, Olsen J. Observer variability in the assessment of type and dysplasia of colorectal adenomas, analyzed using kappa statistics. *Dis Colon Rectum* 1995;**38**(2):195–8.

143. Ergina PL, Barkun JS, McCulloch P, Cook JA, Altman DG. IDEAL Group IDEAL framework for surgical innovation 2: observational studies in the exploration and assessment stages. *BMJ* 2013;**346**:f3011.

144. Vohra RS, Spreadborough P, Johnstone M, Marriott P, Bhangu A, Alderson D, *et al.* Protocol for a multicentre, prospective, population-based cohort study of variation in practice of cholecystectomy and surgical outcomes (the CholeS study). *BMJ Open* 2015;**5**(1):e006399.

145. Lee MJ, Sayers AE, Drake TM, Marriott PJ, Anderson ID, Bach SP, *et al.* National prospective cohort study of the burden of acute small bowel obstruction. *BJS Open* 2019;**3**(3):354–66.

146. Glancz LJ, Poon MTC, Coulter IC, Hutchinson PJ, Kolias AG, Brennan PM. Does drain position and duration influence outcomes in patients undergoing burr-hole evacuation of chronic subdural hematoma? Lessons from a UK multicenter prospective cohort study. *Neurosurgery* 2019;**85**(4):486–93.

147. Milone M, Basso L, Manigrasso M, Pietroletti R, Bondurri A, La Torre M, *et al.* Consensus statement of the Italian society of colorectal surgery (SICCR): management and treatment of pilonidal disease. *Tech Coloproctol* 2021;**25**(12):1269–80.

148. Segre D, Pozzo M, Perinotti R, Roche B, Italian Society of Colorectal Surgery. The treatment of pilonidal disease: guidelines of the Italian Society of Colorectal Surgery (SICCR). *Tech Coloproctol* 2015;**19**(10):607–13.

149. Beal E, Hind D, Bradburn M, Lee E, Howard A, Shackley P, *et al.* Design and rationale of the PIlonidal sinus Treatment – STudying the OPtions (PITSTOP) study: a multicentre cohort, nested mixed-methods case study and discrete choice experiment (poster presentation). *Int J Surg* 2018;**59**:S6.

Appendix 1 PITSTOP Project Management Group

Khalafalla Ali, Mike Bradburn, Richard Brady, Graham Branagan, Steven Brown, Sanjay Chaudri, Francesco Di Fabio, Godwin Dennison, Farhat Din, David Donnelly, Martyn Evans, Francois Gerald, Sarah Gonzalez, Jennie Grainger, Alex Hardy, Mohan Harilingam, Daniel Hind, Philip Hopley, Najam Husain, Helen Jones, Sandeep Kapur, Kenneth Keogh, Ellen Lee, Matt Lee, Michael Lim, Jon Lund, Paul Mackey, Yasuko Maeda, Sanjay Mahaptra, Sudhaker Mangam, Felix Mazarelo, Christine Moffatt, Jon Morton, Karim Muhammad, Nikhill Pawa, Lyndsay Pearce, James Pitt, Raj Rajaganeshan, Asha Senapati, Phil Shackley, Richard Simmonds, Richard Stevenon, Jared Torkington, Peter Vaughan-Shaw, Dale Vimalachandran, Jeremy Wilson, Peter Wysocki.

Appendix 2 Cohort study participating sites

Countess of Chester Hospital

Sheffield Teaching Hospitals

Wirral University Teaching Hospital

University Hospital of Wales – Cardiff

Norfolk and Norwich University Hospitals

Oxford University Hospital

St Mark's Hospital London

Glasgow Royal Infirmary

Queen Alexandra Hospital – Portsmouth

Addenbrookes Hospital – Cambridge

Royal Derby Hospital

Western General Hospital – Edinburgh

Burton Hospital

Tameside and Glossop Integrated Care NHS Foundation Trust

Newcastle Upon Tyne Hospitals NHS Foundation Trust

Manchester Royal Infirmary

York Teaching Hospital NHS Foundation Trust

St Helens and Knowsley Teaching Hospitals NHS Trust

Swansea Bay University Health Board – Morriston Hospital

Salford Royal NHS Foundation Trust

Yeovil District Hospital NHS Foundation Trust

Queen Elizabeth The Queen Mother Hospital – East Kent

Royal Devon and Exeter NHS Foundation Trust

East Suffolk and North Essex NHS Foundation Trust – Ipswich

Musgrove Park Hospital – Taunton

Salisbury NHS Foundation Trust

Leicester General Hospital

Trafford General Hospital

Peterborough City Hospital

Hinchingbrooke Hospital

Chelsea and Westminster Hospital NHS Foundation Trust

Appendix 3 Additional figures and tables

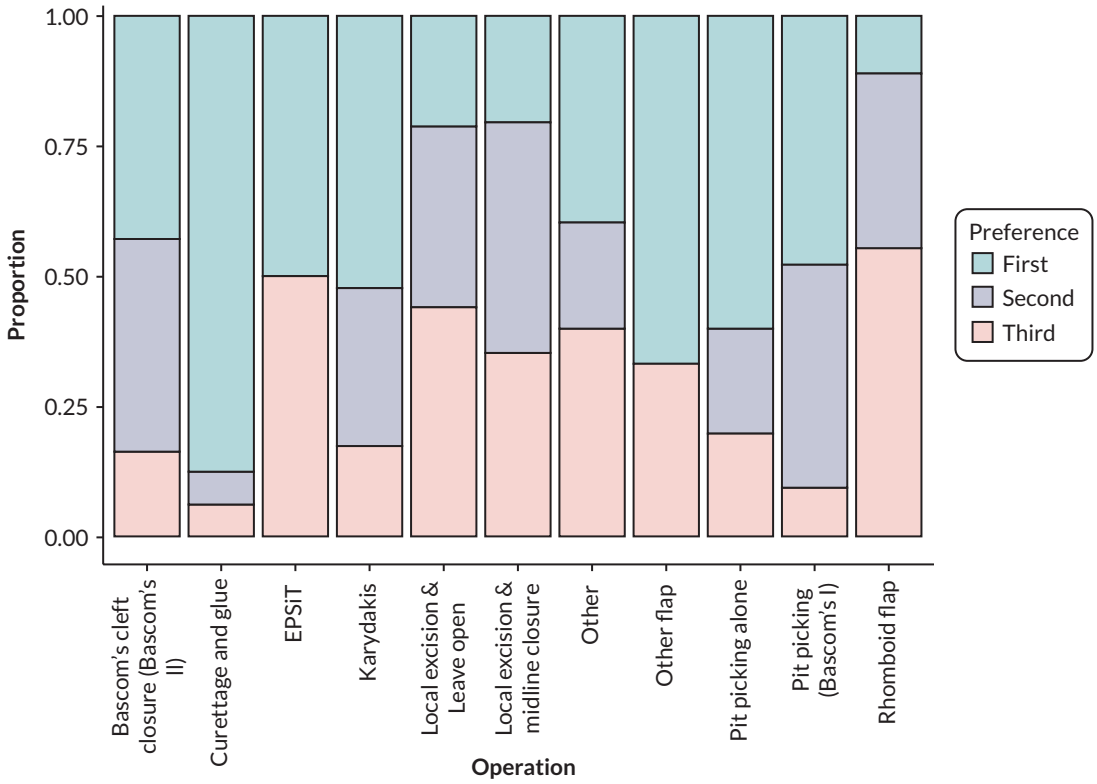


FIGURE 12 Procedure preferences of responding surgeons.

TABLE 25 Outcome completion rates

Outcome, *n* (%) with available data	Time point					
	All	Baseline	Day 1	Day 7	Clinic visit	6 months
Any outcome data		667 (100%)	608 (91%)	577 (87%)	513 (77%)	476 (71%)
Complications			608 (91%)	576 (86%)	510 (76%)	474 (71%)
Pain (today)		666 (100%)	606 (91%)	574 (86%)	501 (75%)	470 (70%)
Pain (worst in last week)		665 (100%)		574 (86%)	502 (75%)	470 (70%)
EQ-5D						
EQ-5D-5L health utility (crosswalk)		654 (98%)		572 (86%)	494 (74%)	466 (70%)
EQ-5D – your health today		658 (99%)		572 (86%)	493 (74%)	466 (70%)
CWIQ						
Physical symptoms and daily living experience					497 (75%)	467 (70%)

continued

TABLE 25 Outcome completion rates (*continued*)

Outcome, *n* (%) with available data	Time point					
	All	Baseline	Day 1	Day 7	Clinic visit	6 months
Physical symptoms and daily living stress					496 (74%)	461 (69%)
Well-being					495 (74%)	460 (69%)
QoL					495 (74%)	465 (70%)
QoL satisfaction					495 (74%)	464 (70%)
Repacking procedures				563 (84%)	496 (74%)	458 (69%)
Replacement/removal of dressing				561 (84%)	486 (73%)	447 (67%)
Service interactions				572 (86%)	501 (75%)	471 (71%)
Returned to normal activities	607 (91%)					
Wound healed	553 (83%)					
Any reported recurrence	629 (94%)					
Any complication during follow-up	643 (96%)					
DR						456 (68%)
Scar spread					246 (37%)	
Scar overall impression					241 (36%)	
Scar itch (in past 24 hours)					412 (62%)	
Scar pain (in past 24 hours)					412 (62%)	

TABLE 26 Procedure information

Procedure information	Asymmetric closure (*n* = 272)	Leave open (*n* = 49)	Midline closure (*n* = 76)	Minimal excision (*n* = 270)	All (*n* = 667)
Length of surgery (minutes)					
N (%)	271 (100%)	47 (96%)	74 (97%)	261 (97%)	653 (98%)
Mean (SD)	47.0 (22.3)	25.2 (14.7)	40.5 (18.8)	19.0 (15.9)	33.5 (23.1)
Median (IQR)	45 (30, 60)	20 (15, 33)	36 (27, 50)	15 (9, 24)	30 (16, 45)
Min, max	10, 171	5, 67	13, 105	2, 136	2, 171
Category of hospital stay					
Day case	251 (92%)	45 (92%)	73 (96%)	266 (99%)	635 (95%)
Inpatient	20 (7%)	4 (8%)	3 (4%)	2 (1%)	29 (4%)
Grade of operating surgeon					
Consultant	172 (63%)	38 (78%)	46 (61%)	199 (74%)	455 (68%)
Non-consultant	100 (37%)	11 (22%)	30 (39%)	71 (26%)	212 (32%)

TABLE 26 Procedure information (*continued*)

Procedure information	Asymmetric closure (*n* = 272)	Leave open (*n* = 49)	Midline closure (*n* = 76)	Minimal excision (*n* = 270)	All (*n* = 667)
Type of anaesthetic during operation					
General	227 (83%)	41 (84%)	74 (97%)	179 (66%)	521 (78%)
Spinal	11 (4%)	4 (8%)	0 (0%)	5 (2%)	20 (3%)
Local	34 (13%)	3 (6%)	2 (3%)	85 (31%)	124 (19%)
Sedation (for those using local anaesthetic)					
No	7 (3%)	1 (2%)	1 (1%)	56 (21%)	65 (10%)
Yes	27 (10%)	1 (2%)	1 (1%)	26 (10%)	55 (8%)
Antibiotics used at induction	232 (85%)	21 (43%)	42 (55%)	108 (40%)	403 (60%)
Antibiotics used post surgery	88 (32%)	1 (2%)	17 (22%)	17 (6%)	123 (18%)

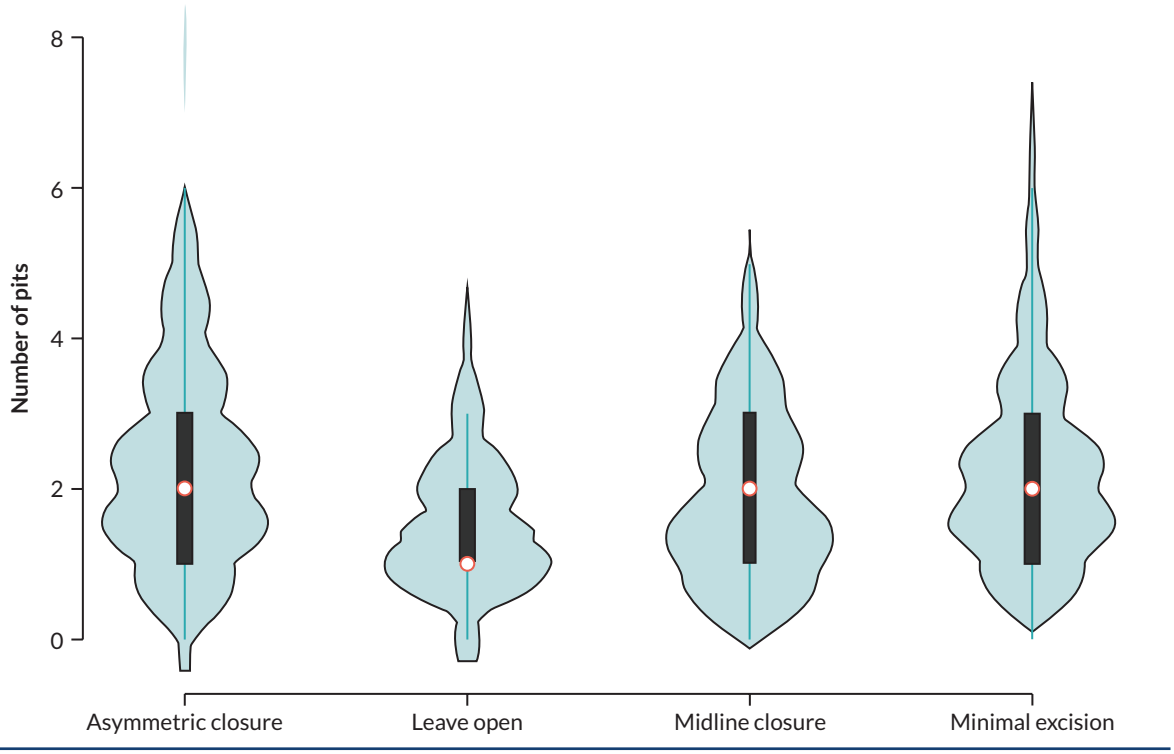


FIGURE 13 Treatment choice by number of pits (*N* = 640).

TABLE 27 Continuous outcomes measured at multiple time points

	Asymmetric closure (*n* = 272)		Leave open (*n* = 49)		Midline closure (*n* = 76)		Minimal excision (*n* = 270)		All (*n* = 667)	
Outcome	***n***	**Mean (SD)**	***n***	**Mean (SD)**	***n***	**Mean (SD)**	***n***	**Mean (SD)**	***n***	**Mean (SD)**
Pain (today)[a]										
Baseline	271	1.9 (2.2)	49	2.9 (2.6)	76	1.6 (2.0)	270	1.8 (2.3)	666	1.9 (2.3)
Day 1	253	4.5 (2.4)	43	2.9 (2.6)	68	3.9 (2.8)	242	2.6 (2.2)	606	3.6 (2.5)
Day 7	246	3.3 (2.3)	41	4.1 (2.8)	66	3.7 (2.8)	221	1.9 (2.2)	574	2.8 (2.5)
Clinic visit	218	1.6 (2.1)	34	1.8 (2.4)	51	2.3 (2.5)	198	0.8 (1.7)	501	1.4 (2.1)
6-month visit	202	0.6 (1.5)	34	1.4 (2.2)	54	0.7 (1.9)	180	0.8 (1.6)	470	0.7 (1.6)
Pain (worst in last week)[a]										
Baseline	270	3.6 (3.0)	49	4.5 (3.2)	76	3.4 (3.2)	270	3.3 (3.0)	665	3.5 (3.0)
Day 7	246	5.4 (2.5)	41	5.9 (2.9)	66	5.5 (3.1)	221	3.0 (2.8)	574	4.5 (3.0)
Clinic visit	218	2.7 (2.8)	34	3.0 (2.6)	51	3.4 (2.9)	199	1.5 (2.5)	502	2.3 (2.8)
6-month visit	202	1.0 (2.0)	34	2.2 (2.7)	54	1.1 (2.3)	180	1.4 (2.5)	470	1.3 (2.3)
EQ-5D-5L health utility (crosswalk)[b]										
Baseline	267	0.79 (0.20)	48	0.76 (0.19)	74	0.81 (0.20)	265	0.82 (0.19)	654	0.80 (0.20)
Day 7	246	0.65 (0.21)	41	0.60 (0.22)	66	0.61 (0.27)	219	0.79 (0.20)	572	0.69 (0.23)
Clinic visit	214	0.80 (0.21)	34	0.75 (0.22)	49	0.75 (0.20)	197	0.89 (0.17)	494	0.83 (0.20)
6-month visit	201	0.90 (0.18)	34	0.82 (0.19)	53	0.89 (0.19)	178	0.89 (0.16)	466	0.89 (0.17)
EQ-5D – your health today[c]										
Baseline	268	76.6 (16.1)	48	74.4 (18.2)	73	76.1 (12.9)	269	77.4 (16.8)	658	76.7 (16.2)
Day 7	247	73.5 (16.1)	41	71.1 (19.6)	64	71.5 (19.4)	220	79.4 (17.7)	572	75.4 (17.6)
Clinic visit	214	78.7 (16.9)	34	81.1 (15.0)	49	80.4 (14.3)	196	85.3 (13.7)	493	81.6 (15.6)
6-month visit	202	81.0 (19.8)	34	79.4 (19.6)	53	82.5 (12.9)	177	84.0 (16.2)	466	82.2 (17.8)

TABLE 27 Continous outcomes measured at multiple time points (*continued*)

	Asymmetric closure		Leave open		Midline closure		Minimal excision		All	
	(*n* = 272)		(*n* = 49)		(*n* = 76)		(*n* = 270)		(*n* = 667)	
Outcome	*n*	Mean (SD)	*n*	Mean (SD)	*n*	Mean (SD)	*n*	Mean (SD)	*n*	Mean (SD)
CWIQ QoL[d]										
Clinic visit	216	7.5 (2.1)	34	7.6 (1.7)	48	7.5 (1.8)	197	8.6 (1.7)	495	7.9 (1.9)
6-month visit	201	8.4 (1.7)	34	7.6 (2.4)	53	8.0 (2.0)	177	8.6 (1.5)	465	8.4 (1.7)
CWIQ QoL satisfaction[d]										
Clinic visit	216	7.4 (2.2)	34	7.5 (2.0)	48	7.8 (1.7)	197	8.6 (1.7)	495	7.9 (2.0)
6-month visit	201	8.3 (1.9)	34	7.5 (2.6)	53	8.1 (2.1)	176	8.5 (1.9)	464	8.3 (2.0)
CWIQ Physical symptoms and daily living experience[e]										
Clinic visit	216	81.2 (22.0)	34	76.5 (21.4)	49	75.1 (22.1)	198	91.8 (15.0)	497	84.5 (20.4)
6-month visit	202	92.7 (14.8)	34	85.2 (22.0)	54	92.5 (14.3)	177	92.7 (12.3)	467	92.1 (14.6)
CWIQ Physical symptoms and daily living stress[e]										
Clinic visit	216	85.0 (22.2)	34	82.1 (21.9)	49	81.2 (20.8)	197	94.9 (12.9)	496	88.4 (19.6)
6-month visit	197	94.8 (14.6)	34	88.3 (18.2)	53	94.8 (12.3)	177	95.7 (9.5)	461	94.7 (13.0)
CWIQ Well-being[e]										
Clinic visit	216	57.9 (23.3)	34	54.1 (24.2)	49	58.6 (22.6)	196	68.7 (22.5)	495	62.0 (23.6)
6-month visit	199	68.5 (22.4)	34	56.5 (22.0)	52	66.7 (24.3)	175	68.1 (22.8)	460	67.3 (22.9)

a Self-reported pain related to pilonidal sinus ranges from 0 (no pain) to 10 (worst imaginable pain).
b EQ-5D-5L ranges from −0.22 to 1; higher scores represent better health.
c EQ-5D your health today ranges from 0 (worst health) to 100 (best health).
d CWIQ QoL and QoL satisfaction range from 0 to 10; higher scores represent better QoL/satisfaction.
e CWIQ scores range from 0 to 100, higher scores indicate greater impact.

TABLE 28 Repacking and re-dressing procedures during follow-up: characteristics

Outcome	Asymmetric closure (*n* = 272)	Leave open (*n* = 49)	Midline closure (*n* = 76)	Minimal excision (*n* = 270)	All (*n* = 667)
Repacking procedure since last follow-up					
Day 1	0/254 (0%)	0/43 (0%)	0/69(0%)	0/242 (0%)	0/608 (0%)
Day 7	16/248 (6%)	25/41 (61%)	5/67(7%)	22/221 (10%)	68/577 (12%)
Clinic visit	37/222 (17%)	21/35 (60%)	14/53(26%)	15/203 (7%)	87/513 (17%)
6-month visit	22/205 (11%)	11/35 (31%)	6/56(11%)	6/180 (3%)	45/476 (9%)
Re-dressing procedure since last follow-up					
Day 1	0/254 (0%)	0/43 (0%)	0/69(0%)	0/242 (0%)	0/608 (0%)
Day 7	122/248 (49%)	18/41 (44%)	35/67(52%)	51/221 (23%)	226/577 (39%)
Clinic visit	105/222 (47%)	16/35 (46%)	21/53(40%)	37/203 (18%)	179/513 (35%)
6-month visit	36/205 (18%)	8/35 (23%)	12/56(21%)	11/180 (6%)	67/476 (14%)

TABLE 29 Postoperative complications during follow-up

Time point	Complication	Asymmetric closure (*n* = 272)	Leave open (*n* = 49)	Midline closure (*n* = 76)	Minimal excision (*n* = 270)	All (*n* = 667)
During follow-up	Any complication	135/272 (50%)	26/49 (53%)	46/76 (61%)	94/270 (35%)	301/667 (45%)
	Bleeding	49/272 (18%)	14/49 (29%)	16/76 (21%)	15/270 (6%)	94/667 (14%)
	Dehiscence	46/272 (17%)	2/49 (4%)	17/76 (22%)	8/270 (3%)	73/667 (11%)
	Discharge	44/272 (16%)	12/49 (24%)	17/76 (22%)	46/270 (17%)	119/667 (18%)
	Seroma	11/272 (4%)	0/49 (0%)	5/76 (7%)	3/270 (1%)	19/667 (3%)
	Infection	83/272 (31%)	15/49 (31%)	26/76 (34%)	51/270 (19%)	175/667 (26%)
Day 1	Any complication	17/254 (7%)	8/44 (18%)	10/69 (14%)	7/242 (3%)	42/609 (7%)
	Bleeding	8/254 (3%)	5/44 (11%)	5/69 (7%)	3/242 (1%)	21/609 (3%)
	Dehiscence	1/254 (0%)	0/44 (0%)	1/69 (1%)	0/242 (0%)	2/609 (0%)
	Discharge	3/254 (1%)	3/44 (7%)	1/69 (1%)	4/242 (2%)	11/609 (2%)
	Seroma	1/254 (0%)	0/44 (0%)	2/69 (3%)	0/242 (0%)	3/609 (0%)
	Infection	3/254 (1%)	0/44 (0%)	0/69 (0%)	0/242 (0%)	3/609 (0%)
Day 7	Any complication	47/248 (19%)	9/42 (21%)	21/67 (31%)	30/220 (14%)	107/577 (19%)
	Bleeding	15/248 (6%)	4/42 (10%)	5/67 (7%)	4/220 (2%)	28/577 (5%)
	Dehiscence	9/248 (4%)	0/42 (0%)	4/67 (6%)	2/220 (1%)	15/577 (3%)
	Discharge	10/248 (4%)	2/42 (5%)	5/67 (7%)	11/220 (5%)	28/577 (5%)
	Seroma	0/248 (0%)	0/42 (0%)	2/67 (3%)	0/220 (0%)	2/577 (0%)
	Infection	20/248 (8%)	4/42 (10%)	9/67 (13%)	16/220 (7%)	49/577 (8%)
Clinic visit	Any complication	100/221 (45%)	12/36 (33%)	30/54 (56%)	48/202 (24%)	190/513 (37%)
	Bleeding	29/221 (13%)	3/36 (8%)	12/54 (22%)	6/202 (3%)	50/513 (10%)

continued

TABLE 29 Postoperative complications during follow-up (*continued*)

Time point	Complication	Asymmetric closure (*n* = 272)	Leave open (*n* = 49)	Midline closure (*n* = 76)	Minimal excision (*n* = 270)	All (*n* = 667)
	Dehiscence	35/221 (16%)	1/36 (3%)	11/54 (20%)	3/202 (1%)	50/513 (10%)
	Discharge	27/221 (12%)	6/36 (17%)	8/54 (15%)	20/202 (10%)	61/513 (12%)
	Seroma	9/221 (4%)	0/36 (0%)	3/54 (6%)	2/202 (1%)	14/513 (3%)
	Infection	61/221 (28%)	10/36 (28%)	15/54 (28%)	26/202 (13%)	112/513 (22%)
6-month visit	Any complication	61/204 (30%)	13/36 (36%)	20/56 (36%)	42/179 (23%)	136/475 (29%)
	Bleeding	16/204 (8%)	4/36 (11%)	6/56 (11%)	4/179 (2%)	30/475 (6%)
	Dehiscence	27/204 (13%)	1/36 (3%)	5/56 (9%)	4/179 (2%)	37/475 (8%)
	Discharge	12/204 (6%)	6/36 (17%)	8/56 (14%)	25/179 (14%)	51/475 (11%)
	Seroma	4/204 (2%)	0/36 (0%)	2/56 (4%)	1/179 (1%)	7/475 (1%)
	Infection	26/204 (13%)	4/36 (11%)	7/56 (13%)	16/179 (9%)	53/475 (11%)

TABLE 30 Outcomes measured once during follow-up and outcomes incorporating the full follow-up period

Characteristic	Asymmetric closure (*n* = 272)	Leave open (*n* = 49)	Midline closure (*n* = 76)	Minimal excision (*n* = 270)	All (*n* = 667)
Scar spread, *N*	130	14	25	77	246
None to near-invisible	13 (10%)	3 (21%)	7 (28%)	28 (36%)	51 (21%)
Pencil-thin line	40 (31%)	0 (0%)	5 (20%)	20 (26%)	65 (26%)
Mild spread, noticeable on close inspection	49 (38%)	6 (43%)	6 (24%)	18 (23%)	79 (32%)
Moderate spread, obvious scarring	24 (18%)	4 (29%)	5 (20%)	11 (14%)	44 (18%)
Severe spread	4 (3%)	1 (7%)	2 (8%)	0 (0%)	7 (3%)
Participant satisfaction, *N*	201	34	51	177	463
Very satisfied	113 (56%)	18 (53%)	21 (41%)	89 (50%)	241 (52%)
Satisfied	61 (30%)	9 (26%)	18 (35%)	54 (31%)	142 (31%)
Neither satisfied nor dissatisfied	15 (7%)	0 (0%)	6 (12%)	19 (11%)	40 (9%)
Dissatisfied	3 (1%)	6 (18%)	6 (12%)	9 (5%)	24 (5%)
Very dissatisfied	9 (4%)	1 (3%)	0 (0%)	6 (3%)	16 (3%)
Scar impression – desirable scar	107/128 (84%)	9/12 (75%)	16/23 (70%)	67/78 (86%)	199/241 (83%)
Scar itch (in past 24 hours)	71/195 (36%)	9/25 (36%)	14/39 (36%)	32/153 (21%)	126/412 (31%)
Scar pain (in past 24 hours)	67/195 (34%)	8/25 (32%)	21/38 (55%)	24/154 (16%)	120/412 (29%)
Returned to normal activities	195/260 (75.0%)	27/44 (61.4%)	48/69 (69.6%)	211/241 (87.6%)	481/614 (78.3%)
Wound healed	176/243 (72.4%)	23/39 (59.0%)	44/64 (68.8%)	167/224 (74.6%)	410/570 (71.9%)
Any complication during follow-up	135/265 (51%)	26/46 (57%)	46/74 (62%)	94/258 (36%)	301/643 (47%)
Any reported recurrence	55/257 (21%)	13/46 (28%)	18/70 (26%)	87/256 (34%)	173/629 (28%)
Recurrence within 6 months	28/226 (12%)	10/44 (23%)	13/67 (19%)	61/229 (27%)	112/566 (20%)
Treatment failure[b]	109/257 (42%)	23/46 (50%)	37/70 (53%)	121/257 (47%)	290/630 (46%)
Recurrence apparent from AE report	12	3	2	16	33

a DR is scored from 0 to 100; higher scores indicate greater regret.
b Treatment failure is defined as having recurred, having not returned to normal activity during follow-up, or wound not healed during follow-up.

TABLE 31 Comparison of pain on day 1 and day 7 between major and minor procedures

Model	Major procedure		Minor procedure		*N*	Mean difference (95% CI)[a]
	n	Mean (SD)	*n*	Mean (SD)		
Pain (day 1)						
Raw difference	364	4.22 (2.53)	242	2.60 (2.24)	606	1.62 (1.23 to 2.02)
Risk-adjusted – Wysocki					601	1.64 (1.24 to 2.05)
Risk-adjusted – chosen model (sex, smoking, Wysocki)					544	1.56 (1.14 to 1.98)
Risk-adjusted – full model					404	1.70 (1.20 to 2.20)
Propensity-adjusted – IPW					591	1.54 (1.08 to 2.00)
Propensity matching					591	1.64 (1.17 to 2.10)
Augmented IPW					536	1.58 (1.14 to 2.01)
Pain (day 7)						
Raw difference	353	3.44 (2.50)	221	1.86 (2.18)	574	1.58 (1.18 to 1.98)
Risk-adjusted – Wysocki					569	1.56 (1.15 to 1.97)
Risk-adjusted – chosen model (lateral distribution, sex, Wysocki)					514	1.45 (1.03 to 1.87)
Risk-adjusted – full model					382	1.47 (0.98 to 1.95)
Propensity-adjusted –IPW					559	1.57 (1.14 to 2.00)
Propensity matching					559	1.65 (1.23 to 2.07)
Augmented IPW					512	1.53 (1.12 to 1.95)

a Reference group: minor procedure. Risk-adjusted models: linear regression model with adjustment for covariates as listed, full model includes adjustment for sex, BMI, natal cleft depth, gluteal hair, smoking status, Wysocki classification, pit density, lateral distance, lateral distribution and presence of pus; propensity-adjusted and propensity matching adjust for sex, Wysocki classification and presence of pus; augmented IPW adjusted for sex, Wysocki classification, lateral distance and presence of pus.

TABLE 32 Recurrence rate for surgeons who operated on at least 10 cohort participants (*N* = 13 surgeons)

N per surgeon with outcome data	Recurrence	Recurrence within 6 months	Treatment failure
	≥ 9	≥ 7	≥ 9
Recurrence, %			
Min, max	0, 61	0, 55	18, 78
Median (IQR)	25 (18–31)	17 (0–27)	39 (34–53)
	Recurrence within 6 months (*N* = 566)	Any reported recurrence (*N* = 629)	Treatment failure (*N* = 630)
Recurrence among surgeons with ≥ 10 procedures [*n* (%)]	45 (40.2%)	73 (42.2%)	116 (40.0%)
Range among 13 surgeons	0–55%	0–61%	18–78%
Recurrence among surgeons with < 10 procedures [*n* (%)]	67 (59.8%)	100 (57.8%)	174 (60.0%)

TABLE 33 Outcomes and reflections ordered by level of DR (0–100, high to low)

Participant information					Decision regret	
ID	Excision	Closure	Time to healing (days)	Pain/post-surgery complications	Score	Sample quote (coding)
16	Local excision	Lateral closure and Karydakis	62	0	0	'Everything was great from that first consultation at the doctors to all the way through my recovery. So yeah, I've not really got anything to change about it.' (CODE: consolidation)
19	Local excision	Primary midline closure, marsupialisation and lateral closure	28	0	0	'I would've done it much earlier. As I say, I waited a very long time, probably 12, 13 years, possibly more!' (CODE: consolidation)
21	Local excision	Midline closure	78	0	0	Follow-up interview not complete
7	Curettage and pit picking	FG	51	0	0	Follow-up interview not complete
17	Local excision	Lateral closure and Karydakis	54	0	0	'I think the first surgery was so quick that I wasn't really able to almost consider what I was getting done … I didn't have any time to think about what was happening so it meant afterwards I didn't really take it seriously enough' (Intervention coherence)
5	Local excision	Lateral closure	60	0	5	'As I say it all, all went well. You know there's, there's no reason for me to want to do anything differently' (Acceptability: perceived effectiveness; CODE: consolidation)
9	Local excision	Lateral closure and Karydakis	Length of time not specified	0	5	'Tried to get it [treatment] sooner' (CODE: consolidation)
10	Seton (no excision)		38	0	5	'[So is there anything that you would have done differently?] No' (Acceptability: perceived effectiveness/ethicality; CODE: consolidation)
11	Curettage	No closure/leave open	112	0	5	'I think surgery was the way to go. I don't think I could have done it differently.' (Acceptability: perceived effectiveness; CODE: consolidation)
1	Local excision	Midline closure	Not healed	2	10	'The end result has been a positive one … I think that I would've rather had been in a position in which the wound had just been left open to be packed … that would've actually caused less pain and discomfort overall as well as avoiding the need to sort of visit the hospital for a follow-up' (Acceptability: perceived effectiveness/opportunity costs)

continued

TABLE 33 Outcomes and reflections ordered by level of DR (0–100, high to low) (*continued*)

Participant information					Decision regret	
ID	Excision	Closure	Time to healing (days)	Pain/post-surgery complications	Score	Sample quote (coding)
3	Pit picking	No closure/leave open	84	1	10	[Is there anything that you would've done differently?] Not really because … it's not a condition that you have knowledge of … if you have tingling in your left hand and you have shortness of breath, you know you're having a heart attack … whereas this is not something you have any knowledge of so (mm) I suppose … you sort of do learn on the job with this sort of condition because it's not that common.' (CODE: consolidation)
14	EPSiT	No closure/leave open	Not healed	3/Discharge	15	'The only thing I could have done is … asked for a different doctor, or…. said it was more urgent, so I could have been got in sooner … I'm pretty convinced that months of waiting around, and getting worse and splitting open my skin is the first problem with why it hasn't healed as well as …' (CODE: consolidation)
8	Local excision	No closure/leave open	49	0	20	'… I did everything like as soon as I could like' (CODE: consolidation)
6	Local excision	No closure/leave open	Not healed	5/Discharge and infection	40	'I don't know what I would do differently but I think the, that is what I did differently to change going from [hospital name] to [hospital name].' (self-efficacy)
18	Pit picking	Pit picking – closed and lateral wound – left open	18	0	50	'I'm glad I waited for the right person and the right procedure.' (CODE: consolidation)
2	Local excision	Lateral closure and Karydakis	LTFU	LTFU	LTFU	Follow-up interview not complete
12	Curettage	FG	14	LTFU	LTFU	Follow-up interview not complete
13	Curettage	FG	LTFU	LTFU	LTFU	Follow-up interview not complete
15	Local excision	Seton and flap (type: fascial)	8	LTFU	LTFU	Follow-up interview not complete
20	Local excision	Flap (type: rhomboid)	LTFU	LTFU	LTFU	Follow-up interview not complete

Note
Pain/post-surgery complications recorded at 6 month follow-up; score – DR score regarding treatment decision recorded at 6-month follow-up – high score = high DR, low score = low DR (0–100) – table orders participants from low to high DR scores; sample quote taken from 6-month follow-up interview.

TABLE 34 Discrete choice experiment modelled preferences

Attributes	Model 1: all attributes categorical Coefficient (SE)	Model 2: risk attribute linear Coefficient (SE)
Constant	0.356*** (0.068)	0.368*** (0.067)
Recovery time		
Week = 12 (reference level)	0.000 (.)	0.000 (.)
Week = 1	1.583*** (0.155)	1.556*** (0.151)
Week = 2	2.054*** (0.154)	2.035*** (0.152)
Week = 6	1.256*** (0.109)	1.250*** (0.109)
Risk of infection/persistence		
Risk (%) = 30 (reference level)	0.000 (.)	–
Risk (%) = 20	1.173*** (0.113)	–
Risk (%) = 10	2.217*** (0.145)	–
Risk (%) = 5	3.042*** (0.160)	–
Risk of infection/persistence as a linear variable		
Risk (%)	-	−0.119*** (0.006)
Observations	3552	3552
Log-likelihood	−768.95	−771.78
BIC	1605.24	1588.45
Attribute importance score: a relative measure of the impact that an attribute has on a respondent's choices within the DCE exercise.		
Risk of infection/persistence	70.10%	
Recovery time	29.90%	

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Note
Positive coefficients show attribute levels that are preferable to patients and negative coefficients indicate attribute levels that decrease the likelihood of choosing a treatment.

TABLE 35 Ratings presented by policy statement within each stakeholder group

Statement	Round 1			Round 2			Round 3		
Policy statement	**Clinicians (%)**	**Patients (%)**	**Both (%)**	**Clinicians (%)**	**Patients (%)**	**Both (%)**	**Clinicians (%)**	**Patients (%)**	**Both (%)**
Any treatment of pilonidal disease should aim to be less disruptive than the disease itself.	82.5	80.0	81.8	To consensus					
Surgeons should have access to opportunities to learn new techniques for the treatment of PSD.	97.5	73.3	90.9	To consensus					
Lay open is associated with slow healing and delayed return to normal activities. It should rarely be considered as the first treatment option.	60.0	60.0	60.0	60.5	76.9	64.7	62.2	78.6	66.7
Minimally invasive techniques should be considered as the first-line intervention, as these are associated with low operative morbidity and comparable recurrence and healing rates to more extensive interventions.	65	86.7	70.9	68.4	84.6	72.5	To consensus		
There is a need for a standard classification system/tool for PSD.	82.5	53.3	74.5	81.6	84.6	82.4	To consensus		
Any classification tool should be easy to use.	92.5	46.7	80.0	To consensus					
A classification tool for pilonidal sinus should help to inform treatment options.	82.5	66.7	78.2	76.3	92.3	80.4	To consensus		
Patients should be counselled about the risk of recurrence.	97.5	80.0	92.7	To consensus					
Patients should be counselled about the impact of treatments on return to normal activities.	95	80.0	90.9	To consensus					
Patients may wish for symptomatic improvement rather than cure, and this should be explored in early discussions.	80.0	53.3	72.7	84.2	69.2	80.4	To consensus		
Clinicians and researchers need to clearly define failure of healing vs. recurrence as the two may present similarly.	57.5	80.0	63.6	73.7	92.3	78.4	To consensus		
Delayed return to work is an important outcome following treatment.	90	73.3	85.5	To consensus					
A tool is needed to measure the impact of treatments/disease on QoL (e.g. a disease-specific PRO measure).	82.5	60.0	76.4	84.2	69.2	80.4	To consensus		

TABLE 35 Ratings presented by policy statement within each stakeholder group (*continued*)

Statement	Round 1			Round 2			Round 3		
Policy statement	**Clinicians (%)**	**Patients (%)**	**Both (%)**	**Clinicians (%)**	**Patients (%)**	**Both (%)**	**Clinicians (%)**	**Patients (%)**	**Both (%)**
We need to determine how long we should wait before deciding wound healing is delayed or failed.	45	60.0	49.1	60.5	76.9	64.7	62.3	71.4	64.7
Follow-up should continue until there is evidence of complete wound healing.				60.5	92.3	68.6	To consensus		
Patients with symptomatic pilonidal disease always require a secondary care referral.				55.3	69.2	58.8	59.5	57.1	58.8
Novel minimally invasive procedures (e.g. laser) should be thoroughly appraised in randomised trials before general adoption.				73.7	53.8	70.6	To consensus		
Imaging is rarely useful in pilonidal disease.				42.1	30.8	39.2	32.4	42.9	35.3
SDM should be employed when discussing treatment options.				86.8	92.3	88.2	To consensus		

TABLE 36 Ratings presented by research statement within each stakeholder group

Statement	Round 1			Round 2			Round 3		
Research statement	**Clinicians (%)**	**Patients (%)**	**Both (%)**	**Clinicians (%)**	**Patients (%)**	**Both (%)**	**Clinicians (%)**	**Patients (%)**	**Both (%)**
A future randomised trial (RCT) in the treatment of pilonidal sinus should compare widely used techniques.	90	53.3	80.0	To consensus					
Postsurgical care (e.g. wound care, follow-up etc.) is an important part of treatment strategy. Further work is required to establish the optimum way to deliver this.	87.5	73.3	83.6	To consensus					
Future research should aim to define an algorithm or decision tree to aid surgeon decision-making.	77.5	80.0	78.2	To consensus					
A future randomised trial (RCT) should include two broad groups of interventions – major (i.e. asymmetric closure, leave open and midline closure) vs. minor (i.e. minimal excision).	67.5	60.0	65.5	71.1	92.3	76.5	To consensus		
A decision aid targeted at patients to help understand treatment options might improve patient satisfaction with treatment.	80.0	80.0	80.0	To consensus					
Classification should include an assessment of symptoms.	82.5	60.0	76.4	89.5	92.3	90.2	To consensus		
Classification systems should include data related to hair type and distribution.	32.5	40.0	34.5	50.0	76.9	56.9	51.4	57.1	52.9
Classification systems should include data on recurrent skin infections in non-pilonidal areas.	45.0	66.7	50.9	44.7	30.8	41.2	54.1	42.9	51.0
Classification systems should include data on extent of disease beyond the natal cleft.	85.0	46.7	74.5	81.6	46.2	72.5	To consensus		
Consistency in reporting patient and disease factors would help us better understand what characteristics are associated with good or bad outcomes.	85.0	66.7	80.0	84.2	92.3	86.3	To consensus		
A core outcome set for pilonidal disease might help us understand what outcomes are important to clinicians and patients following treatment of pilonidal disease. It may also improve future evaluations of treatments.	95.0	53.3	83.6	To consensus					

TABLE 36 Ratings presented by research statement within each stakeholder group (*continued*)

Statement	Round 1			Round 2			Round 3		
Research statement	**Clinicians (%)**	**Patients (%)**	**Both (%)**	**Clinicians (%)**	**Patients (%)**	**Both (%)**	**Clinicians (%)**	**Patients (%)**	**Both (%)**
There is a need for a PRO to be used in future pilonidal sinus research.	90.0	86.7	89.1	To consensus					
Future research should explore whether hair removal reduces the risk of wound complications or recurrence of pilonidal disease.	90.0	86.7	89.1	To consensus					
Future research should explore whether weight loss reduces the risk of wound complications or recurrence of pilonidal disease.	52.5	53.3	52.7	57.9	53.8	56.9	56.8	42.9	52.9
Future research should explore whether smoking behaviours reduce the risk of wound complications and/or recurrence of pilonidal disease.	92.5	46.7	80.0	To consensus					
Future research should assess the role of postoperative antibiotic treatment in wound healing and/or recurrence.	70.0	66.7	69.1	60.5	92.3	68.6	To consensus		
Future research should explore the role wound dressings play in wound healing and/or recurrence.	75	60.0	70.9	60.5	76.9	64.7	73.0	78.6	74.5
A future randomised trial (RCT) should compare procedures in mild or minimal disease where the wound is left open (e.g. pit picking and EPSiT) vs. closure of the wound (e.g. glue).	75	80.0	76.4	To consensus					
A future randomised trial (RCT) should compare non-excisional therapies.	77.5	60.0	72.7	81.6	76.9	80.4	To consensus		
Future research should explore the role of patient characteristics including genetics and microbiome on the pilonidal disease process.				34.2	76.9	45.1	48.6	50.0	49.0
Wide excision and leave open procedures should not be included in any future trial.				44.7	38.5	43.1	45.9	42.9	45.1
Future research should compare major procedures (e.g. flaps) against minor procedures (e.g. pit picking, glue) stratified by disease severity.				76.3	92.3	80.4	To consensus		

TABLE 37 Baseline demographics and disease characteristics by 6-month follow-up attendance

Characteristic	Measure	Attended 6-month follow-up (*n* = 476)	Did not attend 6-month follow-up (*n* = 191)
Age	*N* (%)	476 (100%)	191 (100%)
	Median (IQR)	27.0 (22.0–31.5)	28.0 (23.0–35.0)
Sex	Male	338 (71%)	147 (77%)
	Female	138 (29%)	44 (23%)
Ethnicity	White	406 (85%)	164 (86%)
	Asian/Asian British	45 (9%)	13 (7%)
	Mixed/multiple ethnic groups	12 (3%)	1 (1%)
	Black/African/ Caribbean/Black British	4 (1%)	4 (2%)
	Other ethnic group	3 (1%)	4 (2%)
	Prefer not to say	4 (1%)	1 (1%)
BMI (kg/m^2)	*N* (%)	443 (93%)	169 (88%)
	Median (IQR)	28.4 (24.9–32.8)	28.1 (25.1–31.9)
Number of baths and/or showers in a typical week	*N* (%)	462 (97%)	182 (95%)
	Median (IQR)	7.0 (6.0–7.0)	7.0 (6.0–7.0)
Seated for more than 6 hours in a working day	No	224 (47%)	94 (49%)
	Yes	241 (51%)	91 (48%)
First-degree relatives with history of PSD	No	386 (81%)	156 (82%)
	Yes	88 (18%)	34 (18%)
Smoking status	Non-smoker	283 (59%)	91 (48%)
	Current smoker	123 (26%)	73 (38%)
	Current e-cigarette smoker	26 (5%)	11 (6%)
Number of pits	*N* (%)	461 (97%)	179 (94%)
	Median (IQR)	2.0 (1.0–3.0)	2.0 (1.0–3.0)
Length of pits (spread)	*N* (%)	296 (62%)	110 (58%)
	Median (IQR)	21.0 (10.0–40.0)	25.0 (11.0–45.0)
Number of previous procedures	0	255 (54%)	102 (53%)
	1	124 (26%)	50 (26%)
	2	54 (11%)	24 (13%)
	3 or more	43 (9%)	15 (8%)
Previous procedure	Elective procedure for PSD	98 (21%)	50 (26%)
	Acute drainage for PSD	146 (31%)	49 (26%)

TABLE 37 Baseline demographics and disease characteristics by 6-month follow-up attendance (*continued*)

Characteristic	Measure	Attended 6-month follow-up (*n* = 476)	Did not attend 6-month follow-up (*n* = 191)
	Emergency procedure for PSD	3 (1%)	2 (1%)
Wysocki classification	Type 1	127 (27%)	55 (29%)
	Type 2	240 (50%)	84 (44%)
	Type 3	38 (8%)	12 (6%)
	Type 4	64 (13%)	37 (19%)
	None of the above	2 (0%)	2 (1%)
Distribution of lateral openings	No lateral openings	203 (43%)	92 (48%)
	Unilateral	219 (46%)	68 (36%)
	Bilateral	13 (3%)	7 (4%)

TABLE 38 Comparison of pain on day 1 and day 7 between asymmetric closure and minimal excision

Model	Asymmetric closure		Minimal excision		*N*	Mean difference (95% CI)[a]
	n	Mean (SD)	*n*	Mean (SD)		
Pain (day 1)						
Raw difference	253	4.55 (2.37)	242	2.60 (2.24)	495	1.95 (1.54 to 2.36)
Risk-adjusted – Wysocki					493	1.96 (1.54 to 2.38)
Risk-adjusted – chosen model (sex, smoking, Wysocki)					444	1.91 (1.46 to 2.35)
Risk-adjusted – full model					339	1.98 (1.46 to 2.51)
Propensity-adjusted – IPW					490	1.88 (1.38 to 2.38)
Propensity matching					490	1.99 (1.49 to 2.49)
Augmented IPW					441	1.97 (1.50 to 2.43)
Pain (day 7)						
Raw difference	246	3.26 (2.35)	221	1.86 (2.18)	467	1.40 (0.99 to 1.81)
Risk-adjusted – Wysocki					465	1.35 (0.92 to 1.78)
Risk-adjusted – chosen model (lateral distribution, sex, Wysocki)					437	1.22 (0.78 to 1.66)
Risk-adjusted – full model					321	1.21 (0.69 to 1.72)
Propensity-adjusted – IPW					462	1.39 (0.93 to 1.85)
Propensity matching					462	1.45 (1.00 to 1.91)
Augmented IPW					436	1.33 (0.89 to 1.76)

a Reference group: minimal excision. Risk-adjusted models: linear regression model with adjustment for covariates as listed, full model includes adjustment for sex, BMI, natal cleft depth, gluteal hair, smoking status, Wysocki classification, pit density, lateral distance, lateral distribution and presence of pus; propensity-adjusted and propensity matching adjust for sex, Wysocki classification and presence of pus; augmented IPW adjusted for sex, Wysocki classification, lateral distance and presence of pus.

TABLE 39 Comparison of complications during follow-up between asymmetric closure and minimal excision

Complications	Asymmetric closure	Minimal excision	*n*	Risk difference (95% CI)[a]
Raw difference	135/265 (51%)	94/258 (36%)	523	14.5 (6.1 to 22.9)
Risk-adjusted – Wysocki			521	13.7 (5.0 to 22.4)
Risk-adjusted – chosen model (BMI, Wysocki)			476	14.4 (5.2 to 23.5)
Risk-adjusted – full model			354	14.0 (3.5 to 24.5)
Propensity-adjusted – IPW			518	13.9 (4.8 to 23.1)
Propensity matching			518	12.8 (3.1 to 22.6)
Augmented IPW			473	15.0 (5.8 to 24.3)

a Reference group: minimal excision, risk-adjusted difference estimated using logistic regression with adjustment for covariates as listed, full model includes adjustment for sex, BMI, natal cleft depth, gluteal hair, smoking status, Wysocki classification, pit density, lateral distance, lateral distribution and presence of pus. Propensity-adjusted and propensity matching adjusted for sex, Wysocki classification, presence of pus. Augmented IPW adjusted for sex, Wysocki classification, BMI, and presence of pus.

TABLE 40 Comparison of recurrence between asymmetric closure and minimal excision

	Recurrence				Recurrence (within 6 months)			
Recurrence	Asymmetric closure	Minimal excision	*n*	Risk difference (95% CI)[a]	Asymmetric closure	Minimal excision	*n*	Risk difference (95% CI)[a]
Raw difference	55/257 (21%)	87/256 (34%)	513	−12.6 (−20.3 to −4.9)	28/226 (12%)	61/229 (27%)	455	−14.2 (−21.4 to −7.1)
Risk-adjusted – Wysocki			511	−13.1 (−21.0 to −5.2)			453	−14.5 (−22.0 to −7.1)
Risk-adjusted – Chosen model (Wysocki, pit density)			484	−11.5 (−19.7 to −3.4)			428	−12.8 (−20.5 to −5.1)
Risk-adjusted – full model			343	−10.1 (−20.0 to −0.2)			304	−9.0 (−18.5 to 0.5)
Propensity-adjusted – inverse weighting			508	−16.2 (−25.1 to −7.3)			450	−15.7 (−24.1 to −7.2)
Propensity matching			508	−13.4 (−22.6 to −4.2)			450	−15.2 (−24.2 to −6.2)
Augmented IPW			483	−11.9 (−20.5 to −3.2)			427	−12.3 (−20.6 to −4.0)

a Reference group: minimal excision, risk-adjusted difference estimated using logistic regression with adjustment for covariates as listed, full model includes adjustment for sex, BMI, natal cleft depth, gluteal hair, smoking status, Wysocki classification, pit density, lateral distance, lateral distribution and presence of pus. Propensity-adjusted and propensity matching adjusted for sex, Wysocki classification and presence of pus. Augmented IPW adjusted for sex, Wysocki classification, BMI and presence of pus.

TABLE 41 Comparison of treatment failure between asymmetric closure and minimal excision

Recurrence	Asymmetric closure	Minimal excision	*n*	Risk difference (95% CI)[a]
Raw difference	109/257 (42%)	121/257 (47%)	514	−4.7 (−13.3 to 3.9)
Risk-adjusted – Wysocki			512	−5.4 (−14.3 to 3.4)
Risk-adjusted – chosen model (Wysocki, pit density)			485	−5.7 (−14.8 to 3.4)
Risk-adjusted – full model			344	−3.7 (−14.5 to 7.1)
Propensity-adjusted – inverse weighting			509	−8.4 (−18.0 to 1.2)
Propensity matching			509	−5.8 (−15.7 to 4.2)
Augmented IPW			484	−4.8 (−14.1 to 4.6)

a Reference group: minimal excision, risk-adjusted difference estimated using logistic regression with adjustment for covariates as listed, full model includes adjustment for sex, BMI, natal cleft depth, gluteal hair, smoking status, Wysocki classification, pit density, lateral distance, lateral distribution and presence of pus. Propensity-adjusted and propensity matching adjusted for sex, Wysocki classification and presence of pus. Augmented IPW adjusted for sex, Wysocki classification, BMI and presence of pus.

TABLE 42 Comparison of time to return to normal activities between asymmetric closure and minimal excision

	Asymmetric closure		Minor procedure			Difference
Model	*n*	Median (IQR)	*n*	Median (IQR)	*N*	Mean difference (95% CI)
Raw difference	255	30 (14–60)	241	7 (4–21)	496	18.3 (13.6 to 23.1)
Risk-adjusted – Wysocki					492	17.7 (12.8 to 22.5)
Risk-adjusted – chosen model (Wysocki, lateral distance, natal cleft depth					431	16.6 (11.7 to 21.5)
Risk-adjusted – full model					340	15.5 (10.1 to 20.8)
Propensity-adjusted – IPW					489	25.8 (16.6 to 35.0)
Augmented IPW						(Not estimable)

TABLE 43 Comparison of time to wound healing between asymmetric closure and minimal excision

	Asymmetric closure		Minor procedure			Difference
Model	*n*	Median (IQR)	*n*	Median (IQR)	*N*	Mean difference (95% CI)
Raw difference	239	57 (30–134)	217	30 (14–154)	456	31.3 (18.2 to 44.3)
Risk-adjusted – Wysocki					452	28.1 (14.6 to 41.5)
Risk-adjusted – chosen model (Wysocki, BMI, smoking status, pus)					371	26.3 (10.9 to 41.7)
Risk-adjusted – full model					311	26.6 (10.9 to 42.3)
Propensity-adjusted – IPW					449	30.1 (14.5 to 45.7)
Augmented IPW					371	21.2 (3.1 to 39.3)

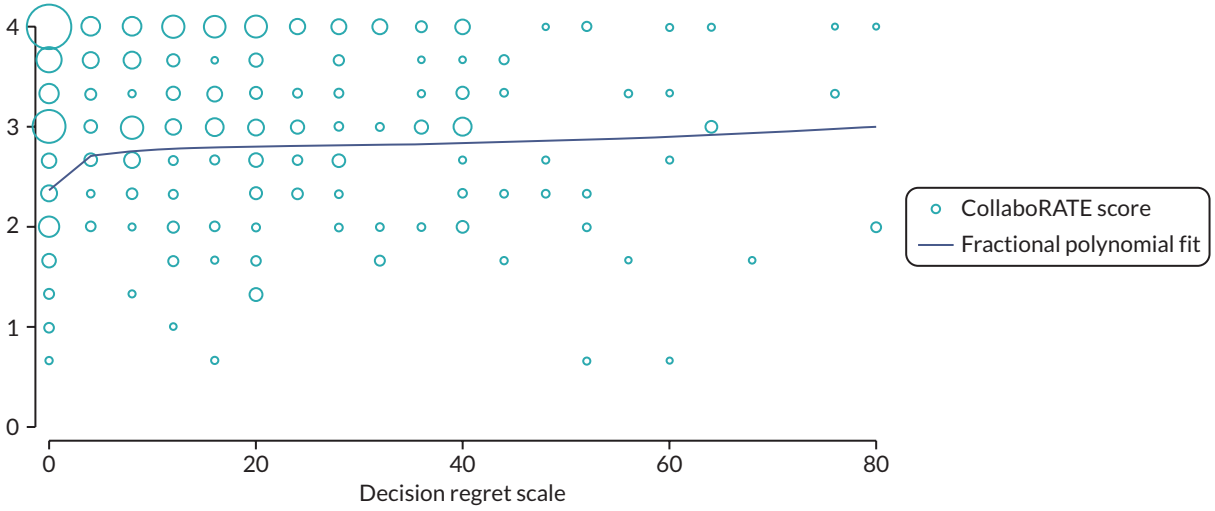


FIGURE 14 CollaboRATE score at baseline by 6-month DR score.

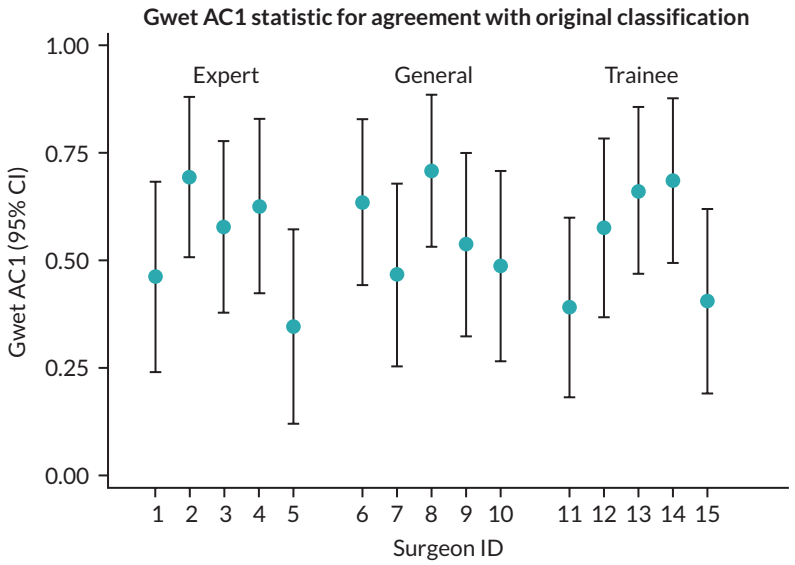


FIGURE 15 Gwet AC1 chance-corrected agreement between assessment and original.

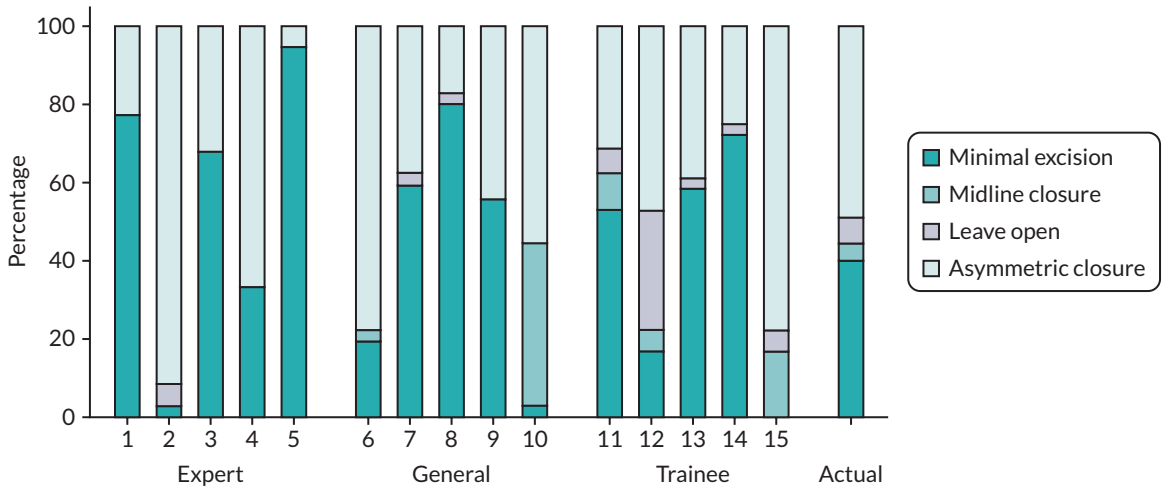


FIGURE 16 Preferred treatment by surgeon and actual treatment within the study.

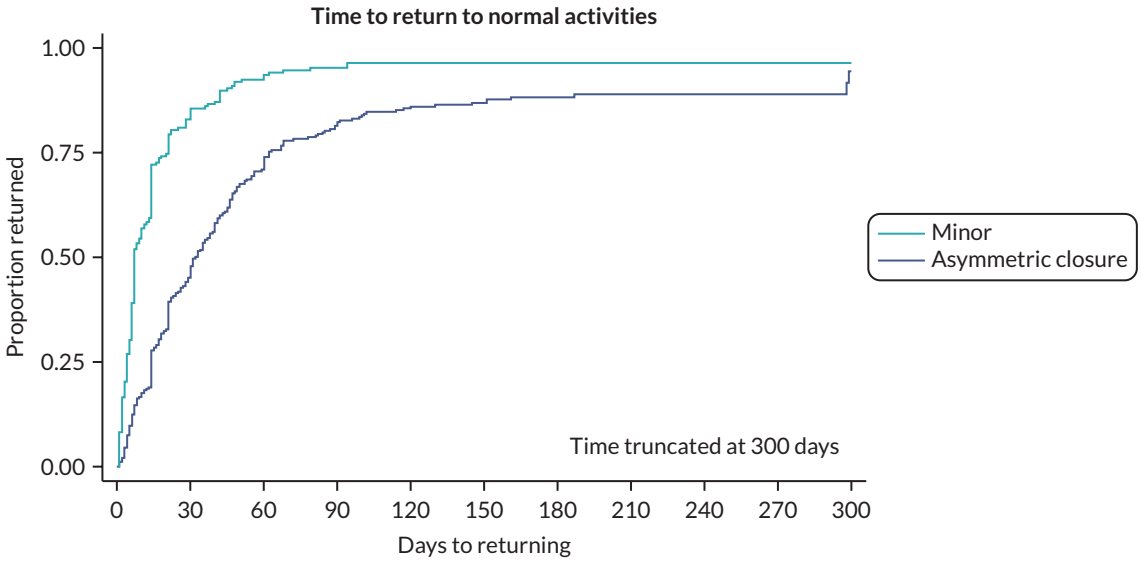


FIGURE 17 Time to return to normal activities between asymmetric closure and minimal excision.

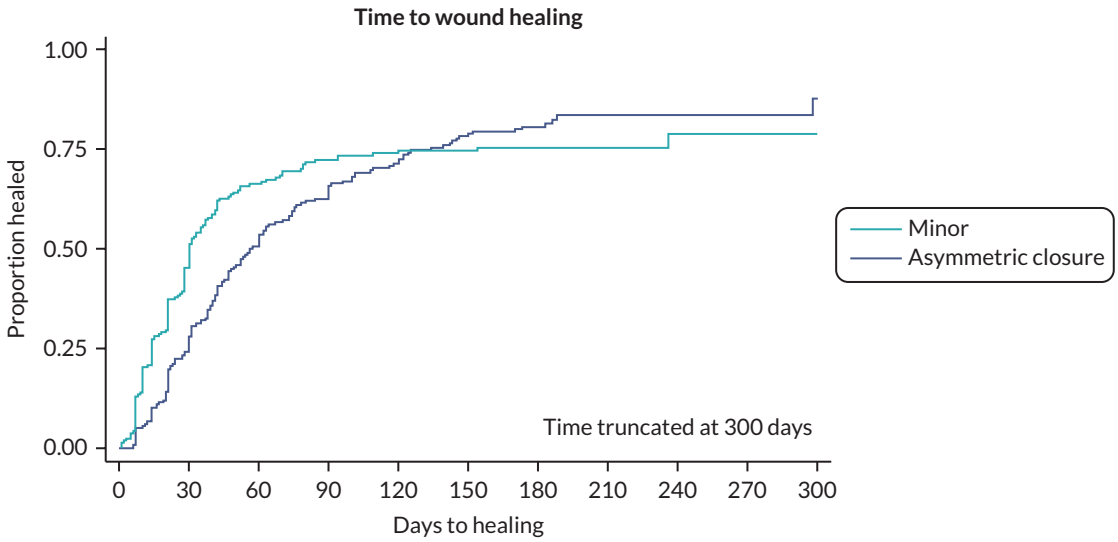


FIGURE 18 Time to wound healing between asymmetric closure and minimal excision.

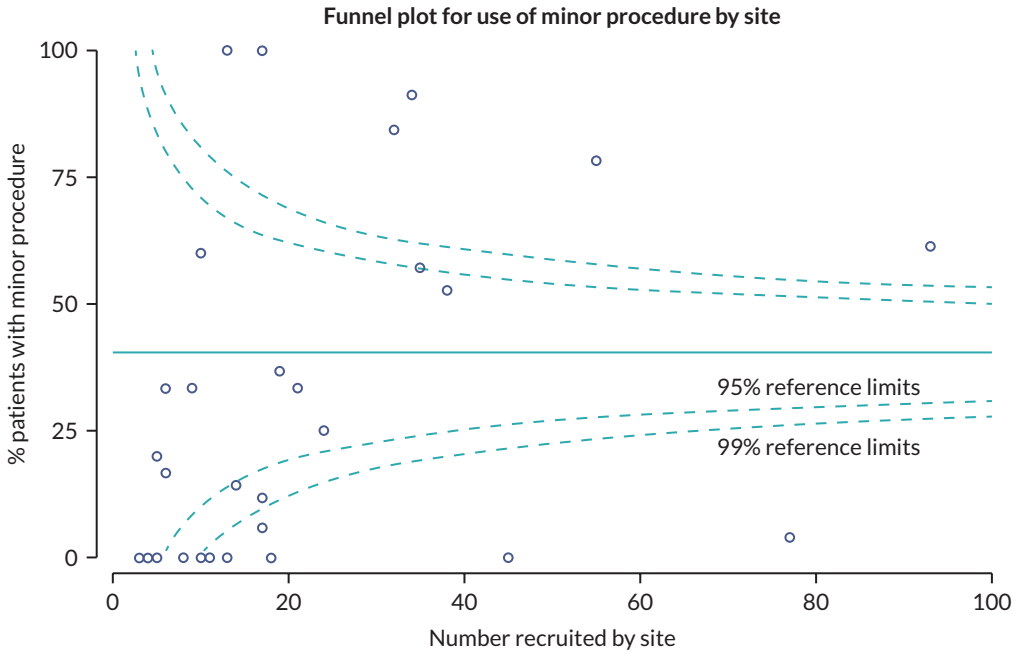


FIGURE 19 Funnel plot of minor procedure use by site.

BOX 1 Example DCE task

Please imagine this scenario:

- You have pilonidal disease.
- You are experiencing pain, itchiness, discharge and some discomfort when you move.
- You've had emergency surgery to drain an abscess, but your symptoms persist.
- You are told that you need further surgery to cure the pilonidal disease.
- You are now considering which treatment option you should choose next to cure this disease.

Now we would like to understand which outcomes would affect your decision to choose a treatment for the scenario described above. To help us understand how important the different outcomes of surgery are to you, we are going to ask you to make a series of 16 choices. In each choice you will be asked to choose between two treatments that you can take.

The two treatments differ in:

- **Recovery time** – which is the usual time a patient takes to get back to doing normal activities without pain, such as bending and putting on socks, being able to go to work or attend school. In the two treatments recovery time can vary from 1 week, 2 weeks, 6 weeks to 12 weeks.
- **Risk of infection and persistence of symptoms** – for a proportion of people who have surgery, the pilonidal sinus does not get better. This means that you will need further treatment. In the two treatments the risk of infection and persistence of symptoms can vary from 5%, 10%, 20% to 30%. For example, a 5% risk means that if 100 people had the same surgery for pilonidal disease, 95 people would be cured of the pilonidal disease but for 5 people the symptoms will continue, and the disease will not get better.

We will now move onto the questions where you have to make a treatment choice. In total there are 16 choices to consider.

	Treatment A	Treatment B
Recovery time	12 weeks	1 week
Risk of infection/persistence	In 5 out of 100 (5%) patients pilonidal sinus will persist	In 30 out of 100 (30%) patients pilonidal sinus will persist

Which treatment would you choose?

Treatment A ○ Treatment B ○

BOX 2 Treatment ranking exercise

Please imagine this scenario:

- You have pilonidal disease.
- You are experiencing pain, itchiness, discharge and some discomfort when you move.
- You've had emergency surgery to drain an abscess, but your symptoms persist.
- You are told that you need further surgery to cure the pilonidal disease.
- You are now considering which treatment option you should choose next to cure this disease.
- Your consultant has presented to you 5 different treatment options.

The following table is a summary of the 5 treatments.

	Excision of skin and leave the wound open	Excision of the skin and closure of the wound with stitches	Excision of the skin and closure of the wound with a skin flap and stitches	Excision of the sinuses and closure of the wound with glue	Excision of the sinuses only and leave open the wound to heal
Type of excision and closure	Cutting out all the diseased skin and leave open	Cutting out all the diseased skin and close with stitches	Cutting out all the disease skin and close with stitches and a skin flap	Scraping away debris and close using glue	Cutting, destroying or scraping the sinus only and leave open to heal
Type of anaesthetic	General anaesthetic	General anaesthetic	General anaesthetic	Local anaesthetic	Local anaesthetic
Length of hospital stay	Day operation	Day operation	Stay one/two nights in hospital	Day operation	Day operation
Wound care	Nurse support 2 or more times a week	No need for nurse support, but you will need to change dressings at home	No need for nurse support, but you will need to change dressings at home	Should not need dressing changes	No need for nurse support, but you will need to change dressings at home
Pain medication required	More than 2 weeks	2 weeks	2 weeks	Not usually required after the initial 1-2 days after surgery	Few days
Infection risk	1 in 20 people	2 to 4 in 20 people	1 in 20 people	1 in 20 people	1 in 20 people
Healing time	Few months	2 to 3 weeks	2 to 3 weeks	A few days	A few days
Risk of recurrence	2 to 4 in 20 people	2 to 4 in 20 people	1 in 20 people	5 in 20 people	1 to 2 in 20 people
Scarring	Large scar	Small scar	Large scar	Small dimple/no scar	Small scar

Now that you know more about the treatment options available to you, we would like you to consider which treatments you would prefer. Please rank these treatments in order of preference: 1 = best preferred treatment and 5 = least preferred treatment.

_____ Excision of skin and leave the wound open

_____ Excision of the skin and closure of the wound with stitches

_____ Excision of the skin and closure of the wound with a skin flap and stitches

_____ Excision of the sinuses and closure of the wound with glue

_____ Excision of the sinuses only and leave open the wound to heal

EME
HSDR
HTA
PGfAR
PHR

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

Published by the NIHR Journals Library