NIHR | National Institute for Health and Care Research

Journals Library

Check for updates

Health Technology Assessment

Volume 29 • Issue 40 • August 2025

ISSN 2046-4924

Clinical and cost-effectiveness of percutaneous nephrolithotomy, flexible ureterorenoscopy and extracorporeal shockwave lithotripsy for lower pole stones: the PUrE RCTs

Oliver Wiseman, Daron Smith, Kath Starr, Lorna Aucott, Rodolfo Hernández, Ruth Thomas, Steven MacLennan, Charles Terry Clark, Graeme MacLennan, Dawn McRae, Victoria Bell, Seonaidh Cotton, Zara Gall, Ben Turney and Samuel McClinton

DOI 10.3310/WFRE6844

Health Technology Assessment

Extended Research Article

Clinical and cost-effectiveness of percutaneous nephrolithotomy, flexible ureterorenoscopy and extracorporeal shockwave lithotripsy for lower pole stones: the PUrE RCTs

Oliver Wiseman,[1] Daron Smith,[2] Kath Starr,[3] Lorna Aucott,[4] Rodolfo Hernández,[5] Ruth Thomas,[6] Steven MacLennan,[7] Charles Terry Clark,[8] Graeme MacLennan,[6] Dawn McRae,[6] Victoria Bell,[6] Seonaidh Cotton,[6] Zara Gall,[9] Ben Turney[10] and Samuel McClinton[7*]

[1]Department of Urology, Addenbrooke's Hospital, Cambridge University Hospitals NHS Trust, Cambridge, UK
[2]Stone and Endourology Unit, University College Hospital, University College London Hospitals NHS Foundation Trust, London, UK
[3]Warwick Medical School, University of Warwick, Warwick, UK
[4]Health Services Research Unit, University of Aberdeen, Aberdeen, UK
[5]Health Economics Research Unit, University of Aberdeen, Aberdeen, UK
[6]Centre for Healthcare Randomised Trials, Health Services Research Unit, University of Aberdeen, Aberdeen, UK
[7]Academic Urology Unit, University of Aberdeen, Aberdeen, UK
[8]Patient/Public Representative
[9]Department of Urology, Stepping Hill Hospital, Stockport NHS Foundation Trust, Stockport, UK
[10]Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK

*Corresponding author mcclintonsam@gmail.com

Published August 2025
DOI: 10.3310/WFRE6844

This report should be referenced as follows:

Wiseman O, Smith D, Starr K, Aucott L, Hernández R, Thomas R, *et al.* Clinical and cost-effectiveness of percutaneous nephrolithotomy, flexible ureterorenoscopy and extracorporeal shockwave lithotripsy for lower pole stones: the PUrE RCTs. *Health Technol Assess* 2025;**29**(40). https://doi.org/10.3310/WFRE6844

Health Technology Assessment

ISSN 2046-4924 (Online)

Impact factor: 4

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 4 and is ranked 30th (out of 174 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 13/152/02. The contractual start date was in December 2015. The draft manuscript began editorial review in June 2023 and was accepted for publication in March 2024. 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.

Copyright © 2025 Wiseman *et al*. This work was produced by Wiseman *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

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

Abstract

Background: Renal tract stone disease is common. The three intervention options are shockwave lithotripsy, flexible ureteroscopic stone treatment and keyhole surgery.

Objectives: To determine which of shockwave lithotripsy, flexible ureteroscopic stone treatment and keyhole surgery offer the best outcomes in terms of health and quality of life, clinical effectiveness and cost-effectiveness for people with lower pole kidney stones.

Design: The PUrE study comprised two pragmatic multicentre, open-label, superiority randomised controlled trials: RCT1 for lower pole stones ≤ 10 mm and RCT2 for lower pole stones > 10 and ≤ 25 mm.

Setting: National Health Service Urology departments.

Participants: Adults presenting with lower pole renal stones, able to undergo any of the treatments and complete trial procedures.

Intervention: Eligible participants were randomised in RCT1 to flexible ureteroscopic stone treatment or shockwave lithotripsy; and in RCT2 to flexible ureteroscopic stone treatment or keyhole surgery.

Main outcome measures: The primary outcome measure was health status 'area under the curve', measured weekly to 12 weeks post intervention with the EuroQol-5 Dimensions, five-level version. The primary economic outcome was the incremental cost per quality-adjusted life-year gained at 12 months from randomisation.

Results: *RCT1:* A total of 461 participants were randomised: 231 to flexible ureteroscopic stone treatment; and 230 to shockwave lithotripsy. *RCT2:* A total of 159 participants were randomised: 73 to flexible ureteroscopic stone treatment; and 86 to keyhole surgery.

Primary outcome: *RCT1:* The mean health status area under the curve was 0.807 (standard deviation 0.205) in the flexible ureteroscopic stone treatment group (n = 164) and 0.826 (standard deviation 0.207) in the shockwave lithotripsy group (n = 188). The between-group difference, 0.024 (95% confidence interval −0.004 to 0.053), was a small difference in favour of flexible ureteroscopic stone treatment after correcting for a baseline imbalance. Complete stone clearance was higher with flexible ureteroscopic stone treatment (72%) than shockwave lithotripsy (36%).

RCT2: The mean health status area under the curve was 0.794 (standard deviation 0.198) in the flexible ureteroscopic stone treatment group (n = 57) and 0.818 (standard deviation 0.217) in the keyhole surgery group (n = 63). The between-group difference, −0.07 (95% confidence interval −0.11 to −0.02), was a borderline meaningful difference favouring keyhole surgery. Complete stone clearance was higher with keyhole surgery (71%) than flexible ureteroscopic stone treatment (48%).

Economic evaluation: *RCT1*: Flexible ureteroscopic stone treatment is more costly (£1138; 95% confidence interval £646 to £1631) and produces 0.017 (95% confidence interval −0.008 to 0.043) additional quality-adjusted life-years; with an incremental cost-effectiveness ratio of £65,163 per quality-adjusted life-year gained. Shockwave lithotripsy has a 99.9% chance of being cost-effective at a £20,000 threshold value.

RCT2: Flexible ureteroscopic stone treatment is more costly (£733; 95% confidence interval −£508 to £1973) and produces fewer quality-adjusted life-years (−0.001; 95% confidence interval −0.044 to 0.042). Keyhole surgery has an 87% chance of being cost-effective at a £20,000 threshold value.

Limitations: Blinding of participants and healthcare providers was not possible. There were differential waiting times between interventions in RCT1; however, adjusting for this gave similar treatment effect estimates.

Conclusions: The PUrE study found in RCT1 that shockwave lithotripsy was more cost-effective than flexible ureteroscopic stone treatment, with no meaningful difference in patient health status even though complete stone-free rates were higher with flexible ureteroscopic stone treatment. In RCT2, keyhole surgery was more cost-effective than flexible ureteroscopic stone treatment on a micro-costing basis, which better reflects treatment cost differences to the NHS. Keyhole surgery was marginally beneficial for health status with higher complete stone-free rates.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Future work: What effect will suction devices, improvements in laser technology, and intraoperative pressure monitoring have on postoperative pain, quality of life, stone-free rates, complications, and costs of flexible ureteroscopic stone treatment?

What effect does miniaturisation of keyhole surgery have on postoperative pain, length of stay, complications, stone-free rates and costs?

Trial registration: This trial is registered as ISRCTN98970319.

Funding: This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: 13/152/02) and is published in full in *Health Technology Assessment*; Vol. 29, No. 40. See the NIHR Funding and Awards website for further award information.

Contents

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

List of tables

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

List of figures

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

List of abbreviations

AUC	area under the curve
BAUS	British Association of Urological Surgeons
BMI	body mass index
CEAC	cost-effectiveness acceptability curve
CHaRT	Centre for Healthcare Randomised Trials
CI	chief investigator
CONSORT	Consolidated Standards of Reporting Trials
CReSP	Cambridge Renal Stone Patient Reported Outcome Measure
CRF	case report form
CTKUB	computed tomography, kidneys, ureters and bladder
CTU	Clinical Trials Unit
DMC	Data Monitoring Committee
EAU	European Association of Urology
EQ-5D	EuroQol-5 Dimensions
EQ-5D-5L	EuroQol-5 Dimensions, five-level version
ESWL	extracorporeal shockwave lithotripsy
FURS	flexible ureterorenoscopy
GCP	good clinical practice
GLM	general linear regression model
GP	general practitioner
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HRG	Health Care Resource Group
HSRU	Health Services Research Unit
HTA	Health Technology Assessment
ICER	incremental cost-effectiveness ratio
ISRCTN	International Standard Randomised Controlled Trial Number
ITT	intention to treat
LPS	lower pole stone
MCID	minimal clinical important difference
MICE	multiple imputation using chained equations
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health and Care Research
NLLS	Non-Linear Least Squares
NRS	numeric rating scale
PCNL	percutaneous nephrolithotomy
PI	principal investigator
PIL	patient information leaflet
PMG	Project Management Group
PO	primary outcome
PPI	patient and public involvement
PROM	patient-reported outcome measure
PSA	probabilistic sensitivity analysis
PUrE	The Percutaneous nephrolithotomy, flexible Ureterorenoscopy and Extracorporeal shockwave lithotripsy for lower pole kidney stone study
QALY	quality-adjusted life-year
QoL	quality of life
RCT	randomised controlled trial
REC	Research Ethics Committee
RoB	risk of bias
SAE	serious adverse event
SAP	statistical analysis plan
SF-12	Short Form questionnaire-12 items
SFR	stone-free rate
SMP	supermini percutaneous nephrolithotomy (approximately 14Fr Track, may be combined with suction)
SWAT	Study Within a Trial
TSC	Trial Steering Committee
UMP	ultramini percutaneous nephrolithotomy (approximately 14Fr track)
URS	ureteroscopy

Plain language summary

About 10% of people will suffer from kidney stones in their lifetime. Approximately 50% of those people will experience symptoms, typically kidney pain, and about 25% of patients with stones will require active treatment. Active treatments include shockwave therapy, telescopic surgery and keyhole surgery. Stones commonly develop in the lower part (pole) of the kidney.

Previous studies suggested that for lower pole stones:

- and stones smaller than 10 mm in size, telescopic surgery is more likely to remove the stone with a single treatment
- and for larger stones, which are > 10 mm but smaller than 25 mm, keyhole surgery is more likely to remove the stone with a single treatment.

Remarkably little evidence was available on the impact of these treatments on quality of life for patients.

What did PUrE do?

We assessed the effect on the participants' health, and the economic impact to the National Health Service, of the different care pathways by conducting 2 studies involving over 600 patients with lower pole stones.

- The first study (smaller stones) compared telescopic surgery with shockwave therapy.
- The second study (larger stones) compared telescopic surgery with keyhole surgery.

A computer program (random allocation) decided which treatment each person received. The patients, and the doctors caring for them, proceeded with the agreed treatment. All patients were requested to fill in quality-of-life questionnaires on a regular basis.

What did PUrE find?

We found that for shockwave therapy and keyhole surgery, improvement in participants' health status was similar to telescopic surgery. Telescopic surgery was less cost-effective for the National Health Service in both studies.

What does this mean?

These studies show that, based on cost and health status, shockwave therapy and keyhole surgery are the first choice for active treatment in the National Health Service.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Scientific summary

Background

Renal tract stone disease is very common, with a lifetime prevalence of approximately 10% across the world. It mainly affects adults of working age, and the incidence has been increasing over recent decades. Approximately 50% of people with renal tract stones will experience symptoms, typically kidney pain, and about 25% of patients with stones will require active treatment. Many of these stones occur in the lower pole of the kidney and the three standard active intervention options are extracorporeal shockwave lithotripsy (ESWL), flexible ureteroscopic stone treatment (FURS) and keyhole surgery [percutaneous nephrolithotomy (PCNL)].

Objectives

The aim of the PUrE study was to determine which of ESWL, FURS and PCNL offer the best treatment outcomes in terms of health status, clinical-effectiveness, and cost-effectiveness for people with lower pole kidney stones seeking treatment within the UK NHS.

Methods

Design

Two pragmatic multicentre, patient-randomised, open-label superiority randomised controlled trials (RCTs): the first (RCT1) for lower pole stones (LPSs) ≤ 10 mm in maximum dimension and the second (RCT2) for LPSs > 10 and ≤ 25 mm.

Setting

National Health Service secondary care units across the UK, with a high volume of patients presenting with LPSs, and able to deliver all active treatments.

Participants

Adults (16 years or over) with lower pole renal stones judged to require active treatment.

Intervention

Treatment following either the ESWL, FURS or PCNL pathways. Participants that were eligible and consented were randomised within RCT1 to FURS or ESWL, or within RCT2 to FURS or PCNL.

Main outcome measures

Clinical: (1) Health status area under the curve (AUC) measured weekly to 12 weeks post intervention using the EuroQol-5 Dimensions, five-level version (EQ-5D-5L) and (2) stone clearance at 12 weeks.

Economic: Incremental cost per quality-adjusted life-years (QALYs) gained at 12-months from randomisation. QALYs are based on the responses to the EQ-5D-5L.

Results

Main outcome

RCT1: The mean health status AUC was 0.807 [standard deviation (SD) 0.205] in the FURS group (n = 164) and 0.826 (SD 0.207) in the ESWL group (n = 188). The adjusted effect estimate was 0.024 [confidence interval (CI) −0.004 to 0.053] and this was not significant (p = 0.097). Complete stone clearance was higher with FURS (72%) than with ESWL (36%).

RCT2: The mean health status AUC was 0.794 (SD 0.198) in the FURS group (n = 57) and was 0.818 (SD 0.217) in the PCNL group (n = 63). The adjusted effect estimate was −0.07 (CI −0.11 to −0.02; p = 0.006). Complete stone clearance was higher with PCNL (71%) than with FURS (48%).

Economic evaluation

RCT1: The mean cost for the NHS were £3362 and £2223 for the intention to treat (ITT) with FURS and ESWL groups, respectively, resulting in an adjusted cost difference of £1138 [95% confidence interval (CI) £646 to £1631]. The mean QALYs per participant were 0.804 and 0.787 for the FURS and ESWL groups, respectively, producing an adjusted QALY difference of 0.017 QALYs (95% CI −0.008 to 0.043) for the 12-month follow-up period. The incremental cost-effectiveness ratio (ICER) between FURS and ESWL was £65,163 per QALY gained by FURS. At a threshold value of £20,000 per QALY, ESWL has a 99.9% chance of being cost-effective.

RCT2: The mean cost for the NHS was £5298 and £4565 for the ITT with FURS and PCNL, respectively; giving an adjusted difference of £733 (95% CI −£508 to £1973). The mean QALY per participant were 0.773 and 0.775 for FURS and PCNL, respectively, yielding an adjusted difference of −0.001 (95% CI −0.044 to 0.042).

Therefore, ITT with FURS is on average more costly and does not produce additional QALYs compared with ITT with PCNL. PCNL is highly likely to be cost-effective at the usual cost-effectiveness threshold values used for decision-making in the UK NHS (e.g. probability of 0.87 at £20,000 per QALY gained). These results, however, are dependent on the method used to estimate the costs of the initial interventions. Therefore, using costings based on the Health Care Resource Group (HRG) the mean cost for the NHS was £5769 for FURS and £6703 for PCNL; resulting in FURS being £934 less costly than PCNL (95% CI −£2582 to £714: ICER: £883,375). That is, ITT with FURS would save on average, £883,375 per QALY forgone.

Comparison with similar randomised trials

The health status outcomes and the economic outcomes, as they relate to the UK NHS, have not been evaluated in previous randomised trials.

Conclusions

The PUrE study shows in RCT1 that ESWL for lower renal pole stones under 10 mm was more cost-effective than FURS, and there was no meaningful difference in patient health status. Stone-free rates (SFRs) were higher, however with FURS. From an overall NHS perspective, the costs savings of treating all patients with these stones with ESWL would be substantial. In RCT2, for larger stones 10–25 mm, PCNL was more cost-effective than FURS when using micro-costing to cost the interventions. Health status was marginally beneficial and SFRs were higher with PCNL.

Implications for health care

The choice of health status assessment as the primary outcome measure provides important data for patient counselling for decision-making and for resource allocation based on the cost per QALY analysis. RCT1 demonstrated that ESWL was the more cost-effective treatment for LPSs < 10 mm. This complements the recommendation from National Institute for Health and Care Excellence (NICE) Renal Stone Guidelines 2019 that ESWL should be offered as the first-line intervention for renal stones of this size.

In RCT2 the cost-effectiveness for FURS and PCNL is less clear-cut, because of the discrepancy between a micro-costing analysis (which showed PCNL was more cost-effective) and HRG-based analysis (which showed that FURS was more cost-effective). It is unclear to us why the HRG costs differ that much between FURS and PCNL given the resources used do not substantially differ between the two procedures. Greater transparency in the way the HRG are costed would be beneficial to inform decision-making in the NHS.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

The consistent drop in health status at week one in both trials, particularly for those in the FURS treatment arms has important implications for practice during consent to treatment. Based on this data it is important for clinicians to emphasise that patients are likely to feel worse before they feel better. While this is true of most surgical interventions, and therefore likely to be expected by most patients, the PUrE study provides detailed information regarding the anticipated post-treatment health status that will help make patients' consent to be more fully informed.

The results from RCT1 show no meaningful difference in health status, while RCT2 indicates marginal benefit for PCNL. However, when considering secondary outcomes, patients should be counselled that for smaller LPSs < 10 mm, FURS leads to a higher stone free rate than ESWL, and for LPSs between 10 and 25 mm, PCNL leads to higher stone free rates than FURS.

Recommendations for research

What effect will suction devices, improvements in laser technology, and intraoperative pressure monitoring have on postoperative pain, QoL, SFRs, complications, and costs of FURS?

What is the effect of minaturisation of PCNL on postoperative pain, length of stay, complications, SFRs, and costs?

What is the clinical and cost-effectiveness of full metabolic assessment compared with standard advice alone, in people who have undergone treatment for LPSs?

Trial registration

This trial is registered as ISRCTN98970319.

Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: 13/152/02) and is published in full in *Health Technology Assessment*; Vol. 29, No. 40. See the NIHR Funding and Awards website for further award information.

Chapter 1 Introduction

Parts of this chapter are reproduced with permission from Wiseman *et al*.[1] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, even for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The text below includes additions and formatting changes to the original text.

Renal tract stone disease is very common, with a lifetime prevalence of approximately 10% across the world.[2] It mainly affects adults of working age and the incidence has been increasing over recent decades.[3,4] Approximately 50% of people with renal tract stones will experience symptoms, typically kidney pain, and about 25% of patients with stones will require active treatment.[5–7] Some people with stones can develop more serious problems including visible blood in the urine (haematuria), uncontrolled pain, infection, sepsis, impaired kidney function and kidney failure. Despite successful removal of the initial stone, there is a lifetime recurrence risk of 50% for the development of a further stone.[8] Kidney pain from stones (renal colic) is the most common cause of emergency admission to urology departments in the UK and, given the age group most commonly affected, it results in time off work and loss of economic activity.[9] The ongoing need for painkillers and the detriment to family, social and work activity reduces quality of life (QoL) and incurs additional costs.

Renal stones are a major burden on the NHS in the UK, resulting in over 88,000 upper urinary tract stone episodes and over 35,000 day case procedures carried out to remove stones in England in 2019–20.[3] Compared to the 5-year period from 2014 to 2015, the number of urinary tract stone episodes has increased by 2.2%, while the number of day case procedures has increased by 14.7% from just over 31,000.[10]

Stones commonly develop in the lower part (pole) of the kidney, accounting for up to 35% of cases.[11] There are currently three technologies available to treat lower pole kidney stones: extracorporeal shockwave lithotripsy (ESWL); flexible ureterorenoscopy (FURS) with laser lithotripsy and percutaneous nephrolithotomy (PCNL). The choice of treatment can be guided by stone size, likely stone composition, and the anatomy of the drainage system of the affected kidney. Clinician and patient preference and availability of equipment and expertise also plays a part.[12] Current evidence indicates that the success rate in terms of stone clearance differs between these technologies particularly in relation to stone size. These technologies are also distinct in terms of degree of invasiveness, anaesthetic requirement, treatment setting, number of procedures required to clear the stone, and type and rate of complications.[12,13]

Extracorporeal shockwave lithotripsy is non-invasive, has a low risk of complications, and does not require general anaesthesia. Current evidence suggests it has reasonable efficacy in terms of stone clearance for smaller lower pole stones (LPSs) at 3 months (63–74% clearance rate for stones ≤ 10 mm).[14] However, 3-month stone clearance rates for LPSs > 10 mm appear to be lower (23–56% for 11–20 mm stones, and 14–33% for 21–30 mm stones).[15] If the stone is not cleared, additional treatments may be required using either repeated ESWL or more invasive options, such as FURS or PCNL. Following ESWL, small residual stone fragments can be left in the kidney and may result in recurrent stone formation over time (20% at 5 years).[7,16]

Flexible ureterorenoscopy (with laser fragmentation) and PCNL are more invasive than ESWL, require a general anaesthetic, and carry a greater risk of complications.[17,18] A single FURS treatment appears to result in good clearance rate for stones up to 15 mm with repeat procedures or combined procedures often required for larger stones. PCNL is the most invasive treatment option and is associated with a higher risk of complications, but it also appears to result in the highest stone clearance rates, which are close to 100% for stones ≤ 10 mm, 93% for stones 11–20 mm, and 86% for stones 21–30 mm.[19] Stone clearance rates for FURS appear to lie between those of ESWL and PCNL.[20–26]

Copyright © 2025 Wiseman *et al*. This work was produced by Wiseman *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Rationale for the trials

Guidance issued by the European Association of Urology (EAU)[12] at the time of study design (in 2014), which was widely followed in UK clinical practice, recommended ESWL or FURS as first-line options for LPSs ≤ 10 mm. For stones > 20 mm, PCNL was recommended as the first-line treatment option. For stones between these two sizes, in the lower pole, the recommendation depended on whether there were unfavourable factors for ESWL, these being detailed as; shockwave resistant stones; the presence of a steep infundibulo-pelvic angle to the lower pole; a long (> 10 mm) lower pole calyx; or where the infundibulum is narrow (< 5 mm). If unfavourable factors were present then PCNL or FURS were the first-line recommended options, with ESWL being second-line. If unfavourable factors were not present, then ESWL or PCNL or FURS were all possible first-line treatment options. The EAU Guideline also commented that there remained considerable uncertainty regarding the management of LPSs, with each treatment option having advantages and disadvantages. The current EAU guidance remains unchanged.[27] Subsequently published National Institute for Health and Care Excellence (NICE) guidance in 2019[28] recommends ESWL for a renal stone < 10 mm but to consider FURS if there are contraindications for ESWL, if a previous course of ESWL has failed or because of adverse intrarenal anatomy. PCNL is regarded as an option to consider if ESWL and FURS have failed to treat the stone. For larger renal stones between 10 and 20 mm, FURS or ESWL are considered equal options, with PCNL the next step if these interventions have failed. Like the EAU guidelines, PCNL is the first-line treatment for a renal stone larger than 20 mm, including staghorn stones, but FURS can be considered if PCNL is not an option.

A Cochrane review and meta-analysis (2014) of randomised controlled trials (RCTs) compared ESWL with either FURS or PCNL for the treatment of stones located anywhere within the kidney, not solely in the lower pole.[13] The review concluded that PCNL had a better stone-free rate (SFR) than ESWL at 3 months [relative risk (RR) 0.39, 95% confidence interval (CI) 0.27 to 0.56], whereas FURS appeared to have similar SFRs to ESWL (RR 0.91, 95% CI 0.64 to 1.30). They included 5 RCTs (338 participants); however, only 3 focused on LPSs. Of these 3 RCTs (160 participants), 2 compared ESWL against PCNL, 1 for stones up to 30 mm[15] and 1 for stones up to 20 mm.[29] The third compared ESWL with FURS for LPSs ≤ 10 mm.[30] The review concluded that the included trials were small and of low methodological quality. The authors had planned to undertake subgroup analyses by size and location of stone, but this was not performed 'because of insufficient data'.[13]

A systematic review performed by some of the PUrE investigators,[31] focusing solely on stones located in the lower pole of the kidney, included trials comparing PCNL with FURS (a comparison not considered in the Cochrane review). This review identified 4 additional relevant trials involving 408 participants[32–35] and we undertook subgroup analyses by stone size (< 10 and 10–20 mm). Taking the seven trials involving participants with lower pole kidney stones as a whole, the Grading of Recommendations Assessment, Development and Evaluation (GRADE) certainty of evidence ratings for the outcome of SFRs indicated they were of 'moderate' quality. The meta-analyses found PCNL and FURS produce significantly higher SFRs than ESWL for LPSs ≤ 20 mm at 3 months. Combining 2 RCTs (155 participants), SFRs for those participants with stones ≤ 20 mm were higher following PCNL than ESWL (RR 2.04, 95% CI 1.50 to 2.77). Combining 5 RCTs (508 participants) showed that FURS resulted in a higher stone clearance rate compared to ESWL (RR 1.31, 95% CI 1.08 to 1.59). However, in a subgroup meta-analysis combining 3 studies (300 participants) for stones ≤ 10 mm, the advantage of FURS over ESWL was less although still statistically significant (RR 1.11, 95% CI 1.03 to 1.19). One RCT (n = 93),[15] which reported SFR categorised by stone size for PCNL versus ESWL, found that the degree of superiority of PCNL was lower for stones sized ≤ 10 mm compared to those sized > 10 to ≤ 20 mm (RR 1.56, 95% CI 1.11 to 2.21; RR 4.02, 95% CI 1.98 to 8.18, respectively). One small RCT (n = 28)[35] showed that the stone free rates were higher when treated with PCNL compared to FURS (11-25 mm LPSs), but with considerable uncertainty around this estimate of effect.

Few of these trials reported data on patient-centred outcomes (such as QoL) or on resource use and none on cost-effectiveness. Pearle[30] suggested that ESWL gave better QoL, shorter convalescence (days to 100% recovered), and had fewer analgesic requirements than FURS (participants had LPSs ≤ 10 mm). Conversely, Singh[36] reported significantly higher participant satisfaction with FURS and comparable convalescence (time to return to routine activity) after having three or fewer ESWL sessions (participants had stones sizes of 10–20 mm). Convalescence was shorter after just a single ESWL session. There were conflicting data on patients' willingness to undergo the procedure again. In one trial, participants[30] favoured ESWL whereas in another[36] FURS was preferred. ESWL (one session) was associated with a shorter hospital stay than either PCNL[15] or FURS.[30] One trial also suggested shorter treatment duration for ESWL (one session) compared to FURS.[30]

In summary, there is some evidence to inform estimates of the relative clinical effectiveness (based on SFR) of ESWL, FURS and PCNL in the treatment of LPSs and to guide clinical practice, but for most outcomes there is only moderate or low certainty in the evidence.[31] There is sparse evidence on the impact of these treatments on patient-reported health status and QoL outcomes (such as severity and duration of pain after intervention), their care pathway (such as the need for additional treatments), and resource use. The EAU Guideline, which differs somewhat from the previously mentioned UK-specific 2019 NICE guidance,[28] recommends either ESWL or FURS for LPSs under 10 mm, FURS or PCNL for LPSs between 10 and 20 mm, where there are unfavourable factors for ESWL, and any of the three treatment options if there is an absence of unfavourable factors for ESWL. This Percutaneous nephrolithotomy, flexible Ureterenoscopy and Extracorporeal shockwave lithotripsy for lower pole stone (PUrE) study aimed to provide robust evidence on health status, QoL, clinical outcomes and resource use to both the NHS and society to close this gap in evidence. The results aimed to benefit patients, clinicians, and the NHS and to inform provision, guidance, and decision-making about which of the competing interventions – ESWL, FURS or PCNL – is the most suitable (clinically effective and cost-effective) for the treatment of people with lower pole kidney stones of varying sizes.

Trial objectives

The aim of the PUrE study was to determine which of ESWL, FURS and PCNL offer the best treatment outcomes in terms of clinical effectiveness and cost-effectiveness for people with lower pole kidney stones seeking treatment within the UK NHS. The research question addressed was: In people requiring treatment for LPSs of the kidney, does FURS result in better health than standard treatment with ESWL or PCNL according to stone size, and is it cost-effective for the UK NHS?

The PUrE study was two separate RCTs run in parallel – the first (referred to as RCT1) for LPSs ≤ 10 mm in maximum dimension and the second (RCT2) for LPSs > 10 and ≤ 25 mm.

PUrE was designed to determine the clinical effectiveness and cost-effectiveness of FURS as the first treatment option in a treatment pathway in comparison to ESWL as a first treatment option in a treatment pathway for stones ≤ 10 mm in maximum dimension, or to PCNL as a first treatment option in a treatment pathway for stones > 10 and ≤ 25 mm in maximum dimension with respect to:

- patient-reported health status measured as area under the curve (AUC) comprised of the score from the EuroQol-5 Dimensions, five-level version (EQ-5D-5L) questionnaire completed at multiple time points up to 12 weeks post intervention
- incremental cost per quality-adjusted life-years (QALYs) at 12 months post randomisation
- successful stone clearance at 12 weeks
- further interventions required to treat stones up to 12 months after randomisation and
- treatment-related harms experienced up to 12 months after randomisation.

The null hypotheses being tested were:

1. The use of FURS to treat lower pole kidney stones ≤ 10 mm will not be different to ESWL as assessed by the EuroQol-5 Dimensions (EQ-5D) score AUC up to 12 weeks post treatment.
2. The use of FURS to treat LPSs of the kidney > 10 and ≤ 25 mm will not be different to PCNL as assessed by the EQ-5D score AUC up to 12 weeks post treatment.

Chapter 2 describes the PUrE study design and methods. *Chapters 3* and *4* present the clinical effectiveness results from RCT1 and RCT2, respectively. The health economic evaluation is presented in *Chapter 5*, and in *Chapter 6* we discuss the results and present conclusions.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Chapter 2 Trial design and methods

Parts of this chapter are reproduced with permission from Wiseman *et al.*[1] and McClinton *et al.*[37] These are Open Access articles distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, even for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

We prospectively registered PUrE on a publicly available website on 11 November 2015 as International Standard Randomised Controlled Trial Number (ISRCTN) ISRCTN98970319.

Trial design

We have published the PUrE study protocol[37] in an open access journal.

PUrE comprised two separate pragmatic, multicentre, individual patient-randomised, open-label superiority RCTs, with an internal pilot phase, to determine which of ESWL, FURS and PCNL offers the best treatment outcomes in terms of clinical effectiveness and cost-effectiveness for people with lower pole kidney stones, seeking treatment within the UK NHS.

We compared the care pathways that started with two active interventions in each RCT.

RCT1: This compared FURS with ESWL for patients with lower pole kidney stones of maximum dimension ≤ 10 mm.

RCT2: This compared FURS with PCNL for patients with lower pole kidney stones of maximum dimension > 10 and ≤ 25 mm.

Participants

Our participants were adults (≥ 16 years), requiring treatment for a stone in the lower pole of either kidney. Adults with a stone ≤ 10 mm, confirmed by non-contrast computed tomography of the kidneys, ureters and bladder (CTKUB), were considered for RCT1. Adults with a stone > 10 and ≤ 25 mm, confirmed by CTKUB, were considered for RCT2. Our aim was to recruit 522 participants to RCT1 and 522 participants to RCT2.

Outcomes

Our primary clinical outcome was health status AUC calculated from multiple measurements of EQ-5D-5L[38] up to 12 weeks post intervention. Our primary economic outcome was the incremental cost per QALYs gained at 12 months post randomisation. The trial design is summarised in *Figure 1*.

Patients were recruited to PUrE and were followed-up to 12 months post randomisation or 12 weeks post intervention, whichever was later.

Population

The PUrE population was adults (≥ 16 years old), with a stone ≤ 25 mm in the lower pole of the kidney confirmed by CTKUB, who presented to NHS urology departments and were considered suitable for intervention.

Potential participants and clinicians had to agree that active intervention was appropriate to remove the LPS and that the potential participant was suitable for either of the interventions for their specific stone size.

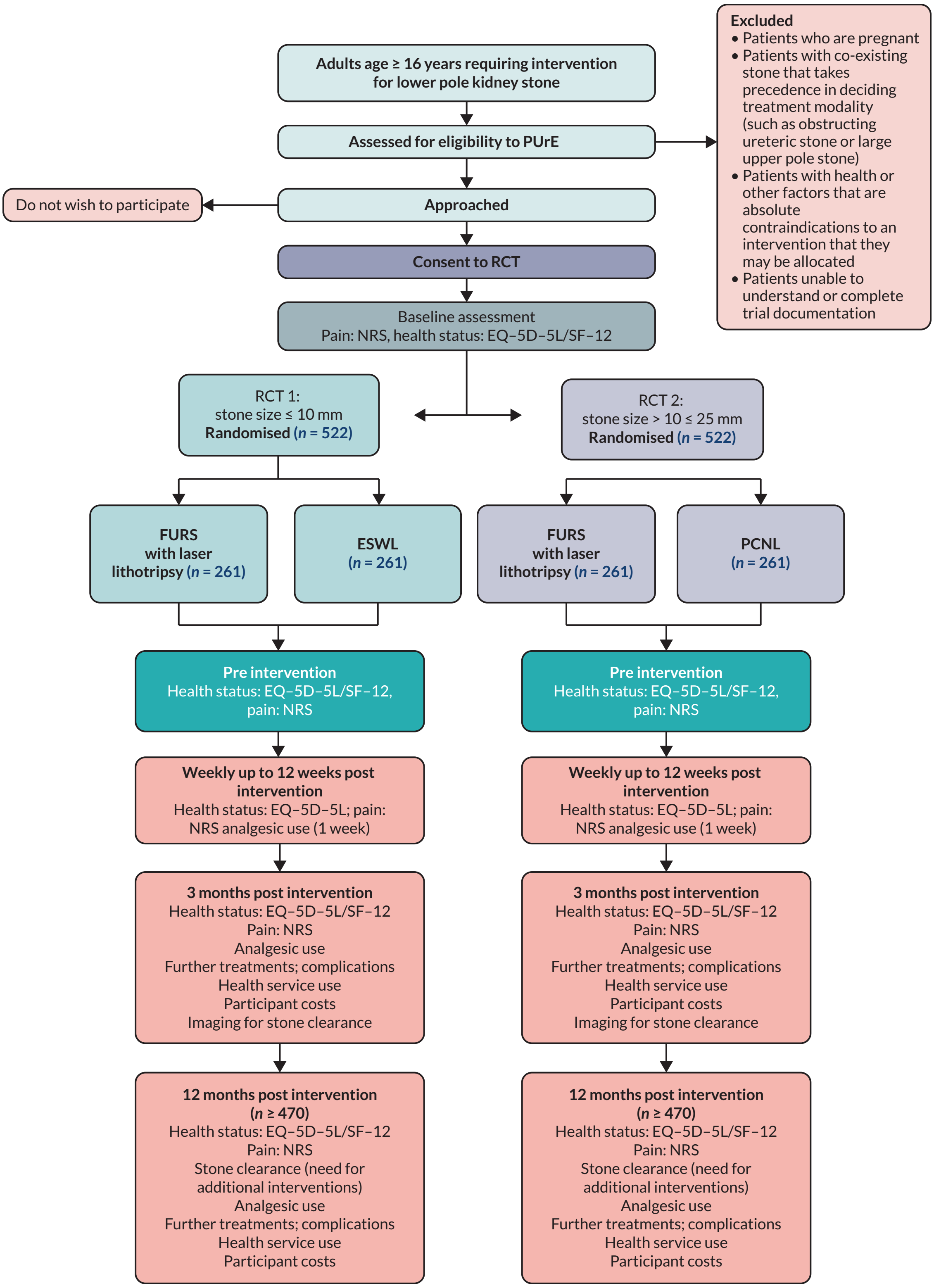


FIGURE 1 PUrE RCTs trial design. SF-12, 12-Item Short Form Survey; EQ-5D-5L, EuroQol-5 Dimensions, five-level version; NRS, Numeric Rating Scale; RCT, Randomised Controlled Trial.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Patients with multiple renal tract stones were eligible to participate in one of the PUrE RCTs provided the stone in the lower pole measured ≤ 25 mm in maximum dimension and that the LPS was the index stone for treatment and follow-up.

The following inclusion criteria were used to identify eligible participants.

Inclusion criteria

- Adults ≥ 16 years of age.
- LPS ≤ 25 mm in maximum dimension with decision to treat that stone.
- Presence of stone previously confirmed by CTKUB.
- Able and willing to undergo either treatment for specified stone size.
- Capacity to give informed consent to participate in trial, which includes adherence to trial requirements.

Exclusion criteria

- Pregnancy.
- Patients with co-existing stone that takes precedence in deciding treatment modality (such as obstructing ureteric stone or large upper pole stone).
- Patients with health or other factors that are absolute contraindications to an intervention that they may be allocated.
- Patients unable to understand or complete trial documentation.

Identification

Sites

United Kingdom NHS hospital sites were identified via the British Association of Urological Surgeons (BAUS) from urology centres with a high rate of LPSs and an ability to deliver all trial interventions. These sites were identified using a feasibility questionnaire, which was reviewed and approved by the PUrE Chief Investigator (CI) and the Trial Office Team. Prior to starting recruitment, a site initiation visit was conducted in person at each site where the local research teams, including the principal investigator (PI), were trained in site processes and procedures by members of the Trial Office team.

Participants

The local research teams at the participating sites identified eligible participants. There were numerous routes to identifying potentially eligible patients, for example, eligible patients could be referred to urology departments because their LPS had been found opportunistically by abdominal imaging, or during investigation of urinary tract symptoms or they could present as an emergency with loin pain or infection.

The PUrE study was publicised at sites by the local research team using posters approved by the Research Ethics Committee (REC).

The local research team sent a flyer to eligible patients with information about the PUrE RCT, for which they were eligible, and at an appropriate point during their initial assessment gave patients the patient information leaflet (PIL) for RCT1 or RCT2 as appropriate.

The PILs explained the purpose and need for the RCT for which the patients were eligible, the RCT interventions and their risks, and the information the patients would be asked to complete during their participation. We developed the two PILs in conjunction with the BAUS Section of Endourology patient group.

The local research team gave patients adequate time to consider participating in the study. If the patient chose to, they could consent to participate in PUrE at the time of their initial assessment. Alternatively, they could choose to take the study information away with them for further consideration, in which case the local research team contacted the patient after at least 24 hours to determine their interest. If the patient was willing to participate in the trial, the local research team arranged with them to complete the appropriate PUrE (RCT1 or RCT2) consent form.

The local research team recorded brief details of the patients they assessed for inclusion on their site screening log. The local research team completed, and entered on the database, brief anonymised details, with reasons if given, for those patients who were assessed for inclusion but not recruited.

Recruitment and consent

All staff involved in recruitment and consent had evidence of up-to-date good clinical practice (GCP) training and had received the appropriate PUrE study training. The local research team ensured that any questions the patient had about the study were answered satisfactorily and asked them to complete the PUrE consent form if they confirmed that they were willing to participate.

Participants who could not give informed consent (e.g. due to incapacity) were not eligible to participate in PUrE. The consent form included consent to inform their general practitioner (GP) that they were taking part in PUrE, and consent for the participant to be approached for further studies on kidney stones, and for long-term follow-up through their local and central NHS clinical records after their active trial participation had finished.

Randomisation and concealment of allocation

Eligible and consenting participants were assigned to one of two treatment groups using simple randomisation (computerised random numbers). A delegated member of the local research team used the telephone 24-hour interactive voice response randomisation application or the web-based application hosted by the fully registered UK Clinical Research Collaboration, Clinical Trials Unit (CTU) at the Centre for Healthcare Randomised Trials (CHaRT), and Health Services Research Unit (HSRU) at the University of Aberdeen to assign the participant to their treatment group. Patients with a stone ≤ 10 mm, entered into RCT1, were randomised 1 : 1 to either FURS or ESWL. Participants with a stone > 10 and ≤ 25 mm, entered into RCT2, were randomised 1 : 1 to either FURS or PCNL.

The local research team informed the participants of their treatment allocation. The participant then followed the standard care pathway for the allocated intervention. In this pragmatic study, the only trial-specific interventions apart from randomised treatment allocation were the participants' completion of outcome questionnaires.

Interventions evaluated

Experimental

- FURS as the first treatment option in a treatment pathway.

Standard

- RCT1 – ESWL as the first treatment option in a treatment pathway
- RCT2 – PCNL as the first treatment option in a treatment pathway.

At the time of funding, all three study interventions were (and still are at the time of reporting), in general use by Urology departments throughout the UK NHS. This study aimed to evaluate the interventions in a standard NHS setting for the results to be generalisable to current routine care in the UK. In line with this aim, all procedures were delivered in NHS facilities and supervised by NHS staff trained and competent in the procedures. All participants were under the care of a named consultant urologist.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

The surgical interventions, FURS and PCNL, were carried out by a trained urologist, or by a trainee urologist under the supervision of a senior urologist, supported by the standard team of ward and theatre staff and radiographers. In some centres, a specialist uro-radiologist also assisted with the procedures, particularly with access to the stone.

The ESWL intervention was delivered using any device approved for this purpose, including both fixed-site and mobile lithotriptors. Delivery of the treatment was according to local practice by staff trained in the procedure; typically, a radiographer and nurse, and supervised by a urologist.

The specific technique and equipment used for FURS, ESWL and PCNL differed in fine detail between different surgeons and departments. Our PUrE protocol did not mandate the use of any specific detailed technical method for each intervention under study but as part of trial initiation of each site the standard procedure including equipment used for FURS, ESWL and PCNL for that site was recorded on a trial proforma.

For FURS, a thin (3 mm diameter) flexible endoscope (ureteroscope) is passed into the kidney via the natural urinary passages (urethra, bladder and ureter) and is used to directly see the stone. A laser fibre is then passed through the working channel of the ureteroscope and laser energy used to fragment the stone within the kidney. Larger fragments can be retrieved with a wire basket device passed through the working channel, while smaller fragments (< 2 mm) may be left to pass spontaneously. Generally, the patient will pass remaining fragments in the urine during the weeks following the procedure. The procedure is usually performed as a day case (or single overnight stay) and usually requires general or spinal anaesthesia. Patients often receive a single dose of antibiotic, at the start of the procedure to reduce the risk of infection. The duration of the operation depends on size of stone but is typically 0.5–1.5 hours. A temporary ureteric stent may be placed at the end of the procedure to protect against blockage of the ureter caused by postoperative reactive oedema or from the passage of residual stone fragments. The operating surgeon monitors the progress and degree of stone clearance during the procedure and this may be checked with imaging. Possible harms of the procedure are urinary tract infection, bleeding and damage to the urinary system, which may require a more prolonged period of stenting. The stent itself can cause pain and urinary symptoms, such as increased urinary frequency and haematuria.

For the purposes of the PUrE study, FURS treatment was expected to be a single procedure in the great majority of cases. However, an additional FURS procedure was considered part of the original FURS treatment strategy in cases of technical complexity or larger stones as long as it took place with 6 weeks of the initial FURS procedure. Any additional procedures beyond this timeframe were recorded separately for study purposes. Once stone clearance was confirmed, the ureteric stent was removed as an outpatient procedure with local anaesthetic. Placement and removal of the stent was considered part of the FURS treatment strategy for the purposes of the PUrE study.

Extracorporeal shockwave lithotripsy involves the generation of an external acoustic (sound) pulse, called a shockwave, outside the body, which is then focused onto the kidney stone, through the patient's flank skin, causing it to fragment. Stone fragments pass down the urinary tract spontaneously, which may take a few weeks. It is routinely performed in an outpatient setting with analgesia, with or without sedation as required. Patients may be given a single dose of antibiotic at the start of the procedure if there is thought to be a higher-than-normal risk of getting an infection afterwards. Each session delivers a set number of shocks (this varies depending on the size of the stone) and stone fragmentation is monitored during the procedure and then by a plain X-ray (or other imaging as standard) taken at a follow-up appointment at approximately 3 weeks.

For the PUrE study up to two separate ESWL treatments were considered as part of the initial ESWL treatment pathway. These were to take place within an 8-week period and each episode was to be recorded separately for trial purposes. The treating urologist could however have decided that further ESWL was not appropriate if stone fragmentation was insufficient following the first or second session. The first session of ESWL was taken as the initial treatment point for the purposes of timing of outcome assessments. If the participant had additional ESWL treatments, the third and any subsequent treatments were not considered to be part of the initial ESWL treatment pathway. Possible harms of ESWL include urinary tract infection, visible bleeding in the urine and blockage of the ureter by the stone fragments. There is also a small risk of bruising surrounding the kidney.

Percutaneous nephrolithotomy is a surgical procedure to remove stones from the kidney by a direct approach. A small (10 mm) incision is made on the skin overlying the kidney, through which a needle is passed into the urine collecting tube system of the kidney. This can be guided either by simultaneous ultrasound imaging of the kidney, or fluoroscopy (X-ray) or both. This can be assisted by the preliminary telescopic placement of a tube through the urethra, bladder and ureter via which contrast fluid can then be injected into the collecting system to guide the needle passage through the skin and into the kidney. Placement of the needle is planned using available imaging (typically a CTKUB) in order to give the best access to successfully remove the stone. For a stone in the lower pole of the kidney, this is usually into the lowermost part of the collecting system. Once the needle is satisfactorily placed, a flexible guide wire is passed into the collecting system of the kidney and used to guide stretching (dilatation) of the needle track to make it wide enough for a hollow rigid access sheath to be passed, creating a channel between the skin and the urine-collecting system of the kidney. A rigid metal telescope (nephroscope) can then be inserted down this channel into the kidney's collecting system to see the stone, which can be retrieved whole using graspers or fragmented using a variety of energy delivery devices; most commonly an ultrasonic probe or pneumatic device or combination device incorporating both energies. After the operation, the kidney may be drained for a period by a nephrostomy tube which is placed either through the access site, or internally using a stent, which is a drainage tube that lies between the kidney and the bladder. In addition, a urinary catheter may be inserted to drain the bladder for a short period after the procedure. The operation is performed under general anaesthesia with a typical duration of one to three hours depending on complexity, and patients may need to stay in hospital for up to a few days. Patients frequently receive antibiotic treatment at the start of the procedure, and this may be continued for a few days after, if there is active infection. The nephrostomy and catheter drainage tubes are usually removed after 24–48 hours without need for further anaesthesia; the stent, typically after 1–2 weeks using a flexible telescope passed into the bladder. Stone clearance is monitored during the procedure and, if necessary, by a plain X-ray (or other imaging) before discharge from hospital. Possible harms include urinary infection, bleeding (which may be severe), and inadvertent puncture of other organs. For the PUrE study, a single PCNL treatment was expected to completely remove stones up to 25 mm.

Apart from the randomised allocation of the initial intervention and participant completion of questionnaires, PUrE did not seek to change or impose any specific protocol regarding the clinical management of participants recruited at each trial site. It was anticipated that at 8 to 12 weeks post intervention, participants would receive imaging to assess stone clearance in accordance with usual standard of care.

We did however record relevant aspects of the participant care during their involvement in the trial up to 12 months after randomisation (or up to 12 weeks post intervention if this time point occurred later) and obtain patient-reported outcome measures (PROMs). In particular, trial participants undergoing any of the stone treatments may require further interventions either to correct harms arising from the initial intervention or because of incomplete stone clearance by the initial intervention. The circumstances, nature and outcome of these additional procedures were recorded, and PROMs obtained where possible before and after each additional intervention to assess effects on health status.

Blinding

This PUrE study was open label. We could not blind participants, clinical staff or the central trial team due to the differing nature of the interventions being tested. Sites were instructed to collect baseline measures (particularly those reported by study participants), before randomisation.

Data collection

We collected outcome data throughout the trial for each participant from consent until 12 months following randomisation or until 12 weeks post intervention, whichever was later. See *Table 1* for the source and timing of measures.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 1 Source and timing of measures

		Timing					
			Intervention (PCNL or first session ESWL/FURS)				
				Weeks post intervention			
Outcome measure	Source	Baseline[a]	Pre intervention	1–11	12	Additional intervention (pre and post if > 12 weeks) or treatment-related hospitalisation	Post randomisation 12 months
Health status (EQ-5D-5L)	PQ	✓	✓	✓	✓	✓	✓
Pain	PQ	✓	✓	✓	✓	✓	✓
Health profile (SF-12)	PQ	✓	✓		✓		✓
Use of analgesics	PQ	✓	✓	✓	✓	✓	✓
Stone clearance (imaging)	CRF				✓[b]		✓
Additional interventions received	CRF and PQ				✓	✓	✓
Complications	CRF and PQ				✓	✓	✓
NHS primary and secondary healthcare use	CRF and PQ		✓		✓	✓	✓
Participant costs	PQ				✓		✓

CRF, case report form; PQ, participant questionnaire.
a Baseline was after the participant had given informed consent but before randomisation.
b Imaging performed at 8 to 12 weeks post treatment.

Baseline

Participants were asked to complete the baseline questionnaire [the self-report measures; EQ-5D-5L[38–44] Short Form questionnaire-12 items (SF-12);[45,46] pain score on a numeric rating scale (NRS) and use of analgesics] by the local research team prior to randomisation.

The local research team also completed a baseline clinical case report form (CRF) for each participant. The baseline CRF recorded the participant's name and full contact details, date of birth, and gender of the participant, stone dimensions, skin-to-stone distance and stone density in Hounsfield units. The local research team entered the data from the baseline questionnaire and CRF into the study database (a web-based trial management and data collection platform).

After randomisation the planned treatment date was added to the CRF and database.

Intervention and follow-up

Participants were asked to complete the questionnaires which included the EQ-5D-5L, pain (NRS) and use of analgesia, at fixed and variable time points during their trial participation.

The fixed time points were: just prior to initial intervention (pre intervention), weekly up to 12 weeks after initial intervention on their treatment pathway (first session of either ESWL, FURS or PCNL) and at 12 months post randomisation.

The variable time points after 12 weeks were just prior to, and 1 week after any additional intervention, including planned additional sessions of ESWL, removal of stent, or during any other hospital admissions related to treatment of their lower pole kidney stone, such as admissions for pain control or infection. Participants were also provided with the additional EQ-5D-5L questionnaire in a take-home study pack to complete at their discretion through their participation in the trial.

In addition, at 12 weeks post initial intervention, and at 12 months post randomisation, participants were asked to complete the SF-12 and questions relating to their primary and secondary care use and their associated travel.

We offered participants a choice of methods of questionnaire completion and reminders including postal, e-mail, web-based, telephone, SMS text with a link and in-person with a member of the local research team when in hospital. We used questionnaire reminders to facilitate questionnaire return. For the earlier time points, this was a text message or e-mail on the day that the questionnaire was due to be completed. For the later time points, 12 weeks and 12 months, this reminder was sent approximately 2 weeks, and again if required 4 weeks, after the questionnaire was due. Participants were sent a £10 multi-retailer shopping voucher as a token of appreciation for completion and return of the questionnaires at 12 weeks post intervention and 12 months post randomisation.

The local research teams completed intervention CRFs on the details and outcome following the initial and second intervention (if part of the treatment strategy). Capturing the date of the first treatment in a treatment pathway in a timely fashion was key to triggering the participants fixed-point weekly questionnaires. For participants where conservative management was agreed on their pathway, we used the date on which conservative management or the care plan was agreed to start their 12-week post-intervention period ('nominated treatment date'). The local research team captured additional treatment sessions and stent removal at 8 to 12 weeks post intervention, after any additional stone-related treatments and admissions due to pain or infection up to 12 months post randomisation.

To measure the secondary clinical outcome of stone clearance, participants had kidney imaging between 8 and 12 weeks after treatment according to clinical need and participant convenience. We also asked local site clinical teams (radiologist/urologist) to state whether there was complete clearance of the target stone from the urinary tract (defined as no further action or observation required for that stone); acceptable clearance, where observation is required but no intervention planned; and unacceptable clearance, where further intervention would be required. We stated a preference for imaging by CTKUB during site set-up and initiation, but renal ultrasound or plain X-ray were acceptable according to patient preference, safety, and local practice.

Data management

The local research team entered the data that were collected locally onto the study database. The Trial Office worked closely with the local research teams to ensure data was as complete and accurate as possible, both through chasing missing data and by querying inconsistencies in the data.

The local research team administered the baseline and the pre-intervention participant questionnaires and the CRFs. The Trial Office administered the weekly post intervention and the 12-month postrandomisation participant questionnaires and reminders. Initially, the local research team provided participants with the 1- and 2-week questionnaires after their intervention to take home, complete and return to the Trial Office. However, it became evident that some sites were unable to provide this for the participants and therefore the Trial Office administered these questionnaires. The local research team and the Trial Office administered the variable time point questionnaires. The Trial Office entered the data from the paper questionnaires that participants returned to the Trial Office.

The Trial Office carried out data accuracy checks on a sample of data entered by sites and by the trial office as per the trial data-monitoring plan.

Participant change of status

Participants were free to withdraw their consent or change their mind about participation at any time. We collected outcome data from participants' medical records (local research team via CRF completion), unless the participant specifically withdrew from such data collection. All data collected up to the point of participants declining further follow-up were retained and used in the analysis. All participants were followed up as per the trial protocol even if they did not get their allocated treatment.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Outcomes

All the EQ-5D-5L scores were converted to the EuroQol-5 Dimensions, three-level version (EQ-5D-3L) scale using the crosswalk as approved by NICE[47] – herein referred to as the EQ-5D score.

Primary outcome

The primary outcome (PO) was patient-reported health status measured using EQ-5D score AUC from intervention to 12 weeks post intervention. The AUC was based on EQ-5D-5L completion at fixed time points: prior to initial intervention (first session of ESWL, FURS or PCNL), and weekly up to 12 weeks after initial intervention. For participants who had conservative management, we used the date when conservative management or the care plan was agreed, to start their 12-week post-intervention period ('nominated treatment date').

Economic outcome

The primary economic outcome was incremental cost per QALY gained at 12 months post randomisation based on the estimated NHS costs and participant responses to the EQ-5D-5L.

Secondary outcomes

Patient-reported

- Patient-reported health status measured using EQ-5D score AUC from randomisation to 12 weeks post intervention.
- Severity of pain (NRS); weekly at 1 to 12 weeks post intervention and at 12 months post randomisation.
- Generic health profile as measured by the SF-12, at 12 weeks and 12 months.
- Use of analgesia weekly at 1 to 12 weeks post intervention and at 12 months post randomisation.

Clinical

- Stone clearance measured at between 8 and 12 weeks post initial intervention using renal imaging (CTKUB preferred but plain X-ray or ultrasound acceptable).
- Maximum dimension of the largest fragment of the treated stone in mm.
- Need for additional treatment (carried out or planned) at 12 weeks post initial treatment and 12 months post randomisation.
- Complications during initial intervention.
- Post-intervention complications at 12 weeks (categorised by Clavien–Dindo classification) post treatment and up to 12 months post randomisation.

Economic

- NHS primary and secondary care resources used and their costs.
- Patient costs (out of pocket) including time off work up to 12 months post randomisation.

Stone clearance was further defined as complete clearance of the target stone from the urinary tract (defined as no further action or observation required for that stone); acceptable clearance, where observation was required but no intervention planned; and unacceptable clearance, where further intervention would be required.

Safety and breaches

PUrE involved procedures for treating LPSs which are all well established in current NHS clinical practice. Adverse effects may occur during or after any type of surgery. We did not record adverse events unless they met the criteria for a serious adverse event (SAE). SAEs were monitored and assessed by the local PI or their delegate against the criteria for seriousness, causality (related/unrelated) and expectedness and reported to the Trial Office via the trial database.

Within PUrE, 'related' was defined as an event that occurs as a result of a procedure required by the protocol, whether it was either (1) the specific intervention allocated at randomisation, or (2) it was administered as an additional intervention as part of normal care.

We did not consider hospitalisations for treatment planned prior to randomisation or for elective treatment of a pre-existing condition as SAEs, although complications occurring during such hospitalisation were reported as appropriate.

Hospital visits (planned or unplanned) and SAEs associated with further interventions or complications of trial treatment (e.g. expected SAEs) due to the LPS were recorded as an outcome, but not reported as an SAE. Planned primary care or hospital visits for conditions other than those associated with the LPS were not collected or reported.

All deaths for any cause (related or otherwise) were recorded on the SAE form.

Expected adverse events

The following events were considered as potentially expected for each intervention:

Flexible ureterorenoscopy: Mild burning pain passing urine; visible bleeding on passing urine; pain and frequent urination associated with a stent; kidney pain; urinary tract infection needing antibiotic treatment; damage to the ureter; leakage of urine into the body; scarring or stricture of the ureter.

Extracorporeal shockwave lithotripsy: Visible blood in urine; pain in the kidney; urinary tract infection; bruising of the skin; stone fragments stuck in ureter; kidney bruising; damage to the pancreas or lungs by the shockwaves requiring further treatment.

Percutaneous nephrolithotomy: Visible blood in the urine; fever (high body temperature); pain and frequent urination associated with a stent; severe kidney bleeding; damage to adjacent organs, such as lung, bowel, spleen or liver; damage to kidney or infection needing further treatment; leakage of urine into the body; leakage of irrigating fluids into the body.

It was a requirement to report any SAEs that were deemed related and unexpected to the sponsor within 24 hours of receiving the signed SAE notification. Such SAEs were also reported to the REC within 15 days of the CI becoming aware of the event. All related SAEs were summarised and reported to the appropriate authorities within their regular progress reports.

Breaches

Sites were asked to report potential breaches of trial protocol or GCP to the trial office. Trial office staff could also report potential breaches. Five breaches were recorded within the study and these are summarised in *Appendix 1*, *Table 32*. The CI and sponsor assessed these breaches as non-serious.

Sample size

The PO was the AUC measured from multiple completion of the EQ-5D-5L by each participant up to 12 weeks post initial intervention (first of allowed session of FURS, ESWL or PCNL). Each RCT was designed to detect a target difference of 0.3 standard deviations (SDs), with 90% power and 5% two-sided alpha. This required outcome data on 235 participants per group (470 total). Such a difference in generic health status was considered clinically relevant and in terms of treatment effect size, in the small to medium range as observed in other clinical studies.[48,49] To allow for completely missing outcome data for approximately 10% of participants, and therefore for whom the AUC could not be calculated, we planned to randomise 522 participants in both RCT1 and RCT2, giving a total trial population of 1044 participants.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Statistical analysis

The trial analysis followed a statistical analysis plan (SAP). The SAP was drafted by the trial statistician in consultation with the Project Management Group (PMG). It was reviewed by the Trial Steering Committee (TSC) and Data Monitoring Committee (DMC) before analysis started. RCT1 and RCT2 were analysed separately. Full definitions of the derived outcomes are given in *Appendix 2* (see *Tables 33–35*).

We based the main statistical analyses on all participants as randomised, intention to treat (ITT), irrespective of subsequent compliance with the treatment allocation. We used a treatment policy approach for our PO estimand which ignores any intercurrent events because the primary question of interest in each trial was a comparison of treatment pathways. In RCT1 there was a longer waiting time for one of the treatments. We tried to account for this by adjusting for waiting time, using a hypothetical strategy approach that attempts to answer the question: which pathway is best if there is no difference in waiting times.

We conducted a single analysis that took place following completion of trial follow-up.

We summarised baseline and follow-up data using appropriate descriptive statistics and graphical summaries.

Primary outcome derivation and analysis

The PO was generated for each participant using the trapezoidal rule. Data for participants who missed a scheduled time point was estimated using a multiple imputation approach to make use of partial outcome data. See *Appendix 2* for the derivation of AUC in general using validated Stata code (package called eq5dmap) approved by NICE[47] in the UK.

To be included in the primary analysis, participants needed responses at three specific time epochs described in *Appendix 2*, along with notes for handling intervening missing data.

The PO was analysed using linear regression with adjustment for centre (using robust variance estimates) and baseline EQ-5D score. Additional analysis adjusted for pre-intervention EQ-5D score and waiting time between randomisation and initial intervention.

Secondary outcome analysis

Secondary outcomes were analysed using the appropriate generalised linear models with adjustment for baseline variables as appropriate. For example, categorical variables were analysed using ordinal regression with robust error variance accounting for centre. Binary variables were analysed using a log-binomial model with robust error variance accounting for centre. Secondary outcomes that were measured repeatedly were analysed using mixed models, with a random effect for participant and a fixed (nominal) effect for time, treatment and design covariates, and modelled with robust error variance accounting for centre.

All treatment effects sizes were summarised by between-group estimates and 95% CIs from the appropriate models. All analyses were carried out using Stata (version 17, StataCorp. *Stata Statistical Software: Release 17*; 2021) (StataCorp LP, College Station, TX, USA).

Subgroup analysis

Subgroup analyses explored the possible modification of treatment effect by important factors; [participant body mass index (BMI), stone size (maximum dimension), stone density on CTKUB (Hounsfield units) and skin-to-stone distance]. Analysis was done by including a subgroup-by-treatment interaction in models. The CIs were set at 99% for this exploratory assessment.

Within group exploratory analysis

We explored within each allocated group whether technical factors modify the treatment effect [access sheath vs. no access sheath and digital vs. non-digital instrument (FURS); fixed site versus mobile device (ESWL); calibre of access tract (PCNL)]. This used treatment-by-factor interactions in the model and were classified as exploratory analyses.

Economic evaluation

We conducted a 'within-trial' and a model-based economic evaluation. We describe these in detail in *Chapter 5*.

Ethics approval and monitoring

PUrE received favourable ethics opinion from the North of Scotland REC on 12 November 2015 (REC reference number 15/NS/0113, IRAS ID 188563) and the Health Research Authority on 9 June 2016. Each participating NHS organisation reviewed the study and gave confirmation of capacity and capability [or research and development (R and D) management approval] before the study commenced at site.

Sponsorship

The University of Aberdeen and NHS Grampian co-sponsored PUrE.

Management of the trial

The trial management team (the Trial Office), consisting of the Trial Managers, a Data Co-ordinator and the CI, were based within CHaRT. The trial management team took responsibility for the day-to-day transaction and coordination of trial activities and processes, including approvals, site set-up and close down, plus day-to-day support for the recruiting sites, provision of approved trial paperwork, administration of participant questionnaires and data entry, chasing and checking using the trial web data entry portal. This trial management team worked with the wider PUrE team, the CHaRT programmers, the trial statistician, health economist and grant holders to deliver the PUrE study.

The PUrE web portal for data collection and trial management was developed and maintained by the programming team within CHaRT to a specification agreed with the trial management team.

Each recruiting site was led by a local PI. The PIs in most cases were supported by research nurses, trial co-ordinators or dedicated staff (the local research team), who were responsible for all aspects of the local organisation, including recruitment of participants, delivery of the interventions and notification of any problems or unexpected developments during the study period.

Oversight of the study

Project Management Group

The PUrE PMG was responsible for overseeing the management of the study. This group consisted of the representatives from the trial management team, grant holders [including the patient and public involvement (PPI) representative] plus the trial statistician. The members of the PMG are listed in the acknowledgements.

Trial Steering Committee

One independent TSC was responsible for monitoring and supervising the progress of the PUrE Study (both RCT1 and RCT2). The TSC defined and agreed their role and responsibilities in the TSC Charter and met nine times between February 2016 and December 2022. The TSC consisted of three independent experts, including a PPI representative, and the CI. Key members of the PMG also attended TSC meetings. The members of the TSC are listed in the acknowledgements.

Data Monitoring Committee

One independent DMC was responsible for monitoring safety and data integrity for both RCT1 and RCT2. The DMC consisted of three independent experts. It was independent of the PMG and TSC. The DMC members defined and

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

agreed their roles and responsibilities, and meeting frequency via the DMC Charter at the outset of the PUrE study. The DMC met six times plus had one e-mail update, between May 2016 and December 2022. The trial statistician provided the data and analyses requested by the DMC prior to each meeting. The DMC members are listed in the acknowledgements.

Patient and public involvement

Our PPI grant co-applicant attended PMG meetings and TSC meetings throughout the study. He provided the other members of the team with his individual perspectives on stone disease, reviewed and commented on the study design, protocol and all study documentation, including patient-facing documents and the regular funder progress reports.

In addition, our PPI TSC member attended TSC meetings and provided review and comment on the study and the TSC papers.

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

Protocol amendments

There were four protocol amendments. These are summarised in *Appendix 3*, *Table 36*. All were reviewed by the sponsor and the study funder before being submitted to REC as appropriate.

Important changes to the methods after the start of the trial

Extension to recruitment and follow-up phase

Recruitment was slower than anticipated in both RCT1 and RCT2, so we requested an extension to the recruitment phase of 18 months, which was granted by the National Institute for Health and Care Research (NIHR) Health Technology Assessment (HTA) programme. Due to the COVID pandemic, recruitment was paused for 4 months, on the direction of the sponsor. The follow-up phase of the study was then extended for 9 months to allow participants who were waiting for their intervention, delayed due to the pandemic, to receive their allocated treatment and complete their 12-week post-intervention follow-up. There were no other important changes to the methods after the start of the trials.

Studies within a trial and satellite studies

Christmas card study within a trial

The PUrE study, alongside seven other host trials, was involved in the Christmas Card Study Within a Trial (SWAT). The aim was to determine the effect on retention of sending Christmas cards to RCT participants and to explore the feasibility of a SWAT across multiple trials simultaneously. Full details of the methods and results have been published.[50,51]

Theoretically informed leaflet study within a trial

The PUrE study was involved in a SWAT to explore the benefit of adding a theoretically informed leaflet to the take-home pack with the 1- and 2-week questionnaires. This SWAT aimed to increase the response rates to the participant-reported outcomes (self-completed questionnaires) for the PUrE study by incorporating techniques based on behaviour change theory into a leaflet, developed from qualitative patient-reported studies and PPI input. *Appendix 4*, describes the methods and results as presented in a poster at the 6th International Clinical Trials Methodology Conference, October 2022.

Cambridge Renal Stone Patient Reported Outcome Measure validation study

We asked a sub-set of participants to complete the Cambridge Renal Stone Patient Reported Outcome Measure (CReSP)[52] a disease-specific health-related QoL measure in development, as part of the validation studies for the CReSP. The validation studies will be reported separately by the CReSP developers and are not included in this report.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Chapter 3 RCT1: flexible ureterorenoscopy versus extracorporeal shockwave lithotripsy: baseline, trial results and clinical effectiveness

Parts of this chapter are reproduced with permission from Wiseman *et al.*[1] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, even for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The text below includes additions and formatting changes to the original text.

RCT1 for stones ≤ 10 mm: flexible ureterorenoscopy versus extracorporeal shockwave lithotripsy

Trial recruitment

RCT1 compares FURS with ESWL. In total 466 participants were recruited to RCT1 from 48 UK recruitment centres. The median number of participants per recruitment centre was 7 [interquartile interval (IQI) 3–11]. We randomised 232 and 234 respectively to the care pathways FURS and ESWL. Participants were recruited between 9 May 2016 and 31 March 2021. The final follow-up was completed in January 2023 (see *Appendix 5*, *Table 37* and *Figure 16*).

Recruitment was on track, but it was paused due to the COVID-19 pandemic between 16 March 2020 and 1 July 2020. Only 20 centres had the capacity to re-open to recruitment after the recruitment pause. A further 13 participants (3% of the total 466 participants) were recruited after the recruitment pause.

Participant flow

Figure 2 shows the Consolidated Standards of Reporting Trials (CONSORT) diagram for RCT1. We screened 1964 potentially eligible patients for inclusion in the trial and 1498 were excluded. Reasons for exclusion are shown in *Figure 2*. A total of 286 patients were ineligible due to stone size or position, 197 were unsuitable (surgical reasons) for FURS or ESWL, 689/1498 (46%) of the patients excluded indicated a preference for treatment as a reason for declining to take part. There were 5 post-randomisation exclusions, 2 in the FURS arm and 3 in the ESWL arm, leaving 230 and 231 participants in each arm respectively.

The majority [382, (83%)] of participants had their intervention and reached their 12-week post-intervention time point prior to the COVID-19 restrictions (*Table 2*). However, 30% of participants completed their follow-up to 12 months post randomisation after COVID restrictions started.

Baseline characteristics

The randomised groups were mostly well balanced at baseline (*Table 3*). The mean age of participants was 53 years. The majority of participants were male. Stone measures from detailed imaging, (including stone Hounsfield unit density, as a marker of 'hardness', and skin-to-stone distance) were assessed at baseline after randomisation and were similar in both treatment arms (see *Appendix 5*, *Table 39*). The level of pain participants experienced related to their stone on the day of completing the baseline questionnaire was relatively low (mean 2 on a scale 1–10) and while 50% of the participants reported having experienced pain in the previous week, they required pain medication on a mean 1.6 days out of 7.

The EQ-5D score was imbalanced between the treatment groups at baseline, the FURS group scored 0.047 lower. This imbalance did not manifest in other health status and QoL measures at baseline, the EuroQol Group's 5 dimensions health status questionnaire – visual analogue scale (EQ-5D-VAS) scale and both the SF-12 scores were similar in both treatment groups.

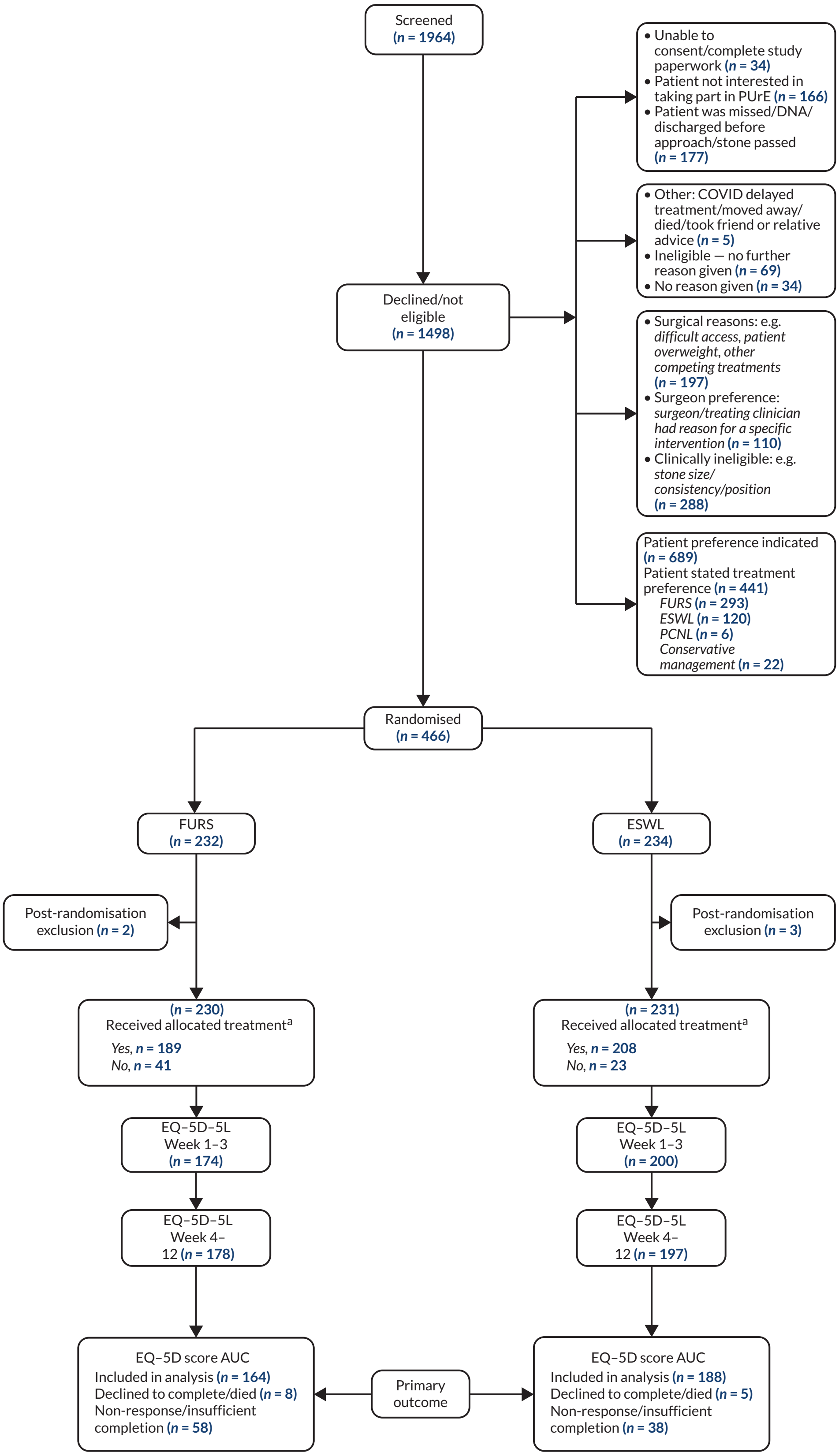


FIGURE 2 Consort flowchart – RCT1. a, See *Table 4*. AUC, area under the curve.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 2 Number of participants at study time points before and after COVID-19 restrictions started

	FURS *N* = 230	ESWL *N* = 231
	n (%)	*n* (%)
Intervention and 12-week post-intervention time point before COVID-19 restrictions	186 (80.9)	196 (84.8)
Intervention before and 12-week post-intervention time point after COVID-19 restrictions	11 (4.8)	10 (4.3)
Randomised before but intervention after COVID-19 restrictions	22 (9.6)	17 (7.4)
Randomised and intervention after COVID-19 restrictions	7 (3.0)	6 (2.6)
Never had intervention	4 (1.7)	2 (0.9)
12-month post-randomisation time point before COVID-19 restrictions	165 (72.7)	160 (69.3)

TABLE 3 Baseline characteristics

Variable	FURS *N* = 230	ESWL *N* = 231
Demographics		
Gender[a]	*n* = 230	*n* = 231
Male	139 (60.4)	148 (64.1)
Female	91 (39.6)	83 (35.9)
Age in years[b]	*n* = 230	*n* = 231
Mean (SD)	54.1 (14.9)	52.8 (14.2)
Median (IQI)	55.8 (45.1–65.3)	54.1 (42.3–63.6)
Minimum, maximum	(20.2, 85.3)	(22.4, 81.2)
BMI[b,c]	*n* = 198	*n* = 183
Mean (SD)	29.1 (5.4)	29.3 (5.5)
Median (IQI)	28.1 (25.5–32.2)	28.7 (25.4–32.4)
Minimum, maximum	(18.9, 51.7)	(18.0, 51.7)
Stone size in mm[b]	*n* = 230	*n* = 231
Mean (SD)	7.1 (1.9)	7.0 (2.0)
Median (IQI)	7.0 (6.0–9.0)	7.0 (5.0–9.0)
Minimum, maximum	(1.0, 10.0)	(1.0, 10.0)
Prior kidney stone treatment[a]	*n* = 230	*n* = 231
Yes	84 (36.5)	74 (32.0)
No	146 (63.5)	156 (67.5)
Missing	-	1 (0.4)
Pain, health status and QoL		
Level of pain today[b,d]	*n* = 229	*n* = 227
Mean (SD)	1.9 (2.5)	1.9 (2.5)
Median (IQI)	0.0 (0.0–4.0)	0.0 (0.0–4.0)
Minimum, maximum	(0, 10)	(0, 10)

continued

TABLE 3 Baseline characteristics (*continued*)

Variable	FURS *N* = 230	ESWL *N* = 231
Stone pain in last 7 days[a]	*n* = 230	*n* = 231
Yes	115 (50.2)	113 (48.9)
No	111 (48.5)	115 (49.8)
Missing	3 (1.3)	3 (1.3)
No. days taken pain relief in last 7 days[b]	*n* = 226	*n* = 228
Mean (SD)	1.6 (2.4)	1.3 (2.3)
Median (IQI)	0.0 (0.0–2.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 7)	(0, 7)
EQ-5D score[b,e]	*n* = 229	*n* = 230
Mean (SD)	0.769 (0.231)	0.816 (0.198)
Median (IQI)	0.819 (0.676–0.985)	0.867 (0.726–0.987)
Minimum, maximum	(-0.209, 0.989)	(-0.174, 0.989)
EQ-5D-VAS[b,f]	*n* = 229	*n* = 230
Mean (SD)	73.2 (20.5)	73.1 (19.5)
Median (IQI)	80.0 (60.0–90.0)	78.0 (60.0–90.0)
Minimum, maximum	(2, 100)	(5, 100)
SF-12 scores[b,g]		
PCS	*n* = 229	*n* = 231
Mean (SD)	45.9 (11.8)	46.9 (10.3)
Median (IQI)	48.2 (37.1–54.8)	49.9 (39.8–54.8)
Minimum, maximum	(16.9, 68.6)	(12.1, 62.4)
MCS	*n* = 227	*n* = 230
Mean (SD)	48.2 (10.8)	49.1 (10.6)
Median (IQI)	49.4 (41.1–56.5)	51.5 (42.4–57.2)
Minimum, maximum	(10.1, 71.4)	(17.9, 65.9)

MCS, mental component score; *n*, number of people who have recorded a response; PCS, physical component score.
a Categorical/ordinal *x* (%); *x* number responding for each category.
b Continuous (or pseudo continuous) variable summaries *n*: mean (SD), median (IQI), (minimum, maximum): (IQI): interquartile interval.
c Collected at time of intervention – participants who did not have an intervention (due to known or assumed passed stone, medical advice for conservative management or by patient decision) had clinical measures retrieved from patient records during study period where possible.
d Level of pain today: 0 to 10 (higher score indicates more pain).
e NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
f EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).
g SF-12 scores range from 0 to 100 (higher score is better). The PCS and MCS are standardised to have a mean of 50 and SD 10.

Care pathway and treatment received

Table 4 describes the allocated pathway and actual treatment received. The majority received their allocated pathway, with 189 (82.2%) on the FURS pathway and 208 (90%) on the ESWL pathway receiving their allocated treatment. However, 21 (9.1%) participants allocated to the FURS arm received ESWL and 10 (4.3%) in ESWL received FURS. One participant in each arm required a PCNL approach. Nineteen (8.3%) and 12 (5.2%) participants did not receive any

Copyright © 2025 Wiseman *et al*. This work was produced by Wiseman *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 4 Allocated pathway and treatment received

	FURS (*N* = 230)	ESWL (*N* = 231)
Treatment received n (%)		
Received FURS	189 (82.2)[a]	10 (4.3)
Received ESWL	21 (9.1)	208 (90.0)[a]
Received PCNL	1 (0.4)	1 (0.4)
Received no initial active treatment	19 (8.3)	12 (5.2)
conservative management	12 (5.2)	4 (1.7)
stone not visible at intervention time	2 (0.9)	5 (2.2)
died before treatment	1 (0.4)	
did not attend	1 (0.4)	1 (0.4)
declined/cancelled active treatment	3 (1.3)	2 (0.9)
Waiting time in weeks:		
mean (sd): *n*	22.8 (23.0): 226	11.1 (13.0): 229
median (IQI)	16.36 (8.3–27.7)	6.6 (4.1–27.9)
Length of stay in days:		
mean (sd): *n*	0.3 (0.6): 211	0.1 (0.5): 222

a Indicates randomised allocated treatment.

active treatment in the FURS and ESWL arms, respectively. Many of these (particularly in the FURS pathway) instead received conservative management as agreed between clinician and participant.

The waiting time to initial treatment (see *Table 4*) was twice as long in the FURS treatment arm compared to ESWL. Length of hospital stay for first intervention treatment was short.

Primary outcome

Figure 3 shows the plots of mean EQ-5D score and the change in EQ-5D score from baseline over time. A similar pattern was seen for the EQ-5D-VAS scores (see *Appendix 6*, *Figure 18*).

The PO and treatment comparison effect sizes are presented in *Table 5*. There were 164/230 (71.3%) on the FURS pathway with sufficient measures within the required time epochs and 188/231 (81.4%) in the ESWL treatment arm. All models show that participants on the FURS pathway may have slightly better outcome up to 12 weeks post intervention compared to ESWL. However, this is not significant in any of the models except for Sensitivity 1, which adjusts for waiting time. For context, the target difference the trial was designed to detect was a 0.3 SD difference, around 0.07 on the AUC score. The per-protocol (treatment as received) analyses are presented in *Appendix 7*, *Table 41*, and results are similar to the ITT analyses. The suite of multiple imputation approaches is described in *Appendix 2*, *Description of different scenarios for multiple imputation* and is presented in *Appendix 7*, *Figure 20*. The results were similar to the ITT analyses.

Secondary outcomes

Definitions and derivations of the secondary outcomes both for clinical outcomes and the patient-reported outcomes (pain, health status and QoL) are given in *Appendix 2*, *Tables 33–35*. Briefly to reiterate: stone clearance was further defined as complete clearance of the target stone from the urinary tract (defined as no further action or observation required for that stone); acceptable clearance, where observation was required but no intervention planned; and unacceptable clearance, where further intervention would be required.

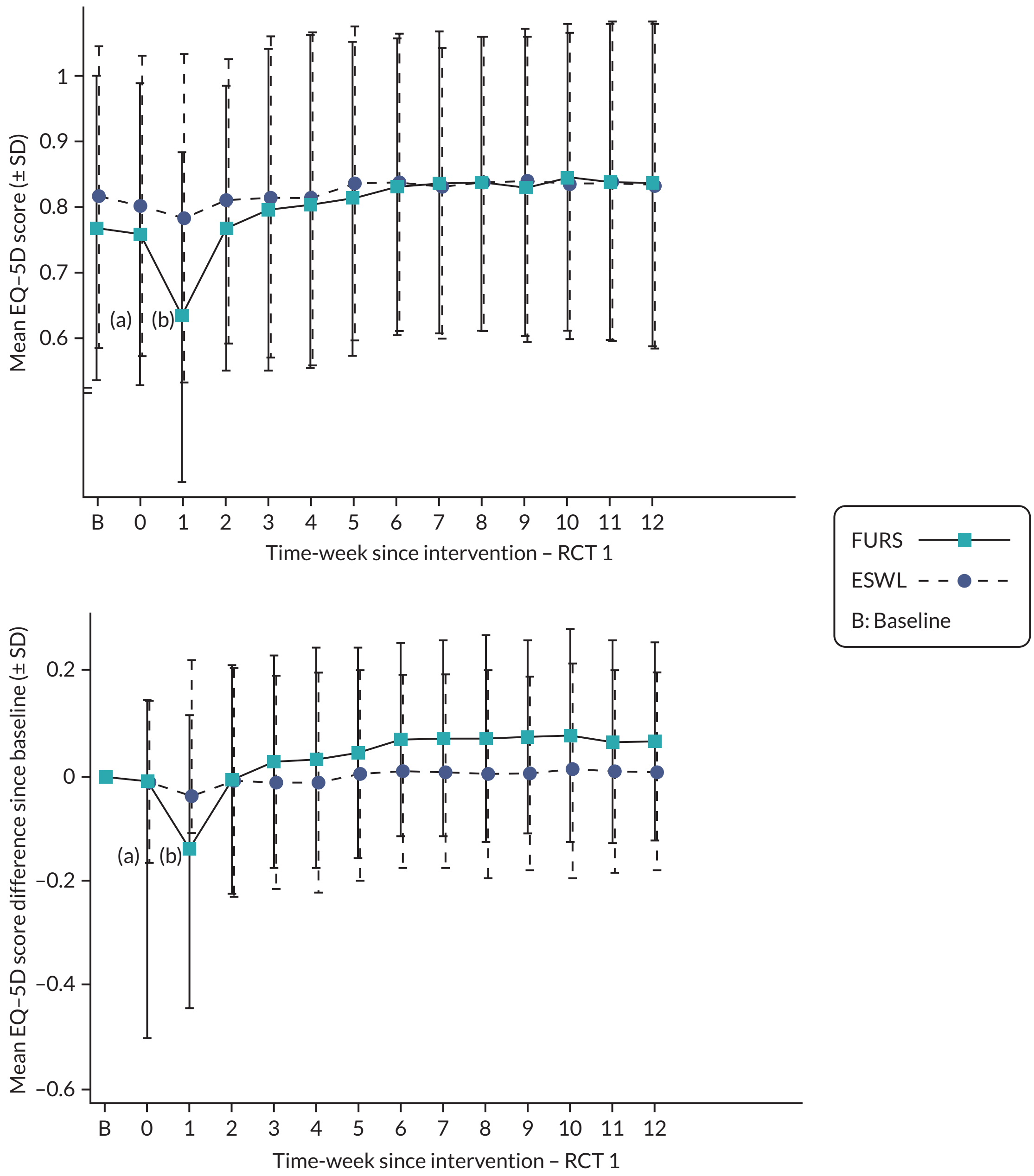


FIGURE 3 Health status plots over time – RCT1. Time points B and 0 are baseline and pre intervention. Areas (a) and (b) do not necessarily represent 1 week, from week 1 to 12 all are 1-week intervals.

Stone clearance and additional treatments by 12 weeks post intervention

At 12 weeks post intervention, more participants had complete stone clearance (as defined in *Appendix 2*, *Table 33*) in the FURS care pathway than for ESWL (71.7% vs. 36.4%) and the overall effect shows FURS to be superior to the ESWL odds ratio (OR): 4.26 (2.74–6.61). The number of participants who required additional treatment (defined in *Appendix 2*, *Table 34*) was lower in FURS group at 12 weeks compared to ESWL [RR: 0.34 (0.18–0.61) $p < 0.001$].

Stone clearance and additional treatments by 12 months post randomisation

The picture for the 12-month post-randomisation time point was complicated by some participants having long waiting times to initial treatment. The majority of participants in both groups reached the 12-month post-randomisation time point after their 12-week post-intervention time point, as intended in the original design. However, some did not have their intervention (or nominated treatment date in the case of conservative management) until after they had been in the trial for 12 months (12 months post randomisation). Others reached the 12-month post-randomisation time point somewhere between intervention and the 12-week post-intervention period. While this happened in both arms, the biggest impact was for the FURS care pathway (*Table 6*). Despite this, we also found that the FURS treatment arm was

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 5 Health status (EQ-5D score) – weekly average area under the curve to 12 weeks post intervention

ITT	FURS *N* = 230 Mean (SD) Valid *n* = 164	ESWL *N* = 231 Mean (SD) Valid *n* = 188	Effect estimate	95% CI	*p*-value
Primary: Main outcome					
EQ-5D score weekly average AUC[a]	0.807 (0.205)	0.826 (0.207)	0.024	(−0.004 to 0.053)	0.097
Adjusting covariate: Baseline EQ-5D score	0.770 (0.217)	0.827 (0.192)			
Primary: Sensitivity 1					
EQ-5D score weekly average AUC[a]	0.807 (0.204)	0.826 (0.207)	0.04	(0.01 to 0.07)	0.011
Adjusting covariate: Baseline EQ-5D score	0.770 (0.217)	0.827 (0.192)			
Adjusting covariate: Waiting times to treatment in weeks	18.890 (15.621)	9.774 (9.712)			
Primary: Sensitivity 2					
EQ-5D score weekly average AUC[a]	0.807 (0.204)	0.826 (0.207)	0.02	(−0.02 to 0.06)	0.307
Adjusting covariate: Pre-intervention EQ-5D score	0.763 (0.219)	0.816 (0.212)			
Secondary: AUC from baseline to 12 weeks					
EQ-5D score weekly average AUC[b]	0.783 (0.202)	0.825 (0.191)	0.004	(−0.016 to 0.024)	0.691
Adjusting covariate: Baseline EQ-5D score	0.770 (0.217)	0.827 (0.192)			

a Time period (number of weeks) is 12, that is from intervention to 12 weeks post intervention.
b The time period (number of weeks) varies for each participant, because the time period is from randomisation/baseline to 12 weeks post intervention and the number of weeks from randomisation/baseline to intervention varies.

Note
Weekly average AUC is the area under the curve of the EQ-5D scores over time (in number of weeks). Effect estimate: positive favours FURS.

superior for both stone clearance and the additional treatment measure (also defined in *Appendix 2*, *Tables 33* and *34*) which (in the 12-month analysis) also includes those with more treatment still planned.

Complications

These were defined as episodes related to an intervention. The number of treatment-related complications was 48 in 44 participants in the FURS arm and 30 in 26 participants in the ESWL arm. There were 7/231 (3%) and 5/230 (2%) in each arm respectively that were graded Clavien 3 or above (*Table 7*).

Patient-reported health status, quality of life and pain

The participant-reported weekly pain and health statuses (EQ-5D score) are summarised in detail in *Appendix 7*, *Tables 43–54*. *Appendix 7* also summarises the pre-intervention (see *Table 42*); and 12-month post-randomisation (see *Table 55*) pain, health status (EQ-5D score) and QoL (SF-12) measures.

All measures of pain (*Table 8*) show post-intervention patterns consistent with EQ-5D-5L plots (see *Figure 3*) where health status in the first couple of weeks post intervention was lower for FURS compared to ESWL then improved to a steady higher level later. Pain was higher in these first weeks for those allocated to FURS but then decreases.

The secondary outcomes of health status, measured by EQ-5D score, rather than AUC, EQ-5D-VAS are reported in *Table 8* modelled over all 12 weeks. EQ-5D scores and VAS have similar health status patterns, with initial detriment of FURS followed by improvement after 2 weeks. *Table 8* also reports the QoL – SF-12 scores (physical and mental domains) but modelled only on the available time points pre intervention and 12 weeks post intervention.

We explored the potential impact of two subgroup factors, reported at baseline determined a-priori: BMI and stone density as measured by the Hounsfield unit on the PO (AUC of the EQ-5D score) and two secondary outcomes: AUC

TABLE 6 Stone clearance and additional treatments – summaries and model estimates

Variable	FURS N = 230 n (%)	ESWL N = 231 n (%)	Effect estimate	95% CI	p-value
Twelve weeks post intervention					
Stone clearance[a]					
Complete	165 (71.7)	84 (36.4)	4.26[b]	(2.74 to 6.61)	< 0.001
Acceptable	9 (3.9)	27 (11.7)			
Unacceptable	51 (21.7)	117 (50.6)			
Missing	5 (2.2)	3 (1.3)			
Complete/acceptable[a]					
Conservative management alone	*4 (2.3)*	*4 (3.6)*			
After first intervention	*164 (94.3)*	*67 (60.4)*			
After second intervention	*6 (3.4)*	*40 (36.0)*			
Additional treatment needed[a]					
Yes	21 (9.1)	63 (27.3)	0.34[c]	(0.18 to 0.61)	< 0.001
No	205 (89.1)	165 (71.4)			
Missing	4 (1.7)	3 (1.3)			
Twelve months post randomisation					
Twelve-months post-randomisation timepoint was:					
Prior to intervention	20	6			
Between intervention and 12 weeks post intervention	9	5			
After 12 weeks post intervention	198	217			
Stone clearance[a]					
Complete	164 (71.3)	88 (38.1)	2.77[b]	(1.60 to 4.78)	< 0.001
Acceptable	4 (1.7)	15 (6.5)			
Unacceptable[d]	59 (25.7)	125 (54.1)			
Missing	3 (1.3)	3 (1.3)			
Treatment/additional treatment done/planned[a,e]					
Yes	46 (20.0)	102 (44.2)	0.45[c]	(0.28 to 0.73)	0.001
No	181 (78.7)	126 (54.5)			
Missing	3 (1.3)	3 (1.3)			

n, number of people with data.
a Categorical/ordinal *x* (%); *x* number in each category.
b Odds ratio from an ordinal regression with robust error variance accounting for centre clusters.
c Risk ratio based on a log-binomial model with robust error variance accounting for centre clusters.
d Assume stone clearance is also unacceptable (unless recorded as passed) if by 12 months participants have not had an intervention or set as conservative management.
e Done/planned: treatments already completed or are still being planned.

Note
Stone clearance at 12 weeks and 12 months: Estimate effect > 1 favours FURS.
Additional Interventions at 12 weeks and 12 months: Estimate effect < 1 favours FURS.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

from baseline and stone clearance to 12 weeks post intervention (definitions are in *Appendix 2*, *Subgroup analyses and technical factors definitions*).

The figures for the AUC outcomes (see *Appendix 7*, *Figure 21*: *Subgroups – main PO*, *Figures 22–24*: *Subgroups – sensitivity 1*, *sensitivity 2*, and *AUC from baseline to 12 weeks*) suggest that these subgroups were not statistically different with respect to treatment arm and their 99% CIs all include the original overall effect. This was also the case for the clinical outcome stone clearance at 12 weeks (see *Figure 25*, *Subgroups – stone clearance at 12 weeks*).

TABLE 7 Complications

	FURS		ESWL	
Number of people in whom a complication was recorded	***44 (19%)***		***26 (11%)***	
Number of complications:	***48***		***30***	
Clavien grade:	*n*	*(%)*	*n*	*(%)*
1	31	(64.6)	19	(63.3)
2	10	(20.8)	6	(20.0)
3a	3	(6.3)	4	(13.3)
3b	4	(8.3)	.	.
4a	.	.	1	(3.3)

TABLE 8 Pain, health status and QoL measures – treatment comparison summaries and model estimates

Variable/week[a]	FURS N = 230	ESWL N = 231	Effect size (95% CI)[b]	*p*-value
Level of pain today[c]	Mean (SD): *n*	Mean (SD): *n*	*n* = 410	
Baseline	2.0 (2.6): 201	1.8 (2.4): 209		
Pre intervention	2.5 (2.5): 172	2.0 (2.5): 182	Ref[d]	
Week 1	3.8 (2.7): 151	2.1 (2.5): 179	1.55 (1.01 to 2.08)	< 0.001
Week 2	1.9 (2.2): 152	1.7 (2.3): 185	0.09 (−0.38 to 0.56)	0.705
Week 3	1.5 (2.2): 153	1.5 (2.3): 169	−0.26 (−0.77 to 0.24)	0.311
Week 4	1.2 (2.0): 157	1.6 (2.3): 166	−0.47 (−0.95 to 0.01)	0.055
Week 5	1.2 (2.0): 155	1.3 (2.2): 171	−0.25 (−0.66 to 0.16)	0.228
Week 6	1.1 (1.9): 160	1.2 (2.1): 170	−0.21 (−0.61 to 0.18)	0.294
Week 7	1.1 (2.1): 152	1.2 (2.0): 166	−0.20 (−0.55 to 0.15)	0.254
Week 8	1.0 (1.9): 152	1.2 (2.2): 161	−0.40 (−0.85 to 0.04)	0.073
Week 9	0.9 (1.9): 150	1.1 (1.8): 165	−0.32 (−0.65 to 0.02)	0.065
Week 10	1.0 (1.9): 143	1.2 (2.2): 167	−0.35 (−0.77 to 0.06)	0.093
Week 11	0.8 (1.8): 145	1.1 (2.1): 166	−0.28 (−0.64 to 0.07)	0.118
Week 12	0.9 (1.9): 154	1.0 (1.9): 166	−0.30 (−0.65 to 0.05)	0.091
Stone pain in last 7 days,[e] Yes	*x* (%): *n*	*x* (%): *n*	*n* = 408	
Baseline	102 (51.5): 198	100 (47.6): 210		
Pre intervention	99 (59.3): 167	86 (47.0): 183	Ref[d]	

TABLE 8 Pain, health status and QoL measures – treatment comparison summaries and model estimates (*continued*)

Variable/week[a]	FURS *N* = 230	ESWL *N* = 231	Effect size (95% CI)[b]	*p*-value
Week 1	118 (83.1): 142	114 (64.8): 176	1.24 (1.05 to 1.47)[f]	0.011
Week 2	87 (59.6): 146	89 (49.4): 180	1.12 (0.87 to 1.44)[f]	0.380
Week 3	52 (35.1): 148	71 (44.1): 161	0.72 (0.52 to 1.01)[f]	0.061
Week 4	45 (30.0): 150	71 (43.3): 164	0.67 (0.50 to 0.89)[f]	0.006
Week 5	39 (26.2): 149	59 (35.3): 167	0.72 (0.53 to 0.98)[f]	0.035
Week 6	41 (26.3): 156	60 (36.4): 165	0.67 (0.47 to 0.96)[f]	0.027
Week 7	30 (20.8): 144	53 (32.7): 162	0.60 (0.42 to 0.87)[f]	0.006
Week 8	27 (18.6): 145	53 (33.3): 159	0.51 (0.35 to 0.74)[f]	< 0.001
Week 9	29 (20.3): 143	47 (29.4): 160	0.64 (0.45 to 0.91)[f]	0.012
Week 10	28 (20.7): 135	44 (27.0): 163	0.74 (0.53 to 1.03)[f]	0.075
Week 11	28 (21.1): 133	42 (25.8): 163	0.77 (0.51 to 1.16)[f]	0.213
Week 12	30 (20.4): 147	51 (31.1): 164	0.62 (0.43 to 0.88)[f]	0.007
No. days needed pain relief in last 7 days	Mean (SD): *n*	Mean (SD): *n*	*n* = 408	
Baseline	1.7 (2.5): 198	1.3 (2.2): 210		
Pre intervention	2.5 (2.7): 166	1.7 (2.5): 179	Ref[d]	
Week 1	5.1 (2.3): 148	2.3 (2.5): 180	2.52 (2.00 to 3.04)	< 0.001
Week 2	2.8 (2.6): 147	1.5 (2.2): 185	0.94 (0.42 to 1.45)	< 0.001
Week 3	1.9 (2.6): 151	1.4 (2.2): 166	0.22 (−0.22 to 0.65)	0.333
Week 4	1.7 (2.6): 153	1.3 (2.1): 166	0.19 (−0.27 to 0.64)	0.424
Week 5	1.6 (2.6): 152	1.1 (2.0): 171	0.25 (−0.25 to 0.75)	0.329
Week 6	1.4 (2.3): 153	1.2 (2.0): 170	−0.04 (−0.50 to 0.41)	0.853
Week 7	1.2 (2.3): 149	1.1 (1.9): 165	−0.05 (−0.40 to 0.29)	0.765
Week 8	1.1 (2.1): 148	1.1 (1.9): 161	−0.33 (−0.79 to 0.14)	0.173
Week 9	1.3 (2.3): 141	1.1 (2.0): 165	−0.13 (−0.63 to 0.36)	0.597
Week 10	1.2 (2.2): 136	1.0 (2.0): 167	−0.05 (−0.48 to 0.38)	0.828
Week 11	1.2 (2.2): 141	1.0 (1.9): 166	0.02 (−0.39 to 0.44)	0.909
Week 12	1.2 (2.1): 146	1.0 (1.9): 161	−0.17 (−0.49 to 0.15)	0.299
Worst level of pain since intervention	Mean (SD): *n*	Mean (SD): *n*	*n* = 394	
Week 1	7.2 (2.7): 151	3.9 (3.1): 184	Ref[d]	
Week 2	6.9 (2.9): 152	3.7 (3.2): 189	2.84 (2.06 to 3.63)	< 0.001
Week 3	6.1 (3.5): 153	3.6 (3.2): 169	2.07 (1.37 to 2.77)	< 0.001
Week 4	6.1 (3.5): 158	4.1 (3.3): 166	1.90 (1.24 to 2.57)	< 0.001
Week 5	5.9 (3.5): 155	3.9 (3.3): 174	1.97 (1.28 to 2.66)	< 0.001
Week 6	6.0 (3.5): 160	4.1 (3.5): 171	1.71 (1.04 to 2.39)	< 0.001
Week 7	6.0 (3.4): 154	4.1 (3.5): 169	1.84 (1.23 to 2.44)	< 0.001

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 8 Pain, health status and QoL measures – treatment comparison summaries and model estimates (*continued*)

Variable/week[a]	FURS *N* = 230	ESWL *N* = 231	Effect size (95% CI)[b]	*p*-value
Week 8	5.8 (3.6): 152	4.3 (3.5): 162	1.31 (0.59 to 2.02)	< 0.001
Week 9	5.9 (3.5): 150	4.1 (3.5): 167	1.78 (0.95 to 2.61)	< 0.001
Week 10	5.9 (3.4): 143	4.2 (3.6): 169	1.65 (0.88 to 2.43)	< 0.001
Week 11	5.6 (3.7): 145	4.3 (3.5): 167	1.28 (0.51 to 2.06)	0.001
EQ-5D score[g]	Mean (SD): *n*	Mean (SD): *n*	*n* = 413	
Baseline	0.771 (0.224): 201	0.825 (0.192): 212		
Pre intervention	0.759 (0.230): 173	0.802 (0.224): 183	Ref[d]	
Week 1	0.634 (0.251): 148	0.784 (0.222): 179	−0.11 (−0.15 to −0.06)	< 0.001
Week 2	0.770 (0.216): 151	0.811 (0.219): 184	−0.01 (−0.05 to 0.03)	0.661
Week 3	0.796 (0.246): 148	0.814 (0.219): 158	0.03 (−0.02 to 0.07)	0.215
Week 4	0.809 (0.253): 147	0.814 (0.234): 164	0.03 (−0.01 to 0.07)	0.111
Week 5	0.815 (0.240): 148	0.839 (0.208): 157	0.02 (−0.01 to 0.06)	0.207
Week 6	0.831 (0.226): 151	0.839 (0.207): 159	0.04 (0.00 to 0.08)	0.027
Week 7	0.836 (0.230): 148	0.833 (0.229): 159	0.05 (0.02 to 0.08)	0.001
Week 8	0.837 (0.225): 150	0.837 (0.225): 155	0.06 (0.02 to 0.09)	0.005
Week 9	0.837 (0.233): 144	0.829 (0.233): 159	0.06 (0.02 to 0.10)	0.002
Week 10	0.844 (0.234): 141	0.832 (0.231): 165	0.05 (0.01 to 0.09)	0.012
Week 11	0.838 (0.241): 143	0.840 (0.233): 162	0.05 (0.00 to 0.09)	0.030
Week 12	0.836 (0.247): 153	0.832 (0.239): 166	0.05 (0.02 to 0.08)	0.003
EQ-5D-VAS[h]	Mean (SD): *n*	Mean (SD): *n*	*n* = 412	
Baseline	74.4 (19.2): 199	73.9 (18.8): 213		
Pre intervention	72.8 (17.9): 173	76.7 (17.1): 185	Ref[d]	
Week 1	65.8 (20.2): 145	72.0 (20.5): 180	−5.82 (−9.59 to−2.05)	0.002
Week 2	73.2 (19.8): 145	74.8 (20.3): 182	−1.92 (−5.01 to 1.16)	0.222
Week 3	75.2 (20.0): 140	76.2 (20.1): 158	1.25 (−1.65 to 4.15)	0.398
Week 4	77.8 (20.3): 140	74.2 (22.4): 163	4.55 (0.71 to 8.38)	0.020
Week 5	78.1 (20.9): 140	76.2 (20.9): 156	3.86 (0.66 to 7.05)	0.018
Week 6	78.9 (20.1): 143	75.5 (21.2): 158	4.71 (1.46 to 7.95)	0.004
Week 7	78.1 (21.9): 140	78.0 (19.1): 156	1.55 (−1.73 to 4.83)	0.355
Week 8	78.5 (21.8): 141	77.3 (20.9): 151	3.02 (0.01 to 6.03)	0.049
Week 9	78.9 (21.3): 137	76.8 (21.1): 158	4.66 (1.20 to 8.13)	0.008
Week 10	79.6 (22.0): 131	77.4 (20.0): 163	2.68 (−0.91 to 6.27)	0.144
Week 11	79.6 (22.7): 135	78.9 (19.7): 160	2.85 (−0.53 to 6.23)	0.098
Week 12	80.4 (20.7): 143	78.6 (20.2): 165	3.19 (0.50 to 5.89)	0.020

TABLE 8 Pain, health status and QoL measures – treatment comparison summaries and model estimates (*continued*)

Variable/week[a]	FURS N = 230	ESWL N = 231	Effect size (95% CI)[b]	p-value
SF-12 scores[i]				
Physical component	Mean (SD): *n*	Mean (SD): *n*	*n = 392*	
Baseline	45.8 (11.9): 190	47.1 (10.3): 206		
Pre intervention	45.7 (11.2): 159	47.9 (9.6): 176	Ref[d]	
Week 12	47.8 (11.6): 144	48.4 (10.7): 165	0.68 (−1.18 to 2.55)	0.473
Mental component	Mean (SD): *n*	Mean (SD): *n*	*n = 392*	
Baseline	47.7 (11.1): 190	49.6 (10.4): 206		
Pre intervention	47.8 (10.9): 159	49.0 (10.4): 176	Ref[d]	
Week 12	48.6 (11.0): 144	48.1 (11.5): 165	1.55 (−0.32 to 3.42)	0.104

n, number of people with a recorded response; *x*, number of 'Yes' responses to 'Stone pain in last 7 days'.
a Participants included in the models on average returned 9 to 10 weekly questionnaires (range 1–12).
b Models are mixed linear regression with robust error variance accounting for centre clusters unless otherwise specified: effect estimates > 0 indicates FURS has a larger effect size for that outcome than ESWL.
c Levels of pain: (higher score indicates more pain).
d Ref = reference time point.
e Days needing pain relief: (higher score indicates more pain).
f Risk ratio based on a modified random effects multilevel Poisson regression with a log-link function and robust error variance to account for centre clusters:[53] > 1 indicates FURS participant having higher pain levels than ESWL participants.
g NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
h EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).
i SF-12 scores range from 0 to 100 (higher score is better). Subgroup analyses.

Technical factors

We report the results for the main outcomes and covariates in the PO models (main, sensitivity 1 and sensitivity 2), the secondary AUC model (AUC from baseline to 12 weeks post intervention) and also for stone clearance at 12 weeks for the technical factors in *Appendix 8* (see *Tables 71–75*). These were not a comparison between treatment arms (since they were within a specified treatment) but summaries of each of the outcomes and covariates for each technical factor across the two RCTs. They were restricted to technical factors for the first intervention.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Chapter 4 RCT2: flexible ureterorenoscopy versus percutaneous nephrolithotomy – baseline, trial results and clinical effectiveness

RCT2 for stones > 10 mm: flexible ureterorenoscopy versus percutaneous nephrolithotomy

Trial recruitment

RCT2 compares FURS with PCNL as kidney stone treatments for larger stones (> 10 mm, ≤ 25 mm). We recruited 159 participants from 39 centres. The median number of participants per centre was 2 (IQI 1–3). We randomised 73 and 86 participants into each of the care pathways FURS and PCNL respectively. Participants were recruited between 9 May 2016 and 31 March 2021, with final follow-up completed in January 2023 (see *Appendix 5*, *Table 38* and *Figure 17*).

Recruitment was paused due to the COVID-19 pandemic between 16 March 2020 and 1 July 2020. Only 20 sites had the capacity to reopen to recruitment after the recruitment pause. A further 9 participants were recruited after the study reopened to recruitment (6% of the total 159 recruited).

Participant flow

The CONSORT diagram for RCT2 (*Figure 4*) shows that 545 potentially eligible patients were screened and 386 were excluded. Reasons for exclusion are shown in *Figure 4*. Of those excluded 60 were deemed unsuitable (surgical reasons) for FURS or PCNL, 50 were ineligible due to stone size or position, and 193/386 (50%) of patients that were excluded, indicated a preference for treatment as a reason for declining to take part. There were no post-randomisation exclusions in RCT2.

The majority [128, (80%)] of participants received their intervention and reached their 12-week post-intervention time point prior to the COVID-19 restrictions (*Table 9*). However, 27% (43/159) participants completed their follow-up to 12 months post randomisation after COVID-restrictions started.

Baseline characteristics

The randomised groups were mostly well balanced at baseline (*Table 10*). The mean age of participants was 58 years, and most were male. This trial was restricted to those with larger kidney stones (> 10 mm). Stone measures from detailed imaging (including stone Hounsfield unit density, as a marker of 'hardness', and skin-to-stone distance) were assessed at baseline after randomisation and were similar in both treatment arms, see *Appendix 5*, *Table 40*. In the FURS group, 29% of the participants had undergone previous stone treatments compared to 43% in the PCNL group. The level of stone-related pain that participants experienced on the day (scale 0–10) completed for the baseline questionnaire was low with median 0 days (IQI 0–3). Pain in the previous week was experienced by 44% of the participants reported. Participants reported taking pain medication on a median of 0 days (IQI 0–7) in the previous 7 days.

The health status and QoL measures taken at baseline these were comparable in both treatment groups for all the measures.

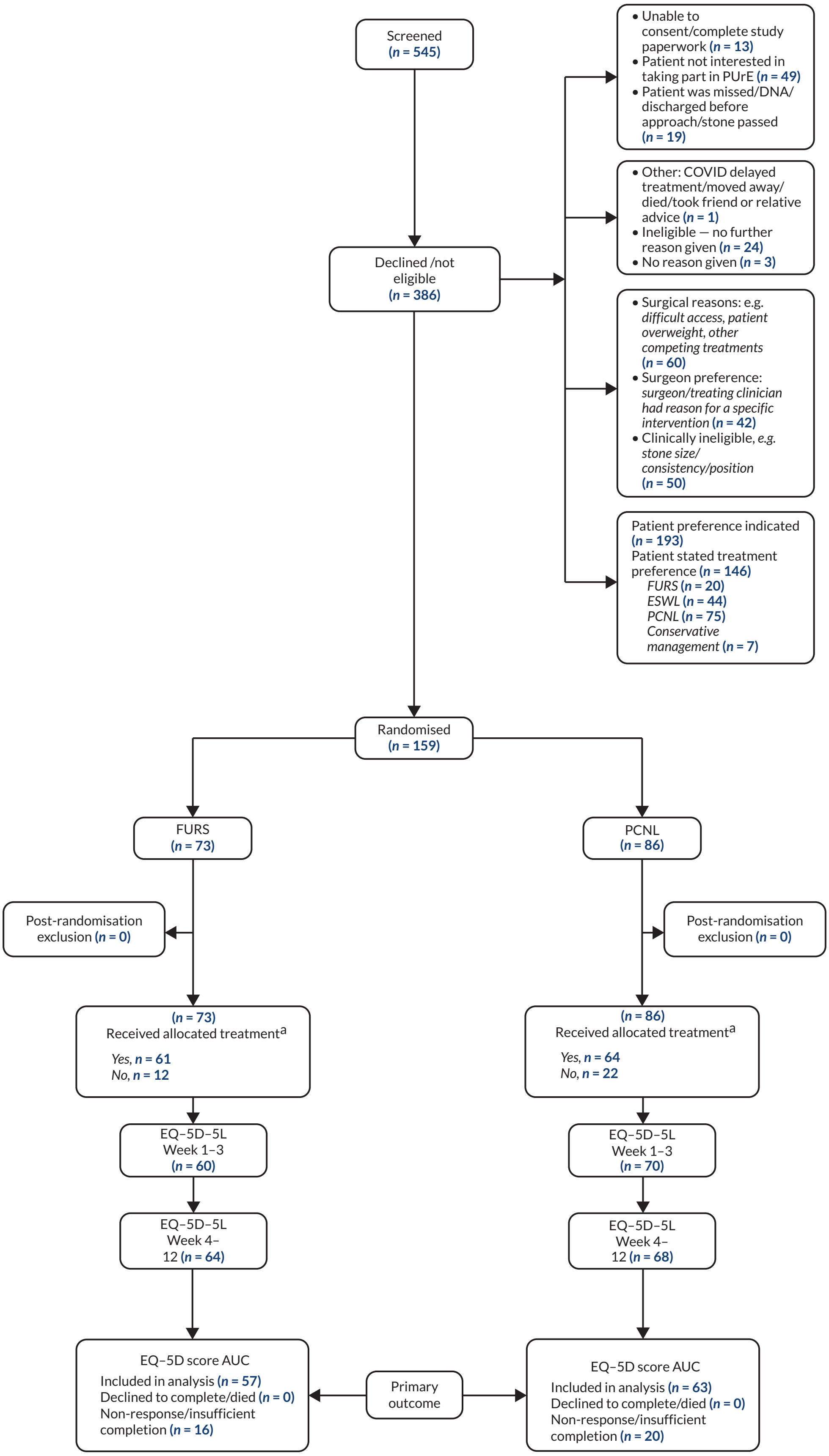


FIGURE 4 Consort flowchart – RCT2. a, See *Table 11*. AUC, area under the curve.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 9 Number of participants at study time points before and after COVID-19 restrictions started

	FURS *N* = 73	PCNL *N* = 86
	n (%)	*n* (%)
Intervention and 12-week post-intervention time point before COVID-19 restrictions	60 (82.2)	68 (79.1)
Intervention before and 12-week post-intervention time-point after COVID-19 restrictions	1 (1.4)	3 (3.5)
Randomised before but intervention after COVID-19 restrictions	8 (11.0)	10 (11.6)
Randomised and intervention after COVID-19 restrictions	4 (5.5)	4 (4.7)
Never had intervention		1 (1.2)
12-month post-randomisation time point before COVID-19 restrictions	53 (72.6)	63 (73.3)

TABLE 10 Baseline characteristics

Demographics	FURS *N* = 73	PCNL *N* = 86
Gender[a]	*n* = 73	*n* = 86
Male *n* (%)	47 (64.4)	53 (61.6)
Female	26 (35.6)	33 (38.4)
Age in years[b]	*n* = 73	*n* = 86
Mean (SD)	57.1 (13.3)	60.3 (12.0)
Median (IQI)	60.8 (49.6–67.0)	61.0 (52.0–68.8)
Minimum, maximum	(24.2, 76.3)	(19.4, 83.3)
BMI[b,c]	*n* = 70	*n* = 79
Mean (SD)	30.4 (6.9)	31.4 (6.3)
Median (IQI)	29.1 (25.7–33.7)	30.8 (26.5–35.5)
Minimum, maximum	(16.2, 52.0)	(21.0, 53.1)
Stone size in mm[a]	*n* = 73	*n* = 86
Mean (SD)	13.9 (2.9)	14.5 (3.2)
Median (IQI)	13.0 (12.0–15.0)	13.0 (12.0–17.0)
Minimum, maximum	(11.0–22.0)	(11.0–24.0)
Prior kidney stone treatment[a]	*n* = 73	*n* = 86
Yes	21 (28.8)	37 (43.0)
No	52 (71.2)	49 (57.0)
Pain, health status and QoL		
Level of pain today[b,d]	*n* = 73	*n* = 85
Mean (SD)	1.6 (2.6)	1.8 (2.7)
Median (IQI)	0.0 (0.0–2.0)	0.0 (0.0–3.0)
Minimum, maximum	(0, 10)	(0, 10)
Stone pain in last 7 days[a]	*n* = 73	*n* = 86
Yes	32 (43.8)	38 (44.2)
No	41 (56.2)	48 (55.8)

TABLE 10 Baseline characteristics (*continued*)

Demographics	FURS *N* = 73	PCNL *N* = 86
No days taken pain relief in last 7 days[b]	*n* = 73	*n* = 86
Mean (SD)	1.3 (2.2)	1.5 (2.5)
Median (IQI)	0.0 (0.0–3.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 7)	(0, 7)
EQ-5D score[a,e]	*n* = 73	*n* = 86
Mean (SD)	0.775 (0.275)	0.758 (0.268)
Median (IQI)	0.860 (0.697–0.986)	0.827 (0.673–0.987)
Minimum, maximum	(-0.308, 0.989)	(-0.241, 0.989)
EQ-5D-VAS[a,f]	*n* = 73	*n* = 86
Mean (SD)	72.2 (19.4)	70.6 (21.4)
Median (IQI)	75.0 (60.0–90.0)	75.0 (55.0–90.0)
Minimum, maximum	(10, 100)	(10, 100)
SF-12 scores[a,g]		
PCS	*n* = 72	*n* = 84
Mean (SD)	45.3 (10.6)	43.6 (10.9)
Median (IQI)	46.9 (39.1–54.4)	44.8 (36.1–52.9)
Minimum, maximum	(20.8, 62.0)	(15.4, 63.5)
MCS	*n* = 72	*n* = 85
Mean (SD)	48.7 (11.0)	49.2 (11.4)
Median (IQI)[d]	51.2 (40.9–56.9)	53.7 (44.5–57.4)
Minimum, maximum	(16.0, 65.6)	(16.0, 63.1)

MCS, mental component score; *n*, number of people who have recorded a response; PCS, physical component score.
a Categorical/ordinal *x* (%); *x* number responding for each category.
b Continuous (or pseudo continuous) variable summaries *n*: mean (SD), median (IQI), (minimum, maximum): (IQI): interquartile interval.
c Mostly collected at time of intervention – participants who did not have an intervention (known or assumed passed stone, medical advice for conservative management or by patient decision) had clinical measures retrieved from patient records during study period where possible.
d Level of pain today: 0 to 10 (higher score indicates more pain).
e NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
f EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).
g SF-12 scores range from 0 to 100 (higher score is better). The PCS and MCS are standardised to have a mean of 50 and SD 10.

Care pathway and treatment received

The allocated treatment compared to the received treatment is described in *Table 11*. A total of 61 (83.6%) received their allocated pathway in the FURS arm and 64 (74.4%) received their allocated pathway in the PCNL arm. A total of 8 (11%) participants allocated to FURS had PCNL while 16 (18.6%) allocated to PCNL had FURS. A total of 10 participants did not receive any active treatment: 4 in the FURS arm and 6 in the PCNL arm. Most of these decided on conservative management with their clinician. For those participants, we used the date when conservative management or the care plan was agreed to start their 12-week post-intervention period.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 11 Care pathway allocated and treatment received

	FURS *N* = (73)	PCNL *N* = (86)
Treatment received n (%)		
Received FURS	61 (83.6)[a]	16 (18.6)
Received PCNL	8 (11.0)	64 (74.4)[a]
Received no initial active treatment	4 (5.5)	6 (7.0)
conservative management	3 (4.1)	3 (3.5)
stone not visible at intervention time		2 (2.3)
died before treatment		1 (1.2)
declined/cancelled active treatment	1 (1.4)	
Waiting time in weeks		
mean (sd): *n*	27.1 (24.4): 73	27.2 (21.1): 85
median (IQI)	20.4 (12.3–34.1)	18.8 (13.8–36.3)
Length of stay in days:		
mean (sd): *n*	0.7 (1.3): 69	1.7 (1.8): 79

a Indicates randomised allocated treatment.

The waiting times to initial treatment (see *Table 11*) were very similar in both treatment groups, both being surgical procedures (mean around 27 weeks) but with an equally large SD indicating a slight skewed distribution. Length of hospital stay for first intervention treatment differed slightly between treatment groups.

Primary outcome

Figure 5 shows plots of the mean EQ-5D score and the change in EQ-5D score from baseline over time. A similar pattern was seen for the EQ-5D-VAS scores (see *Appendix 6*, *Figure 19*).

The PO and treatment comparison effect sizes are presented in *Table 12*; in the FURS arm 57/73 (78.1%) had sufficient measures and 63/86 (73.3%) in the PCNL arm had sufficient measures. The model effect estimates should be considered along with the original minimal clinical important difference (MCID) the study was powered to be around 0.07 (based on 0.3 SD difference and baseline data). The difference between groups consistently favours PCNL.

The per-protocol (treatment as received) analyses are presented in *Appendix 7*, *Table 56*. The treatment effects were similar to the ITT analyses. The suite of multiple imputation approaches is described in *Appendix 2*, *Description of different scenarios for multiple imputation* and presented in *Appendix 7*, *Figure 26*. The results were similar to the original ITT analyses.

Secondary outcomes

Appendix 2 gives definitions and derivations of the secondary outcomes both for clinical outcomes (stone clearance and additional treatments) and the patient reported outcomes (pain, health status and QoL). Briefly to reiterate: stone clearance was further defined as complete clearance of the target stone from the urinary tract defined, and no further action or observation required for that stone; acceptable clearance, where observation is required but no intervention planned; and unacceptable clearance, where further intervention will be required.

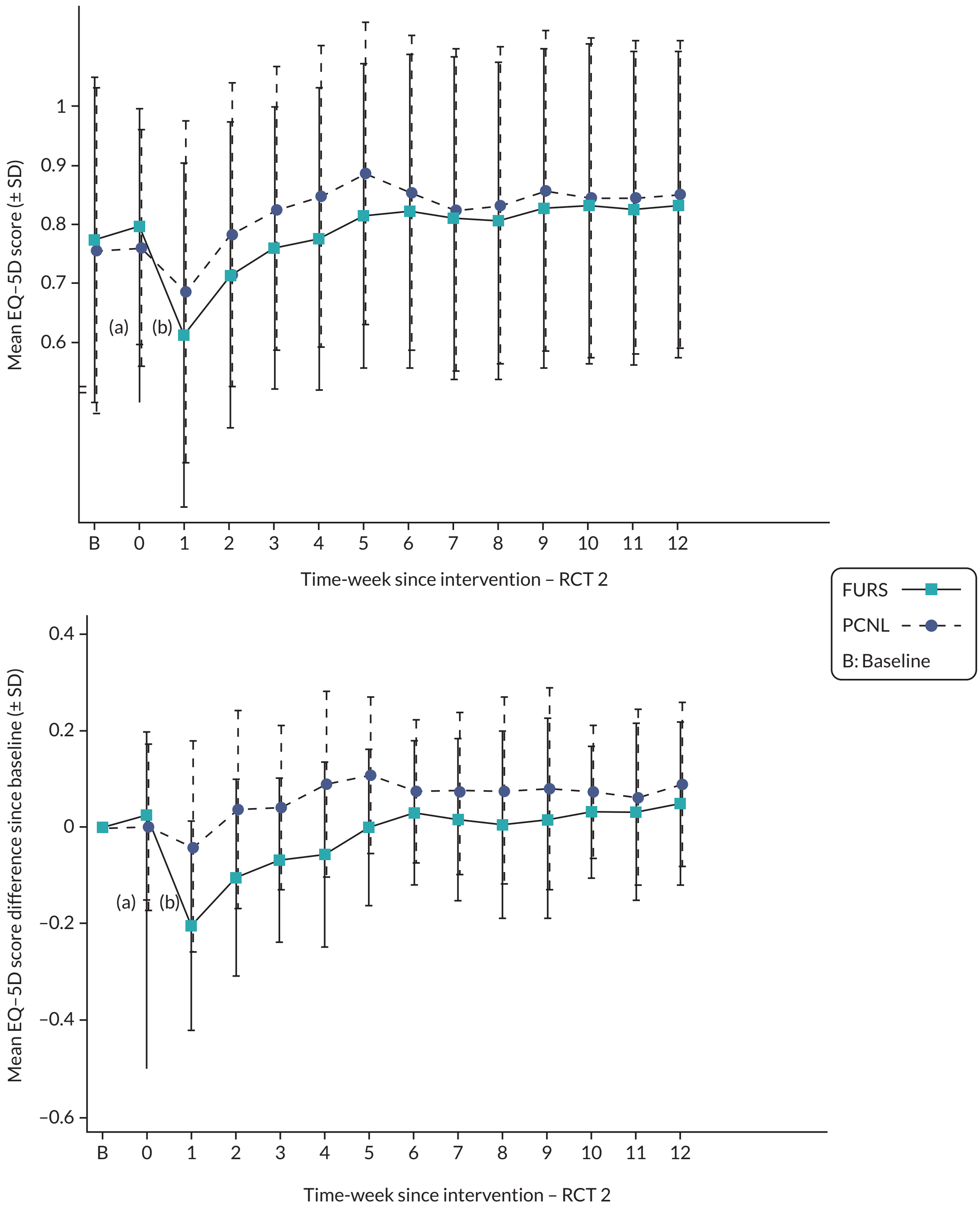


FIGURE 5 Health status plots over time – RCT2. Time points B and 0 are baseline and pre intervention. Areas (a) and (b) do not necessarily represent 1 week from week 1 to 12; all are 1-week intervals (EQ-5D score: observed and differences since baseline).

Stone clearance and additional treatments by 12 weeks post intervention

At 12 weeks post intervention, the FURS care pathway had fewer participants with complete stone clearance than PCNL (47.9% vs. 70.9%) with the overall effect indicating that for these larger stones PCNL is superior to FURS [OR: 0.27 (0.16, 0.43); $p < 0.001$] (*Table 13*). Further, the FURS group had a larger risk of having had additional treatments at this point compared to PCNL [RR: 3.96 (2.07, 7.57); $p < 0.001$].

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 12 Health status (EQ-5D score) – weekly average AUC to 12 weeks post intervention

ITT	FURS *N* = 73 Mean (SD) Valid *n* = 57	PCNL *N* = 86 Mean (SD) Valid *n* = 63	Effect estimate	95% CI	*p*-value
Primary: main outcome					
EQ-5D score weekly average AUC[a]	0.794 (0.198)	0.818 (0.217)	−0.07	(−0.11 to −0.02)	0.006
Adjusting covariate: baseline EQ-5D score	0.822 (0.211)	0.756 (0.259)			
Primary: sensitivity 1					
EQ-5D score weekly average AUC[a]	0.794 (0.198)	0.818 (0.217)	−0.07	(−0.12 to −0.02)	0.006
Adjusting covariate: baseline EQ-5D score	0.822 (0.211)	0.756 (0.259)			
Adjusting covariate: waiting times to treatment in weeks	25.496 (24.379)	23.711 (16.437)			
Primary: sensitivity 2					
EQ-5D score weekly average AUC[a]	0.794 (0.198)	0.818 (0.217)	−0.06	(−0.13 to 0.001)	0.053
Adjusting covariate: pre-intervention EQ-5D score	0.826 (0.141)	0.765 (0.226)			
Secondary: AUC from baseline-12 weeks					
EQ-5D score weekly average AUC[b]	0.794 (0.175)	0.771 (0.216)	−0.03	(−0.05 to 0.002)	0.071
Adjusting covariate: baseline EQ-5D score	0.822 (0.211)	0.756 (0.259)			

a Time period (number of weeks) is 12, that is from intervention to 12 weeks post intervention.
b The time period (number of weeks) varies for each participant, because the time period is from randomisation/baseline to 12 weeks post intervention and the number of weeks from randomisation/baseline to intervention varies.

Note
Weekly average AUC is the area under the curve of the EQ-5D scores over time (in number of weeks). Effect estimate: positive favours FURS.

Stone clearance and additional treatments by 12 months post randomisation

For most (80.5%) participants in this trial, the 12-month post randomisation measures were after the 12-week post-intervention time point as intended in the original design. There were a few participants [18 (12%)] who had not had their intervention (or set nominated treatment date in the case conservative management) until after they had been in the trial for 12 months. Twelve (8%), reached their 12-month post-randomisation time point somewhere between intervention and the 12-week post-intervention period. This happened in both arms (see *Table 13*). The individuals still waiting for treatment were classed as requiring treatment unless evidence of their stone passing was given. Again though, as with the 12-week results, FURS had a lower rate of stone clearance compared to PCNL [OR: 0.24 (0.16, 0.35); $p < 0.001$] and an increased rate of requiring additional treatment and/or had additional treatment [RR: 2.78 (1.69, 4.57); $p < 0.001$]. These were assessed in the context of where in the timeline the 12-month measures fell, taking into account that some were only half-way through their allowed care pathway.

Complications

The number of treatment-related complications was small: 16/73 (22%) participants randomised to FURS and 17/86 (20%) randomised to PCNL experiencing a complication. There were 17 complications (*Table 14*) in each arm with 2/73 (3%) in the FURS arm and 4/86 (5%) in the PCNL arm graded as Clavien–Dindo grade 3 or above.

TABLE 13 Stone clearance and additional treatments – summaries and model estimates

Variable	FURS *N* = 73 *n* (%)	PCNL *N* = 86 *n* (%)	Effect estimate	95% CI	*p*-value
12 weeks post intervention					
Stone clearance[a]					
Complete	35 (47.9)	61 (70.9)	0.27[b]	(0.16 to 0.43)	< 0.001
Acceptable	5 (6.8)	11 (12.8)			
Unacceptable	31 (42.5)	9 (10.5)			
Missing	2 (2.7)	5 (5.8)			
Complete/acceptable[a]					
Conservative management alone	–	3 (4.2)			
After 1st intervention	35 (87.5)	68 (94.4)			
After 2nd intervention	5 (12.5)	1 (1.4)			
Additional intervention[a]					
Yes	17 (23.3)	5 (5.8)	3.96[c]	(2.07 to 7.57)	< 0.001
No	56 (76.7)	80 (93.0)			
Missing	–	1 (1.2)			
12 months post randomisation					
Twelve-months post-randomisation timepoint was:					
Prior to intervention	10	8			
Between intervention and 12 weeks post intervention	3	9			
After 12 weeks post intervention	60	68			
Stone clearance[a]					
Complete	34 (46.6)	59 (68.6)	0.24[b]	(0.16 to 0.35)	< 0.001
Acceptable	1 (1.4)	8 (9.3)			
Unacceptable[d]	38 (52.1)	18 (20.9)			
Missing	–	1 (1.2)			
Treatment/additional treatment done/planned[e]					
Yes	31 (42.5)	13 (15.1)	2.78[c]	(1.69 to 4.57)	< 0.001
No	42 (57.5)	72 (83.7)			
Missing	–	1 (1.2)			

n, number of people with data.
a Categorical/ordinal *x* (%); *x* number in each category.
b Odds ratio from an ordinal regression with robust error variance accounting for centre clusters.
c Risk ratio based on a log-binomial model with robust error variance accounting for centre clusters.
d Assume stone clearance is also unacceptable (unless recorded as passed) if by 12 months participants have not had an intervention or set as conservative management. Stone clearance at 12 weeks and 12 months: Estimate effect > 1 favours FURS.
e Done/planned: treatments already completed or are still being planned. Additional interventions (at 12 weeks and 12 months): Estimate effect < 1 favours FURS.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Patient-reported pain, health status and quality of life

The patient-reported weekly pain and health status (EQ-5D score) are summarised in detail in *Appendix 7*, *Tables 58–69*. *Appendix 7* also summarises the pre-intervention (see *Table 57*) and 12-month post-randomisation (see *Table 70*) pain, health status (EQ-5D score) and QoL (SF-12) measures.

The weekly pain measures (*Table 15*) reflect the health status plots (see *Figure 5*) – the pain measures indicate very slightly higher pain in the first weeks for those allocated to FURS compared to those on PCNL but then decrease. The secondary outcomes of health status, measured by the EQ-5D score and the EQ-5D-VAS scale, and QoL as measured by the SF-12 scores (physical and mental domains) are also reported *Table 15*. Both EQ-5D-5L measures show similar patterns with initial detriment seen in FURS followed by improvement after 2 to 5 weeks, statistically comparable in the PCNL group. The SF-12 measures do not show any difference by 12 weeks.

TABLE 14 Complications

	FURS		PCNL	
Number of participants in whom a complication was recorded	***16 (22%)***		***17 (20%)***	
Number of complications:	***17***		***17***	
Clavien grade:	*n*	*%*	*n*	*%*
1	9	52.9	6	35.3
2	6	35.3	7	41.2
3a	1	5.9	2	11.8
3b	1	5.9	2	11.8

TABLE 15 Pain, health status and QoL measures – treatment comparison summaries over time and model estimates

Variable/week[a]	FURS N = 73	PCNL *N* = 86	Effect size (95% CI)[b]	*p*-value
Level of pain today[c]	Mean (SD): *n*	Mean (SD): *n*	*n* = 141	
Baseline	1.5 (2.5): 67	1.9 (2.7): 74		
Pre intervention	2.0 (2.7): 60	2.5 (3.1): 62	Ref[d]	
Week 1	3.7 (2.6): 52	3.1 (2.5): 59	0.81 (0.08 to 1.53)	0.029
Week 2	2.8 (2.5): 56	1.8 (2.4): 58	1.04 (0.11 to 1.96)	0.028
Week 3	1.9 (2.4): 53	1.4 (1.8): 61	0.63 (−0.12 to 1.37)	0.098
Week 4	1.9 (2.4): 50	1.3 (2.0): 61	0.71 (−0.13 to 1.54)	0.098
Week 5	1.3 (2.0): 54	0.9 (1.6): 59	0.36 (−0.23 to 0.95)	0.237
Week 6	1.3 (2.2): 52	1.1 (1.9): 58	0.31 (−0.45 to 1.06)	0.426
Week 7	1.2 (2.4): 51	1.3 (2.2): 60	−0.06 (−0.70 to 0.59)	0.857
Week 8	1.5 (2.7): 52	1.3 (2.3): 60	0.40 (−0.48 to 1.28)	0.378
Week 9	1.1 (2.1): 52	1.0 (1.9): 61	0.07 (−0.64 to 0.77)	0.852
Week 10	1.0 (2.0): 53	0.9 (1.9): 56	0.09 (−0.52 to 0.71)	0.770
Week 11	0.7 (1.7): 55	1.1 (1.9): 57	−0.15 (−0.75 to 0.44)	0.619
Week 12	0.8 (1.8): 54	0.9 (2.0): 60	−0.13 (−0.72 to 0.46)	0.672

TABLE 15 Pain, health status, and QoL measures – treatment comparison summaries over time and model estimates (*continued*)

Variable/week[a]	FURS N = 73	PCNL N = 86	Effect size (95% CI)[b]	p-value
***Stone pain in last 7 days*[e] Yes**	x (%): n	x (%): n	n = 142	
Baseline	29 (43.3): 67	34 (45.3): 75		
Pre intervention	29 (49.2): 59	34 (54.8): 62	Ref[d]	
Week 1	42 (89.4): 47	39 (63.9): 61	1.44 (1.11 to 1.86)[f]	0.005
Week 2	37 (68.5): 54	22 (36.7): 60	1.91 (1.21 to 3.02)[f]	0.005
Week 3	24 (48.0): 50	19 (31.1): 61	1.62 (0.92 to 2.85)[f]	0.097
Week 4	21 (45.7): 46	12 (20.7): 58	2.28 (1.23 to 4.22)[f]	0.008
Week 5	15 (29.4): 51	13 (22.4): 58	1.51 (0.88 to 2.59)[f]	0.132
Week 6	15 (30.0): 50	11 (19.6): 56	1.67 (0.84 to 3.33)[f]	0.144
Week 7	15 (29.4): 51	10 (17.5): 57	1.82 (0.96 to 3.46)[f]	0.066
Week 8	14 (27.5): 51	10 (18.5): 54	1.73 (0.86 to 3.49)[f]	0.124
Week 9	14 (27.5): 51	11 (19.0): 58	1.70 (0.99 to 2.91)[f]	0.053
Week 10	12 (22.6): 53	5 (9.3): 54	2.77 (1.30 to 5.88)[f]	0.008
Week 11	9 (16.7): 54	11 (20.4): 54	0.92 (0.48 to 1.75)[f]	0.797
Week 12	12 (23.1): 52	11 (18.0): 61	1.39 (0.68 to 2.86)[f]	0.369
No. days needed pain relief in last 7 days	Mean (SD): n	Mean (SD): n	n = 142	
Baseline	1.3 (2.2): 67	1.6 (2.5): 75		
Pre intervention	1.8 (2.5): 60	2.6 (2.9): 61	Ref[d]	
Week 1	4.8 (2.3): 52	4.8 (2.5): 60	0.01 (−0.87 to 0.89)	0.985
Week 2	3.4 (2.7): 56	2.9 (2.9): 59	0.50 (−0.81 to 1.81)	0.455
Week 3	2.0 (2.5): 52	2.2 (2.6): 61	−0.22 (−1.27 to 0.84)	0.686
Week 4	1.6 (2.4): 50	1.4 (2.3): 59	0.24 (−0.67 to 1.14)	0.609
Week 5	1.4 (2.3): 54	1.1 (2.2): 58	0.24 (−0.60 to 1.08)	0.573
Week 6	1.4 (2.5): 53	1.2 (2.2): 56	0.12 (−0.58 to 0.82)	0.739
Week 7	1.5 (2.5): 50	1.4 (2.5): 58	0.08 (−0.78 to 0.94)	0.853
Week 8	1.1 (2.2): 52	1.6 (2.7): 60	−0.46 (−1.24 to 0.31)	0.243
Week 9	1.1 (2.0): 51	1.4 (2.4): 61	−0.33 (−1.09 to 0.43)	0.397
Week 10	0.9 (1.9): 53	1.4 (2.5): 55	−0.46 (−1.18 to 0.26)	0.213
Week 11	1.0 (1.8): 55	1.2 (2.3): 56	−0.10 (−0.81 to 0.62)	0.789
Week 12	0.9 (2.2): 51	0.9 (2.0): 60	−0.05 (−0.76 to 0.67)	0.895
Worst level of pain since intervention	Mean (SD): n	Mean (SD): n	n = 139	
Week 1	7.8 (2.1): 52	6.0 (3.2): 60	Ref[d]	
Week 2	7.2 (2.5): 56	5.0 (3.3): 58	1.59 (0.50 to 2.69)	0.004
Week 3	6.5 (3.3): 53	4.7 (3.7): 60	1.43 (0.14 to 2.72)	0.030
Week 4	6.8 (3.0): 50	4.3 (3.6): 60	2.14 (1.10 to 3.18)	< 0.001

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 15 Pain, health status, and QoL measures – treatment comparison summaries over time and model estimates (*continued*)

Variable/week[a]	FURS *N* = 73	PCNL *N* = 86	Effect size (95% CI)[b]	*p*-value
Week 5	6.2 (3.4): 54	4.2 (3.5): 57	1.46 (0.35 to 2.56)	0.010
Week 6	6.5 (3.5): 53	4.0 (3.4): 57	2.08 (0.98 to 3.17)	< 0.001
Week 7	6.6 (3.4): 51	4.5 (3.5): 60	1.67 (0.53 to 2.81)	0.004
Week 8	6.8 (3.4): 52	4.5 (3.6): 59	1.88 (0.72 to 3.04)	0.002
Week 9	6.8 (3.3): 52	4.3 (3.6): 60	2.18 (1.06 to 3.29)	< 0.001
Week 10	6.6 (3.5): 52	4.4 (3.4): 56	2.01 (0.87 to 3.15)	0.001
Week 11	6.3 (3.6): 55	4.2 (3.5): 56	1.97 (0.87 to 3.06)	< 0.001
EQ-5D score[g]	Mean (SD): *n*	Mean (SD): *n*	*n* = 142	
Baseline	0.767 (0.283): 67	0.758 (0.274): 75		
Pre intervention	0.798 (0.199): 60	0.760 (0.230): 63	Ref[d]	
Week 1	0.614 (0.289): 52	0.687 (0.216): 60	−0.11 (−0.18 to −0.04)	0.002
Week 2	0.715 (0.257): 56	0.783 (0.172): 60	−0.09 (−0.17 to −0.01)	0.021
Week 3	0.760 (0.239): 50	0.826 (0.179): 61	−0.09 (−0.16 to −0.03)	0.004
Week 4	0.777 (0.254): 47	0.848 (0.175): 57	−0.10 (−0.18 to −0.02)	0.013
Week 5	0.814 (0.255): 54	0.886 (0.155): 56	−0.09 (−0.16 to −0.01)	0.027
Week 6	0.822 (0.265): 52	0.853 (0.202): 54	−0.03 (−0.09 to 0.03)	0.333
Week 7	0.810 (0.272): 51	0.824 (0.253): 58	−0.03 (−0.10 to 0.05)	0.455
Week 8	0.807 (0.268): 53	0.833 (0.238): 58	−0.04 (−0.13 to 0.04)	0.323
Week 9	0.828 (0.270): 51	0.858 (0.234): 57	−0.04 (−0.12 to 0.03)	0.279
Week 10	0.835 (0.271): 52	0.845 (0.249): 55	−0.03 (−0.10 to 0.04)	0.393
Week 11	0.828 (0.264): 51	0.846 (0.247): 55	−0.03 (−0.11 to 0.04)	0.414
Week 12	0.833 (0.258): 55	0.851 (0.236): 61	−0.01 (−0.09 to 0.07)	0.813
EQ-5D-VAS[h]	Mean (SD): *n*	Mean (SD): *n*	*n* = 138	
Baseline	73.2 (19.2): 66	71.3 (20.8): 72		
Pre intervention	75.1 (18.8): 59	73.9 (20.4): 62	Ref[d]	
Week 1	62.8 (21.5): 49	68.5 (21.1): 53	−5.52 (−13.76 to 2.72)	0.189
Week 2	70.2 (20.7): 53	75.4 (18.5): 56	−5.01 (−10.61 to 0.59)	0.079
Week 3	75.5 (20.1): 47	80.9 (18.6): 56	−5.54 (−12.16 to 1.08)	0.101
Week 4	73.9 (23.9): 43	79.5 (22.2): 54	−8.04 (−14.61 to −1.48)	0.016
Week 5	74.6 (25.3): 48	80.8 (22.4): 55	−7.66 (−14.77 to −0.56)	0.035
Week 6	78.4 (23.3): 45	79.9 (22.8): 54	−4.06 (−11.21 to 3.09)	0.265
Week 7	78.3 (22.8): 47	76.7 (24.1): 57	−0.59 (−7.52 to 6.34)	0.867
Week 8	76.4 (23.4): 49	76.2 (25.0): 56	−3.13 (−10.17 to 3.91)	0.383
Week 9	82.2 (18.0): 48	77.9 (24.1): 56	1.28 (−3.53 to 6.09)	0.602
Week 10	83.8 (18.7): 48	76.5 (23.8): 54	3.88 (−2.57 to 10.32)	0.238

TABLE 15 Pain, health status, and QoL measures – treatment comparison summaries over time and model estimates (*continued*)

Variable/week[a]	FURS N = 73	PCNL N = 86	Effect size (95% CI)[b]	p-value
Week 11	84.6 (18.5): 48	78.4 (24.3): 53	2.35 (−4.07 to 8.77)	0.474
Week 12	83.5 (15.6): 50	79.1 (22.3): 57	3.05 (−2.89 to 8.99)	0.315
SF-12 scores[i]				
Physical component	Mean (SD): *n*	Mean (SD): *n*	*n = 132*	
Baseline	45.7 (10.8): 63	44.0 (11.0): 71		
Pre intervention	46.4 (9.8): 56	44.5 (11.3): 59	Ref[d]	
Week 12	47.9 (11.2): 53	46.5 (11.5): 58	−1.34 (−4.56 to 1.87)	0.414
Mental component	Mean (SD): *n*	Mean (SD): *n*	*n = 132*	
Baseline	49.0 (10.5): 63	50.0 (10.9): 71		
Pre intervention	50.0 (10.2): 56	49.7 (10.5): 59	Ref[d]	
Week 12	50.3 (10.4): 53	50.2 (12.0): 58	−0.27 (−2.84 to 2.30)	0.836

n, number of people with a recorded response; *x*, number of 'Yes' responses to 'Stone pain in last 7 days'.
a Participants included in the models on average returned 9 to 10 weekly questionnaires (range 1–12).
b Models are mixed linear regression with time as an interaction robust error variance accounting for centre clusters unless otherwise specified: > 0 indicates FURS has a larger effect size for that outcome than PCNL.
c Levels of pain: (higher score indicates more pain).
d Ref = reference time point.
e Days needing pain relief: (higher score indicates more pain).
f Risk ratio based on a modified random effects multilevel Poisson regression with a log-link function and robust error variance to account for centre clusters:[53] > 1 indicates FURS participants having higher pain levels than PCNL participants.
g NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
h EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).
i SF-12 scores range from 0 to 100 (higher score is better).

Subgroup analyses

We explored the potential impact of three subgroup factors reported at baseline: BMI, stone density as measured by the Hounsfield unit, and stone size (the definitions are in *Appendix 2*, *Subgroup analyses and technical factors definitions*). *Appendix 7* gives the plots of the three PO models (main, *Figure 27*, sensitivity 1, *Figure 28* and sensitivity 2, *Figure 29*) and the secondary models (AUC from baseline, *Figure 30*, and the clinical measure of stone clearance, *Figure 31*) with 99% CIs. The AUC outcomes (*Figures 27–30*) all favour PNCL, but not statistically and span the original overall effect estimate. This is also the case for stone clearance at 12 weeks (*Figure 31*), although there the division into the stone density categories maintained some significance but gave little additional information.

Technical factors

We report the results for the main outcomes and covariates in the PO models (main and two sensitivities), the secondary AUC model (AUC from baseline to 12 weeks post intervention) and also for stone clearance at 12 weeks for the technical factors in *Appendix 8* (see *Tables 71–75*). These were not a between treatment arm comparison (since they are within a specified treatment) but were summaries of each of the outcomes and covariates for each technical factor across the two RCTs. They were restricted to time of the first intervention.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Chapter 5 Economic evaluation

Some parts of this chapter are reproduced with permission from Wiseman *et al.*[1] Cooper *et al.*[54] These are Open Access articles distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, even for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Some parts of the Methods section in this chapter are also reproduced with permission from Ahmed *et al.*[55] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY NC 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for non-commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by-nc/4.0/. The text below includes minor additions and formatting changes to the original text.

Introduction

Economic evaluation helps to make better informed decisions to adopt new healthcare technologies by providing information on the relative efficiency of the technologies under consideration. Resources are scarce in the NHS and healthcare technologies should be adopted or retained only if they provide good value for money. Therefore, the economic analysis is interested in all relevant cost and consequences resulting from the compared care pathways. Consequently, a longer time horizon was selected compared to that used in the statistical analysis. This chapter reports on the economic evaluation of FURS as the first treatment option in comparison to ESWL for stones size ≤ 10 mm in maximum dimension (RCT1) and to PCNL for stones size > 10 and ≤ 25 mm in maximum dimension (RCT2).[37] The chapter is divided into two main sections: (1) a within trial analysis that considers costs and consequences up to 12 months post randomisation and (2) a modelling extrapolation extending the analysis for a time horizon of 5 years.

Within-trial analysis

Objectives of the economic evaluation

The primary economic objective of the within-trial cost-effectiveness analysis was to assess the cost-effectiveness of FURS as the first treatment option in comparison to ESWL for stones ≤ 10 mm in maximum dimension or PCNL for stones > 10 and ≤ 25 mm in maximum dimension at 12 months post randomisation.

Methods

Established methods were followed for this economic evaluation[56,57] as well as to report its findings[58] with a prespecified protocol and health economics analysis plan (HEAP).

Study design and participants

The PUrE trial protocol[37] and *Chapter 2* provide details of the design. The within-trial economic analysis follows the ITT principle and is based on the same participants randomised and considered for the main trial analysis.

Cost and outcome assessment

Costs and outcomes were assessed via the trial CRFs and participant-completed questionnaires (see *Figure 1* and *Table 1*). The research nurses completed a CRF at the time of surgery providing details of the operative procedures, complications and resource used in hospital. Costs of the initial intervention procedures were estimated from resource use data recorded on the CRFs coupled with routine unit cost data. Costs associated with subsequent contacts with primary care (due to kidney stones) were estimated from patient questionnaires at pre intervention, 12 weeks post intervention and 12 months post randomisation. QALYs were estimated from patients' responses to the EQ-5D-5L at baseline, pre intervention, weekly up to 12 weeks post intervention and at 12 months post randomisation. EQ-5D-5L data were also collected pre and post intervention for additional interventions occurring outside the 12-week post-intervention window or any hospitalisation. The schedule for economic data collection is shown in *Table 1*.

Assessment of health service costs

NHS costs for health service use in both primary and secondary care for each trial participant were estimated. The economic evaluation seeks to inform the efficient allocation of the NHS budget; therefore, the base case analysis adopted a health service perspective. Costs are expressed in Great British pounds for the 2020–1 price year.

Cost of the primary interventions

Costs of the initial intervention were estimated from resource use data recorded in the PUrE Intervention CRF for each participant. The CRF captured date of admission, treatment and discharge, the type of procedure carried out ESWL, FURS or PCNL, procedure time (i.e. time procedure started and finished for ESWL, and time of entry into anaesthetic room, entry to and exit from the operating room, and time of exit recovery room for FURS and PCNL), intraoperative complications, and destination post operation or recovery room.

The primary costing approach assigned costs to individual components of resource use to capture patient-level variation in costs (*Table 16*). A number of assumptions have been made on discussion with the clinical colleagues in the project management team (three clinicians). In the UK, ESWL is conducted as a day case procedure by a nurse and a technician. The PUrE intervention CRF stated the total time of the procedure and to this, 20 minutes were added to allow for patient preparation and discharge. Details were collected on the type of lithotripter (i.e. mobile or fixed) and corresponding pro rata costs were allocated assuming throughput of 200 patients per year and a 5-year equipment depreciation. For FURS and PCNL, the time in the anaesthetic room was costed assuming a cost per hour (incorporating overheads) for a consultant anaesthetist and an anaesthetist nurse.[59] For time in theatre, we defined the grade of the surgeon conducting the intervention through discussion with the clinical collaborators and costed using the unit cost per hour applied to the final total time in theatre.[60] For PCNL, a radiographer was assumed for whole procedure for all cases plus 1 hour for an interventional radiologist for half of the cases. Nursing staff was costed at the requirement for general day surgery: one anaesthetic nurse, a scrub nurse, and two further theatre nurses. In addition, a published unit cost was applied for time in theatre to reflect the average cost of other staff, supplies and consumables, and allocated capital charges and overheads.[61] This detailed unit cost of theatre time is only available for Scottish hospitals. Therefore, the average Scottish estimate (£596 per hour, excluding medical and nursing staff) was applied to the time in theatre for the base-case analysis. In addition to this, we applied the unit costs of major consumable items specific to the alternative procedures (see *Appendix 9*, *Table 76*).

Time in recovery following surgery was costed using the unit cost of a Band 6 nurse (inclusive of overheads) assuming one-to-one care and the time on the ward following recovery using an estimate of the cost per excess bed day following index procedure[62] adjusted by inflation. Equipment costs were allocated assuming a 200-, 100- and 40-patient-per-year

TABLE 16 Unit costs (NHS perspective)

Resource	How measured	Source of measurement	Unit cost (£)	Source of valuation
Time in anaesthetic room	Time in hours	PUrE intervention form	171 per hour	Band 6 nurse (£51) + consultant anaesthetist (£120); Unit Costs of Health and Social Care, 2021[59]
Time in theatre	Time in hours	PUrE intervention form		
Surgery time		PUrE intervention form	Consultant urological surgeon (122 per hour) Consultant anaesthetist (120) Senior nurse (62) Nurse (floor) (51) 2 nurses (41)	Unit Costs of Health and Social Care, 2021[59]
Anaesthetist time			Consultant: (120 per hour) Anaesthetic nurse (62)	Unit Costs of Health and Social Care, 2021[59]

continued

Copyright © 2025 Wiseman *et al*. This work was produced by Wiseman *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 16 Unit costs (NHS perspective) (*continued*)

Resource	How measured	Source of measurement	Unit cost (£)	Source of valuation
Theatre costs (excluding medical and nursing staff)			596 per hour	Table R140, ISD 2021
Recovery room	Time in hours	PUrE intervention form	51 per hour	Band 6 nurse. Unit Costs of Health and Social Care, 2021[59]
Procedure consumables	Source: personal communication Samuel McClinton, Daron Smith and Oliver Wiseman, personal communication, May 2023			
Perioperative complication costs	Number and type	PUrE intervention form	Various based on recorded reasons and procedures	NHS Reference costs[62]
Readmissions	Number and type of admission	PUrE CRFs	Various based on recorded reasons and procedures	NHS Reference costs[62]
Outpatient appointments	Number of appointments	PUrE supplementary form	149 per attendance	Code 101, Urology, consultant led. NHS Reference costs[62]
Accident and emergency	Number of attendances		175 per attendance	Code 180, Accident and Emergency – non-consultant led NHS Reference costs[62]
Primary care contacts		Patient questionnaires		Unit Costs of Health and Social Care, 2021[59]
GP contact	Number of contacts		39.23 (per contact lasting 9.22 minutes)	
GP nurse	Number of contacts		42 (per hour)	

ISD, information statistics division.

throughput for ESWL, FURS and PCNL, respectively (Samuel McClinton, Daron Smith and Oliver Wiseman, personal communication, 26 May 2023) and 5-year equipment depreciation (see *Appendix 9*, *Table 76*).

As an alternative approach to costing the initial procedure episodes, each patient record was mapped to the appropriate Health Care Resource Group (HRG) and costed using the relevant NHS reference cost.[62] The core HRG codes LB36Z (Extracorporeal Lithotripsy) either as day case (elective with admission and discharge on the same day), elective inpatient, or non-elective short stay (if admitted as elective or emergency with overnight stay, respectively). For FURS and PCNL the HRG codes LB65E (Major Endoscopic, Kidney or Ureter Procedures, 19 years and over, with CC Score 0–2) and LB75B (PCNL with CC Score 0–1) were used, respectively. The unit costs for FURS and PCNL were applied as either a day case (patient admitted and discharged on the same day) or as an elective inpatient admission (stay ≥ 1 day).

Costs of hospital admissions, perioperative complications and hospital readmissions

The clinical management of any procedure resulting from perioperative complications leading to prolonged hospital stay as well as any hospital admissions were based on the NHS reference cost. These included events, such as urinary tract infection, haematuria, abdominal pain (*Chapter 3*, *Complications* and *Chapter 4*, *Complications*). Details for the type of complication experienced and the procedures undertaken (e.g., stent insertion, stent removal, ESWL, FURS) were obtained from the PUrE Intervention CRF, the second treatment session CRF, the follow-up CRF, the supplementary treatment CRF, and the SAEs forms. These events were costed using HRG-based reference costs.[62]

Costs of hospital outpatient and primary care healthcare utilisation

Secondary care outpatient visits incurred over the 12-month follow-up period were obtained from hospital records using supplementary CRFs costed using the NHS reference cost for standard surgery outpatient visit.[62] Primary care visits over the follow-up period were collected from the patient questionnaires at 12 weeks post intervention and 12 months post randomisation. These contacts were costed using the Unit Costs of Health and Social Care[60] by multiplying the number of visits by the appropriate unit cost.

While participants were asked about prescribed medications for their kidney stones condition, the questionnaire had no details of the actual medication prescribed. On discussion with the trial management team and due to the possible very low cost of these medications (i.e. mostly over the counter painkillers), a decision was made to exclude the medication prescribed in primary care from the costing exercise.

Indirect costs

Participants were asked about the number of sick days needed from the day of their intervention and up to 12 months post randomisation. The number of days were multiplied by the average hourly pay (gross) by age group for 2022 published by UK Office of National Statistics,[63] assuming a full-time shift of 8 hours per day of work, to value the time off work. In addition, participants were asked about out-of-pocket expenditure on medications, private health care and the cost of travelling for their last hospital admission, hospital consultation and GP appointment. Participants were asked for the cost of the trip, for those travelling on public transport. Individuals using private cars were asked to estimate their trip mileage with the trip mileage valued at £0.45 per mile following the valuation of cost of using private cars for business by the UK government.[64] As these costs are not costs to the NHS, these are reported separately and were not incorporated in the main economic analysis.

Outcome measures

Quality-adjusted life-years were estimated using the EQ-5D-5L questionnaire completed by participants at baseline, pre intervention, weekly up to 12 weeks post intervention, and at 12 months post randomisation. In addition, EQ-5D-5L was administered pre and post interventions or hospitalisations occurring outside the 12-week post-index intervention window (see *Table 1*). The methods recommended by NICE[65] to map from the EQ-5D-5L to the EQ-5D-3L utility scores and based on Hernández-Alava *et al.* 2020[66] were used to assign a utility score to each participant response. QALYs were then estimated calculating the AUC approach, assuming linear change in utility between the observed follow-up time points.

Statistical analysis of trial economic data

Aggregating costs and effects

Following the ITT principle, resource use, cost and health outcome data were summed up for each participant for the 12-month follow-up period and compared by treatment allocation group. Continuous and count variables are reported as means (± SDs), and dichotomous and categorical variables as absolute numbers (and percentages).

Missing data

Unless the data are missing completely at random, reliance on complete case data for the cost-effectiveness analysis can introduce bias. The total estimated cost is the sum of numerous components over the observed follow-up period of the trial. Besides, QALYs can be computed only where participants have responded to the QoL questionnaires at every relevant follow-up point. Therefore, economic evaluations based on trial participant level data are likely to suffer missing data. We implemented multiple imputation as part of the primary analysis, using chained equations to generate multiple data sets with plausible fitted values assigned for missing cost and utility elements. The imputation was conducted at the cost category variable level (e.g., cost for anaesthetic, operating, and recovery rooms, equipment, consumables, length of stay, etc.) and the EQ-5D score level. The imputation model included all the variables in the analysis model. Rubin's rules were used to pool estimates across the multiple imputation data sets.

Incremental cost-effectiveness analysis

The trial-based economic analysis estimated the incremental cost and QALYs between (1) FURS as the first treatment option in comparison to ESWL for stones ≤ 10 mm in maximum dimension (RCT1) and (2) FURS as the first treatment option in comparison to PCNL for stones > 10 and ≤ 25 mm in maximum dimension (RCT2) up to 12 months post randomisation. General linear regression models (GLM) with appropriate variance and link functions were fitted, with adjustment for baseline EQ-5D score. The selection of the family function was based on modified Park test, while a modified Hosmer and Lemeshow, Pearson correlation, and Pregibon link tests were used to select the link function[57]

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

(see *Appendix 9*, *Table 77* for details). These test results suggested a Poisson family with identity link for costs and a Gaussian family with identity link for QALY for RCT1 and Gamma family with identity link for costs and Gaussian family with identity link for QALYs for RCT2. Adjusted mean values by treatment allocation group, and the incremental differences between groups were obtained using the methods of recycled predictions.[57] The incremental cost-effectiveness ratio (ICER) for FURS and ESWL (RCT1) and FURS and PCNL (RCT2) were calculated as the difference in the (arithmetic) mean cost divided by the difference in (arithmetic) mean QALYs. Variance surrounding the joint incremental costs and effects was characterised using non-parametric bootstrapping. Bootstrapping involves sampling with reposition from the original trial participants a new sample of equal size of the initial sample (e.g. *n* = 230 and *n* = 231 for FURS and ESWL, respectively, for RCT1). These data are analysed in a similar way as for the original trial data and results for average cost and QALYs obtained. The process is repeated several times (i.e. 1000) and a distribution for the average cost, average QALYs, and average cost and QALY differences between trial groups are obtained. The simulation results are presented graphically using the cost-effectiveness plane (scatterplots) and cost-effectiveness acceptability curves (CEACs). The probability of FURS being cost-effective at 12 months is reported for threshold values of willingness to pay per additional QALY of £13,000, £20,000 and £30,000.[67]

Sensitivity analyses

Subgroup analysis was conducted to explore cost-effectiveness of FURS according to predefined subgroups: Normal/overweight/obese/extreme obesity defined according to BMI score; stone size (RCT2 only: 10–15 vs. > 15 mm); stone density (CTKUB Hounsfield units comparing ≤ 1000 vs. > 1000); skin-to-stone distance (< 120 vs. ≥ 120 mm) (see *Appendix 2* for subgroup definition). Treatment allocation by subgroup interaction terms were defined in the regression models for this analysis. A further analysis was conducted using HRG-based costing for the index surgical intervention.

Results

RCT1

Resource use and costs

Resource use and cost from the NHS perspective by treatment allocation are reported in *Table 17*. Data for participants going through surgical procedures resulted in similar average time in anaesthetic room, theatre room and recovery room. As expected, fewer participants went through surgery in the ESWL group (*n* = 7) compared to the FURS group. Conversely, while ESWL were conducted using similar length of time in both study groups, fewer participants went through ESWL in the FURS allocation group (*n* = 17) compared with ESWL group (*n* = 191).

Table 17 also reports follow-up resource use up to 12 months post randomisation. A similar number of participants were readmitted into hospital in both groups (36%, *n* = 83 for FURS vs. 34%, *n* = 79 for ESWL), but those participants admitted in the latter group spent slightly more time in hospital. More participants attended outpatient appointments in the ESWL (30%, *n* = 69) compared with FURS group (24%, *n* = 55) although there was a higher mean number of visits for those in the FURS group. Participants reported a slightly lower mean number of GP and mean number of GP nurse contacts in the FURS group.

Costs for the index treatment episode were higher for FURS when either micro-costing or the HRG-based costing approach were used. However, additional procedures and hospital readmissions costs were higher for the ESWL group but the cost for outpatient appointments or GP visits are similar between the study groups. The 12-month care pathway mean cost for the NHS was £3408 and £2041 for the ITT with FURS and ESWL, respectively; giving an unadjusted difference of £1367 in favour of ESWL. Over 70% of the PUrE participants had complete data on costs (*n* = 330 of 461) and contributed to these mean costs calculation. A cost of zero for the index intervention was included for those individuals waiting for their index intervention at 12 months (20 for the FURS group and 6 for the ESWL group). *Table 17* also shows the care pathway cost when HRGs were used to value the initial procedures. In this case, the 12-month care pathway mean cost for the NHS was £4084 and £2517 for the ITT with FURS and ESWL, respectively; resulting in FURS being £1567 more costly than ESWL.

TABLE 17 Health service resource use and costs by treatment allocation

Variable	No. of observations	FURS *N* = 230		ESWL *N* = 231	
Resource use – index procedure[a]					
Time in anaesthetic room (minutes); mean [SD; median; (IQR); *n*]	178	15 [12; 15; (9–20); 171]		15 [5; 15; (10–20); 7]	
Time in theatre (minutes); mean [SD; median; (IQR); *n*]	174	61 [26; 56; (42–75); 167]		59 [23; 60; (45–63); 7]	
Time in recovery (minutes); mean [SD; median; (IQR); *n*]	175	62 [52; 49; (30–84); 168]		48 [47; 35; (20–65); 7]	
Length of stay (days); mean [SD; median; (IQR); *n*]	417	0.29 [0.6; 0; (0–0); 199]		0.06 [0.5; 0; (0–0); 218]	
Total procedure time ESWL (minutes); mean [SD; median; (IQR); *n*]	208	30 [12; 30; (20–40); 17]		35 [16; 30; (24–44); 191]	
Resource use – additional procedures		ESWL	FURS	ESWL	FURS
Number of further Procedures; *n*	461	8	10	128	3
Hospital readmission; *n* (%)	461	83 (36.1)		79 (34.2)	
Readmission days; mean (SD)	162	1.2 [1.7; 0.5; (0.5–1)]		1.34 [2; 0.5; (0.5–1)]	
Outpatient visits; *n* (%)	461	55 (23.9)		69 (29.9)	
Outpatient visits; mean [SD; median; (IQR)]	124	1.33 [0.7; 1; (1–1)]		1.23 [0.5; 1; (1–1)]	
GP contacts; mean [SD; median; (IQR); *n*]	358	0.57 [1.5; 0; (0–1); 183]		0.61 [1.4; 0; (0–1); 175]	
Practice nurse contacts; mean [SD; median; (IQR); *n*]	359	0.11 [0.4; 0; (0–0); 184]		0.07 [0.3; 0; (0–0); 175]	
Costs (£)					
Index episode cost; mean (SD)					
Micro-costing	425	2418 (1242)		740 (566)	
HRG-based costing	456	2944 (1462)		956 (679)	
Additional treatment costs; mean (SD)					
Micro-costing	461	118 (561)		377 (513)	
HRG-based costing	461	208 (855)		558 (651)	
Readmission costs for further treatment; mean (SD)	*461*	*663 (1283)*		*939 (1878)*	
Outpatient costs; mean (SD)	*461*	*55 (118)*		*69 (125)*	
Primary care costs; mean (SD)	*358*	*27 (62)*		*27 (57)*	
Total NHS cost (£); mean (SD)					
Micro-costing	330	3408 (1832)		2041 (1866)	
HRG-based costing	356	4084 (2052)		2517 (2097)	

IQR, interquartile range.
a Times in anaesthetic, operating, recovery room for ESWL column correspond to participants going through FURS or PCNL as first procedure; procedure time ESWL for FURS column corresponds to individuals going through ESWL in the FURS group.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Utility scores and quality-adjusted life-years

The EQ-5D mean health state utility scores based on the responses to the EQ-5D-5L and the mapping algorithm[66] to the EQ-5D-3L scores by treatment allocation are reported in *Table 18*. QoL data were collected at baseline, pre intervention, weekly post intervention up to 12 weeks (3 months), and 12 months post randomisation. EQ-5D-5L data were collected when participants needed additional interventions beyond the 12-week window and treatment-related hospitalisations. Data attached to the index intervention are shown for observations within the 12-month trial follow-up only, with QALYs calculated with data from baseline, 12 months and any observation between these two time points. For the QALY calculation in *Table 18*, complete data were defined as those individuals with data for, at least, baseline and 12-month EQ-5D scores.

An imbalance in the mean EQ-5D scores favouring the ESWL group is seen at baseline and pre intervention. This difference favouring ESWL continues up to 6 weeks post treatment when the difference turns to favour FURS; probably reflecting the negative QoL impact of needing repeated ESWL sessions in the ESWL group. At 12 months post randomisation, the difference in the mean EQ-5D score favours ESWL. To be noted, the data include 20 and 6 participants that were waiting for their index intervention (see *Table 6*) at 12 months post randomisation. In the last row in *Table 18* the average QALYs for the 12-month follow-up are reported. From a visual inspection of the SDs in *Table 18*, a small non-significant unadjusted mean QALY difference of 0.011 QALYs favouring ESWL emerges. However, this QALY calculation is based on data for 60% of participants and is not adjusted for baseline QoL.

Cost–utility analysis results

The base-case cost–utility analysis results based on the multiple imputation data set and adjusted by baseline EQ-5D utility score are reported in *Table 19*. The 12-month mean costs per participant were £2223 and £3362 for the ITT with ESWL and FURS groups, respectively, resulting in an adjusted cost difference of £1138 [95% CI £646 to £1631]. The mean QALYs per participant were 0.787 and 0.804 for the ESWL and FURS groups, respectively, producing an adjusted QALY difference of 0.017 QALYs (95% CI −0.008 to 0.043) for the 12-month follow-up period. Therefore, ITT with FURS results in significantly higher mean costs and non-significantly higher mean QALYs compared with ESWL.

TABLE 18 Health state utility scores by treatment allocation

Variable	No. of observations	FURS *N* = 230	ESWL *N* = 231
EQ-5D score			
Baseline; mean (SD)	459	0.769 (0.23)	0.817 (0.2)
Pre treatment; mean (SD)	358	0.765 (0.23)	0.808 (0.22)
Post treatment			
1 week; mean (SD)	318	0.628 (0.26)	0.789 (0.22)
2 weeks; mean (SD)	323	0.771 (0.22)	0.814 (0.22)
3 weeks; mean (SD)	290	0.801 (0.25)	0.816 (0.22)
4 weeks; mean (SD)	295	0.812 (0.25)	0.819 (0.23)
5 weeks; mean (SD)	296	0.816 (0.24)	0.84 (0.2)
6 weeks; mean (SD)	297	0.836 (0.23)	0.839 (0.2)
7 weeks; mean (SD)	295	0.844 (0.23)	0.831 (0.23)
8 weeks; mean (SD)	286	0.848 (0.22)	0.837 (0.22)
9 weeks; mean (SD)	290	0.844 (0.23)	0.826 (0.23)
10 weeks; mean (SD)	288	0.856 (0.23)	0.83 (0.23)
11 weeks; mean (SD)	289	0.844 (0.24)	0.84 (0.23)

TABLE 18 Health state utility scores by treatment allocation (*continued*)

Variable	No. of observations	FURS *N* = 230	ESWL *N* = 231
Three months post treatment; mean (SD)	302	0.844 (0.25)	0.829 (0.24)
Additional EQ-5D form; mean (SD)	248	0.698 (0.3)	0.759 (0.21)
Pre-supplementary treatment; mean (SD)	167	0.694 (0.26)	0.815 (0.24)
Twelve months; mean (SD)	283	0.809 (0.24)	0.798 (0.25)
QALY; mean (SD)	282	0.801 (0.21)	0.812 (0.2)

The ICER between FURS and ESWL was £65,163. Thus, ITT with FURS results, on average, in higher QALYs at a cost of £65,163 per QALY gained. This is above the usual cost-effectiveness threshold considered in the UK decision-making (i.e. below £30,000 per QALY gained).[65] Moreover, given the cost and QALY differences seen in *Table 19*, either a difference of 0.057 QALYs or £349 are needed to reach an ICER of £20,000. These differences are outside the 95% CIs for QALYs (−0.008; 0.043) and costs (646; 1631) (see *Table 19*).

Figure 6 shows the cost-effectiveness plane with results from the 1000 bootstrapped iterations of the regression analysis nested multiple imputation. The incremental costs for FURS against ESWL are reported in the vertical axis and the incremental QALYs in the horizontal axis. All the iterations show positive mean costs differences, meaning that FURS was more costly than ESWL for all the iterations. For 11% of the iterations ESWL produced more QALYs resulting in a non-statistically significant difference of 0.017 favouring FURS. The red dashed line in *Figure 6* represents the £20,000 cost-effectiveness threshold and separates iterations that result in either trial group being cost-effective. Those iterations sitting to the right and below the red dashed line indicate that FURS is cost-effective. Conversely, iterations to the left and above the red dashed line indicate that ESWL is cost-effective. All but five of the 1000 iterations sit above and to the left of the £20,000 threshold dashed line; that is just 0.5% of the iterations show QALY differences large enough (and cost difference small enough) for FURS to be considered cost-effective (see *Figure 6*). Increasing the threshold value (i.e. moving the red line anticlockwise from the zero point), would result in higher proportions of the iterations crossing the red line to the south and right in *Figure 6*. The effect of this is represented in *Figure 7*.

The CEACs for FURS and ESWL are reported in *Figure 7*. The curves show the probability of either trial group being cost-effective at alternative cost-effectiveness threshold values. The CEAC for ESWL shows a 99.9% probability of being cost-effective at £20,000 threshold value. This probability decreases when the threshold value increases, reflecting the higher valuation of the additional QALYs resulting from FURS. However, the probability of ESWL being cost-effective is above 65% at relatively high cost-effectiveness threshold (i.e. £50,000) and the cost-effectiveness threshold has to be increased to £68,000 for a 50% chance of being cost-effective for both strategies (data not shown).

Alternative costing method

The results for the HRG-based alternative costing methodology for the index interventions are summarised in *Table 20*. Mean cost for ESWL and FURS are higher with the HRG-based costing compared with the base case using micro-costing. However, FURS cost increases more than ESWL; therefore, the analysis generates a larger cost difference between the study groups (£1372 compared to £1138) resulting in a higher ICER (£78,537) and an even lower probability for FURS being cost-effective at the usual threshold values used for decision-making in the UK NHS.

Indirect costs

Results on productivity costs (time off work) are reported in *Table 78* in *Appendix 9*. These show that participants took an average of 2.3 and 6.6 days off work in the ESWL and FURS groups, respectively for the initial intervention and about half a day in both groups for any other admissions up to 12 months post randomisation. The total mean productivity cost associated to time off work was £416 (SD 2028) and £1012 (SD 4686) for ESWL and FURS groups,

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 19 Base-case analysis results – multiple imputation: RCT1

Intervention	Mean total cost (£)	Incremental cost (£)[a]	Mean total QALYs	Incremental QALY[a]	ICER (£/QALY)
ESWL	2223		0.787		
FURS	3362	1138 (646; 1631)	0.804	0.017 (−0.008; 0.043)	65,163

a Differences adjusted by baseline EQ-5D score (95% CI).

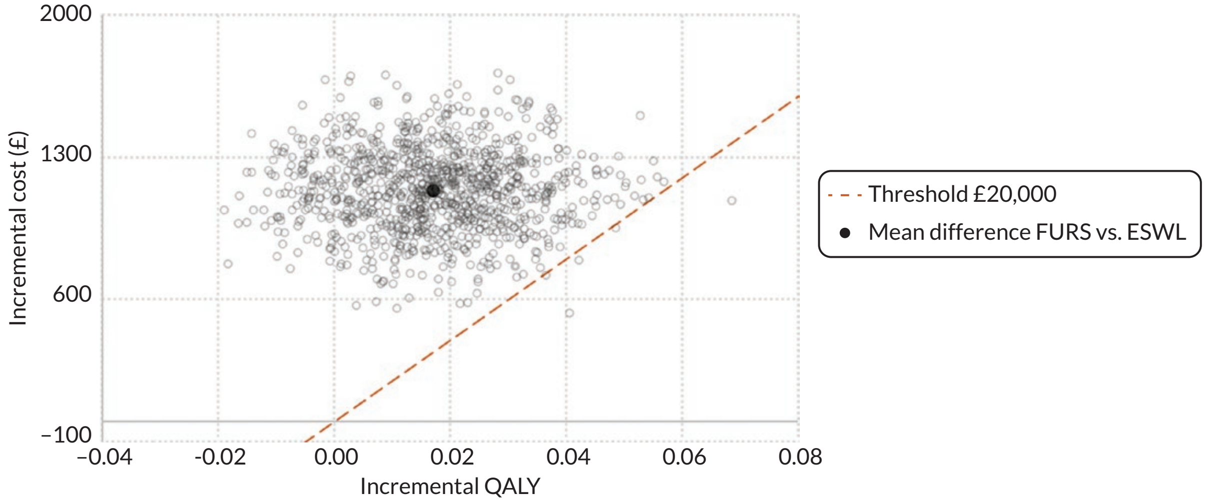


FIGURE 6 Trial-based incremental cost-effectiveness scatterplot for FURS vs. ESWL (base case; imputed data – 1000 bootstrap iterations). Note: Iterations below the red dashed £20,000 threshold line means FURS is cost-effective and above the line ESWL is cost-effective.

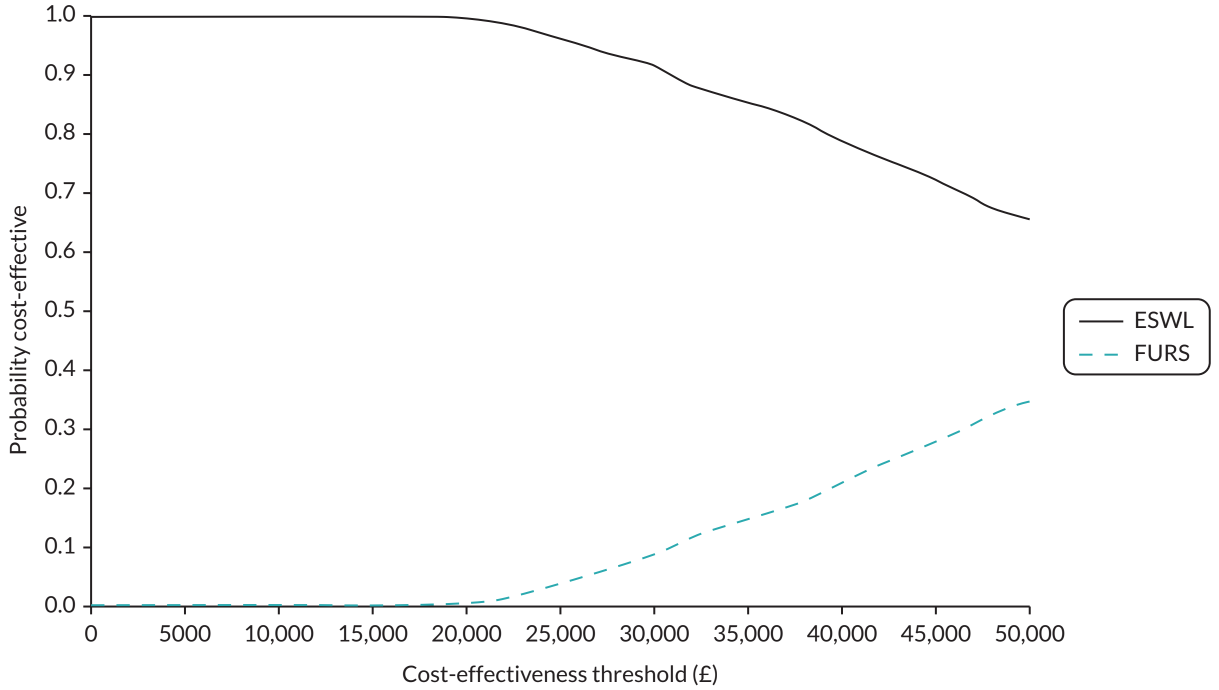


FIGURE 7 Cost-effectiveness acceptability curves for FURS and ESWL groups (trial-based analysis; base case; imputed data – 1000 bootstrap iterations).

TABLE 20 Trial-based incremental cost-effectiveness analysis by alternative costing methods

Intervention	Mean total cost (£)	Incremental cost[a] (£)	Mean total QALYs	Incremental QALY[a]	ICER (£/QALY)	Probability cost-effective £13,000 (%)	£20,000 (%)	£30,000 (%)
Base-case analysis (HRG-based costing)								
ESWL	2527		0.787			100	99.8	96.5
FURS	3899	1372 (856; 1888)	0.804	0.017 (−0.008; 0.043)	78537	0	0	0

Note
a Differences adjusted by baseline EQ-5D score (95% CI)..

respectively. *Table 78* also report the out-of-pocket expenditure. Only 45 individuals from 183 answering the question declared buying medications over the counter in the ESWL group and 52 of 173 in the FURS group, with the mean expenditure in medications declared being £4.60 and £7.60 for the ESWL and FURS groups, respectively. Only two participants declared buying private health care in the ESWL group and five in the FURS group, resulting in average costs of < £1. Finally, *Table 78* presents out-of-pocket travelling costs for hospital admission, outpatient consultation, and GP visits. The mean costs of travel are all below £9 with no meaningful differences between the groups on visual inspection of the standard errors.

RCT2

Resource use and costs

In *Table 21* the resource use and costs form the NHS perspective for RCT2 are reported. PCNL surgery took longer than FURS, with 20, 118 and 84 average minute times in anaesthetic, operating, and recovery rooms, against 15, 88 and 58 minutes for FURS. Length of stay in hospital for the initial operation was also longer for PCNL. However, more procedures were needed after the initial intervention in the FURS group and 53% (*n* = 39) participants needed to be readmitted into hospital (vs. 40%, *n* = 34 in PCNL group) with relatively longer average hospital stay (1.82 days for FURS vs. 1.22 fir PCNL). The FURS group also showed a higher proportion of participants attending hospital outpatient visits (33%, *n* = 24 for FURS vs. 27%, *n* = 23, for PCNL). Questionnaire respondents reported a higher number of average GP visits for PCNL and a similar average number of GP nurse visits.

The average costs of the initial intervention were higher for PCNL regardless of the method used to estimate costs (i.e., micro-costing or HRG-based costing). However, the average cost difference was larger when HRG-based costing was used (£2416 for HRG-based costing vs. £618 for micro-costing). The average costs of subsequent procedures, hospital readmissions and outpatient visits were higher for the FURS group, but the average cost of primary care visits was lower. These higher costs of subsequent procedures and hospital readmissions reverse the cost difference in favour of FURS for the initial intervention when micro-costing was used. The 12-month care pathway mean cost for the NHS was £5685 and £4708 for the ITT with FURS and PCNL, respectively; giving an unadjusted difference of £977 in favour of PCNL. *Table 21* also shows the care pathway cost when HRGs were used to value the initial procedures. In this case, the 12-month care pathway mean cost for the NHS was £6091 and £6907 for the ITT with FURS and PCNL, respectively, resulting in FURS being £817 less costly than PCNL.

Utility scores and quality-adjusted life-years

The utility scores generated from responses to the EQ-5D-5L and mapped to the EQ-5D-3L scores are reported in *Table 22*. The mean EQ-5D scores are slightly higher for the FURS group at baseline and pre-index interventions. Mean EQ-5D scores reduced for both study groups at 1 week after their initial intervention and show a steady recovery up to week 6, when mean scores stabilise. However, week-by-week post-treatment scores are slightly higher for PCNL. At 3 months post treatment, the EQ-5D scores are higher for both groups compared with the scores at baseline and pre treatment. Mean EQ-5D scores converge to the baseline values at 12 months; however, scores at 12 months included individuals waiting for their index intervention (10 and 8 for the FURS and PCNL groups, respectively). QALYs reported

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 21 Health service resource use and costs by treatment allocation: RCT2

Variable	No. of observations	FURS *N* = 73		PCNL *N* = 86	
Resource use – index procedure					
Time in anaesthetic room (minutes); mean [SD; median; (IQR); *n*]	132	15 [12; 13; (9–21); 61]		20 [15; 18; (10–28); 71]	
Time in theatre (minutes); mean [SD; median; (IQR); *n*]	130	88 [39; 88; (60–105); 61]		118 [47; 104; (85–149); 69]	
Time in recovery (minutes); mean [SD; median; (IQR); *n*]	131	58 [59; 40; (26–62); 62]		84 [73; 63; (44–105); 69]	
Length of stay (days); mean [SD; median; (IQR); *n*]	136	0.6 [1.3; 0; (0–1); 62]		1.77 [1.8; 1; (1–2); 74]	
Resource use – additional procedures		**ESWL**	FURS	**ESWL**	FURS
Number of further procedures; *n*	159	1	14	0	3
Hospital readmission; *n* (%)	159	39 (53.4)		34 (39.5)	
Readmission days; mean (SD)	73	1.82 [2.5; 0.5; (0.5–1.5)]		1.22 [1.5; 0.5; (0.5–1)]	
Outpatient visits; *n* (%)	159	24 (32.9)		23 (26.7)	
Outpatient visits; mean [SD; median; (IQR)]	47	1.5 [1.1; 1; (1–1.5)]		1.35 [0.7; 1; (1–2)]	
GP contacts; mean [SD; median; (IQR); *n*]	130	0.97 [2.4; 0; (0–1); 70]		1.27 [3.9; 0; (0–1); 60]	
Practice nurse contacts; mean [SD; median; (IQR); *n*]	130	0.4 [1.2; 0; (0–0); 70]		0.38 [2.3; 0; (0–0); 60]	
Costs (£)					
Index surgical episode cost; mean (SD)					
Micro-costing	152	3067 (1585)		3685 (1858)	
HRG-based costing	159	3463 (1914)		5879 (3032)	
Additional treatment costs; mean (SD)					
Micro-costing	159	606 (1339)		113 (608)	
HRG-based costing	159	665 (1433)		117 (618)	
Readmission costs for further treatment; mean (SD)	*159*	*1494 (2573)*		*596 (1571)*	
Outpatient costs; mean (SD)	*159*	*83 (166)*		*63 (122)*	
Primary care costs; mean (SD)	*130*	*55 (135)*		*66 (242)*	
Total NHS cost (£); mean (SD)					
Micro-costing	126	5685 (3565)		4708 (2848)	
HRG-based costing	130	6091 (3724)		6907 (3488)	

IQR, interquartile range.

in *Table 22* were calculated for 66% for the participants (*n* = 105 of 159) resulting in 0.777 and 0.781 QALYs for FURS and PCNL, respectively, and an unadjusted mean difference of 0.004 QALYs.

Cost–utility analysis results

The cost–utility analysis results for the base case are reported in *Table 23*. Micro-costing was used for the index intervention and multiple imputation to deal with missing data. The 12-month mean cost per participant were £4565

TABLE 22 Health state utility scores by treatment allocation: RCT2

Variable	No. of observations	FURS *N* = 73	PCNL *N* = 86
EQ-5D score			
Baseline; mean (SD)	158	0.775 (0.28)	0.76 (0.27)
Pre treatment; mean (SD)	121	0.809 (0.2)	0.774 (0.22)
Post treatment			
1 week; mean (SD)	100	0.602 (0.32)	0.678 (0.22)
2 weeks; mean (SD)	105	0.727 (0.24)	0.763 (0.19)
3 weeks; mean (SD)	100	0.77 (0.23)	0.818 (0.18)
4 weeks; mean (SD)	92	0.801 (0.24)	0.839 (0.18)
5 weeks; mean (SD)	98	0.814 (0.26)	0.878 (0.16)
6 weeks; mean (SD)	94	0.813 (0.28)	0.836 (0.21)
7 weeks; mean (SD)	95	0.804 (0.28)	0.807 (0.27)
8 weeks; mean (SD)	93	0.796 (0.28)	0.843 (0.21)
9 weeks; mean (SD)	91	0.832 (0.27)	0.872 (0.19)
10 weeks; mean (SD)	91	0.821 (0.29)	0.864 (0.22)
11 weeks; mean (SD)	88	0.824 (0.28)	0.853 (0.22)
Three months post treatment; mean (SD)	95	0.827 (0.27)	0.859 (0.22)
Additional EQ-5D form; mean (SD)	102	0.768 (0.24)	0.752 (0.25)
Pre-supplementary treatment; mean (SD)	46	0.743 (0.22)	0.738 (0.18)
Twelve-month questionnaire; mean (SD)	106	0.759 (0.36)	0.763 (0.28)
QALY; mean (SD)	105	0.777 (0.25)	0.781 (0.21)

and £5298 for ITT with PCNL and FURS, respectively, resulting in an adjusted non-significant mean difference of £733 (PCNL less costly). Mean QALYs were 0.775 and 0.773 for PCNL and FURS, respectively, generating an adjusted non-significant QALY difference of −0.001 QALYs in favour of PCNL. Therefore, ITT with FURS is on average more costly and does not produce additional QALYs compared with ITT with PCNL.

The results for the 1000 bootstrapped iterations of the regression analysis nested multiple imputation are shown in *Figure 8*. The mean QALY differences between FURS and PCNL are presented in the horizontal axis and mean cost differences in the vertical axis. For only 6.5% of the 1000 iterations PCNL is more costly than FURS (negative cost difference in *Figure 8*). Besides, 51% of the iterations result in FURS producing more QALYs than PCNL. The red dashed line in *Figure 8* represents the £20,000 cost-effectiveness threshold. Iterations sitting to the right and below the red dashed line indicate that ITT with FURS is cost-effective and iterations to the left and above indicate that ITT with PCNL is cost-effective. Only 13% of the result are below the threshold line meaning that FURS is unlikely to be considered cost-effective compared with PCNL. The effect of increasing the threshold above £20,000, that is, moving the red dashed line clockwise with centre in the zero point, is shown in *Figure 9*.

The CEACs showing the probability of ITT with FURS or PCNL being cost-effective at alternative cost-effectiveness threshold values are reported in *Figure 9*. PCNL is less costly for the great majority of the simulations in *Figure 8* and given the small differences in QALYs between the groups, the CEACs in *Figure 9* show PCNL as highly likely to be cost-effective at the usual cost-effectiveness threshold values used for decision-making in the UK NHS (e.g. probability

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 23 Base-case analysis results – multiple imputation: RCT2

Intervention	Mean total cost (£)	Incremental cost (£)[a]	Mean total QALYs	Incremental QALY[a]	ICER (£/QALY)
PCNL	4565		0.775		
FURS	5298	733 (−508; 1973)	0.773	−0.001 (−0.044; 0.042)	−692,724[b]

a Differences adjusted by baseline EQ-5D score. (95% CI)
b FURS is more costly and produces fewer QALYs and therefore is dominated by PCNL.

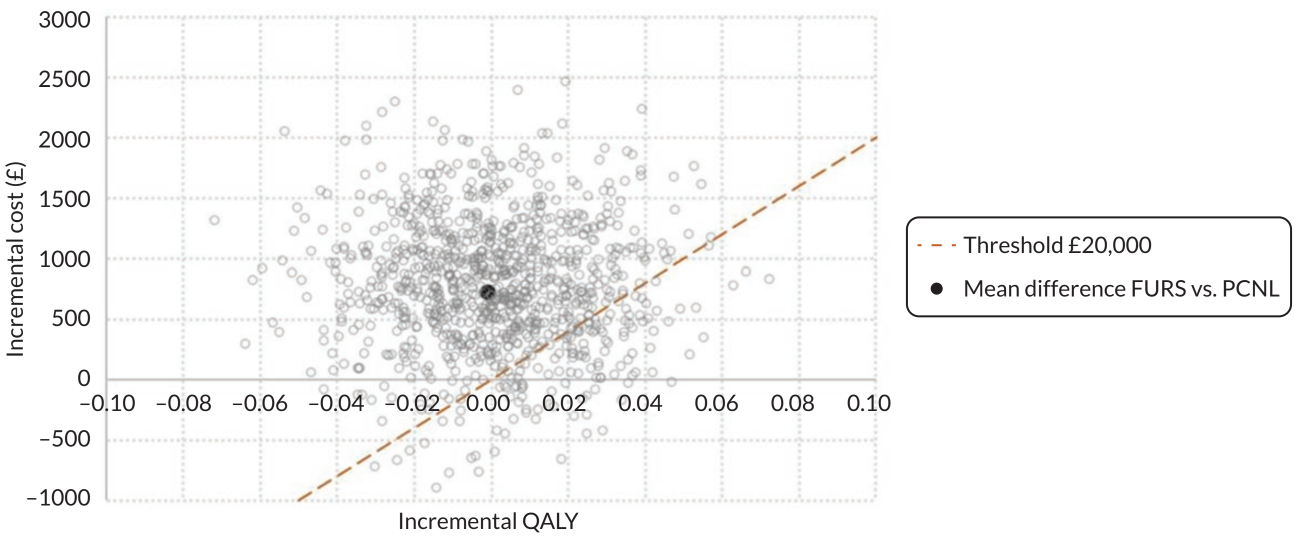


FIGURE 8 Trial-based incremental cost-effectiveness scatterplot for FURS vs. PCNL (base case; imputed data – 1000 bootstrap iterations). Note: Iterations below the red dashed £20,000 threshold line means FURS is cost-effective and above the line that PCNL is cost-effective.

of 0.87 at £20,000 per QALY gained).[67] The probability of PCNL being cost-effective reduces as the cost-effectiveness threshold increases. However, only 51% of the iterations result in FURS producing more QALYs than PCNL; therefore, even at very high threshold values, the probability of FURS being cost-effective cannot be higher than 51%.

Alternative costing method

Sensitivity analysis was conducted using the HRG codes to allocate costs for the index and additional interventions instead of the micro-costing approach. The result for this analysis is reported in *Table 24*. The QALY mean differences remain unchanged with respect to the base-case analysis as this analysis only varied the calculation of costs. ITT with PCNL mean total cost was £6703 and with FURS was £5769. These mean costs produce an adjusted non-significant cost difference of £934 in favour of FURS (PCNL more costly). These cost changes reflect in a downward shift of the simulated iterations in *Figure 9* cost-effectiveness scatterplot. This is presented in *Appendix 9*, *Figure 32*. The great majority of the iterations have crossed to the south and right of the £20,000 threshold represented by the red dotted line, meaning that FURS is highly likely to be considered cost-effective. This is expected: given the very small difference in QALYs the probability of being cost-effective will follow the distribution of the difference in costs. The CEACs for this analysis are reported in *Appendix 9*, *Figure 33*. ITT with FURs has over 80% probability of being cost-effective at the usual threshold values used in the UK NHS for decision making (e.g. up to £30,000 per QALY gained).[65]

Indirect costs

Results on productivity costs (time off work) for RCT2 are reported in *Table 79* in *Appendix 9*. These show that participants took an average of 2.4 and 7.6 days off work in the PCNL and FURS groups respectively, for the 12 weeks after the initial intervention, and 7 and 4 additional days, respectively, for the period up to 12 months post randomisation. The total mean productivity cost associated to time off work were £1200 (SD: 6593) and £2146 (SD:

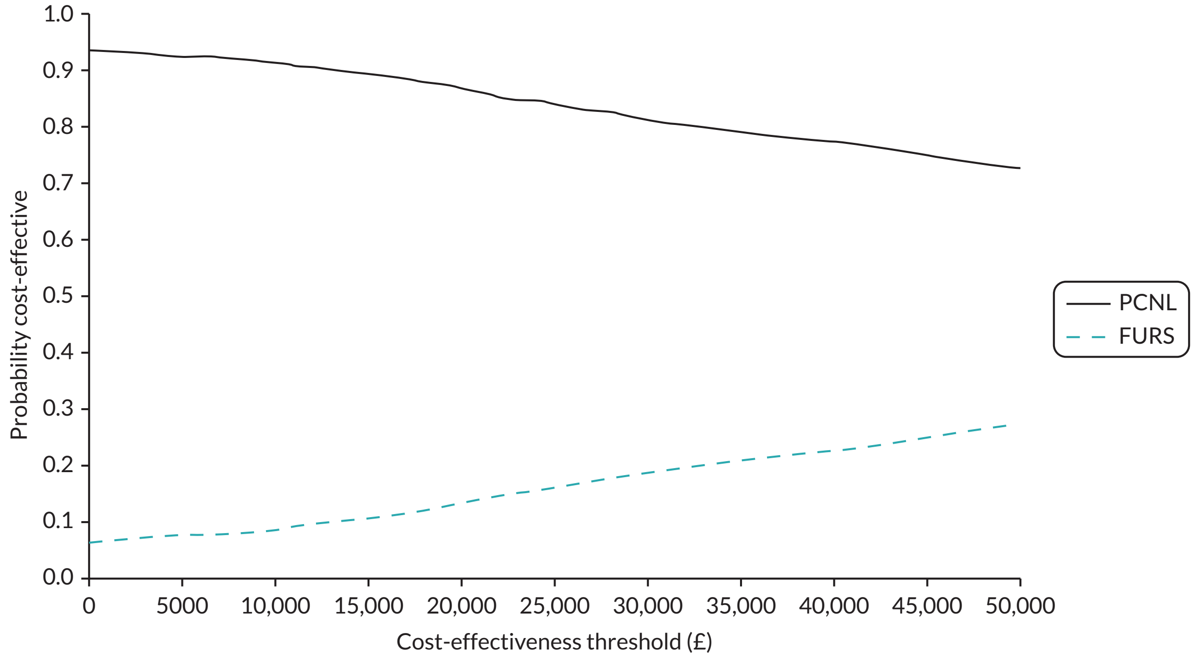


FIGURE 9 Cost-effectiveness acceptability curves for FURS and PCNL groups (trial-based analysis; base case; imputed data – 1000 bootstrap iterations).

TABLE 24 Trial-based incremental cost-effectiveness analysis by alternative costing methods: RCT2

Intervention	Mean total cost (£)	Incremental cost (£)[a]	Mean total QALYs	Incremental QALY[a]	ICER (£/QALY)[a]	Probability cost-effective £13,000 (%)	£20,000 (%)	£30,000 (%)
Base-case analysis (HRG-based costing)								
FURS	5769		0.773			92.3	89.8	86
PCNL	6703	934 (–714; 2582)	0.775	0.001 (–0.042; 0.044)	883,375	7.7	10.2	14

a Differences adjusted by baseline EQ-5D score. (95% CI).

6499) for PCNL and FURS groups, respectively. *Table 79* also reports the out-of-pocket expenditure. Only 9 individuals from 69 answering the question declared buying medications over the counter in the PCNL group and 15 of 60 in the FURS group, with the mean expenditure in medications declared being £0.60 and £5.30 for PCNL and FURS groups, respectively. No participant declared buying private health care in the PCNL group and 2 in the FURS group, resulting in average costs of £1.80. Finally, *Table 79* presents out-of-pocket travelling costs for hospital admission, outpatient consultation, and GP visits. The mean costs of travel are all below £10, with no meaningful differences between the groups on visual inspection of the standard errors.

Economic modelling extrapolation

The statistical analysis in *Chapters 3* and *4* showed differences between the trial groups in the number of participants with residual fragments 12 weeks after the initial treatment and 12 months after randomisation (see *Tables 6* and *13*). The presence and size of residual fragments are associated with higher rates of recurrence,[67] potentially leading to utility and resource use differences beyond the trial follow-up period. The aim of this section is to report on the modelling extrapolation performed to extend the time horizon for the economic analysis beyond the 12-month trial follow-up to 5 years.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Methods

Decision analytic model

A cohort Markov model was developed using TreeAge Pro software (TreeAge Software, Inc., Williamstown, MA, USA). The structure of the model was informed by discussions within the trial management group and clinical collaborators (*Figure 10*). A similar model structure was used for RCT1 and RCT2 and all the compared treatment strategies (i.e. ITT with ESWL, FURS and PCNL). Individuals referred to secondary care and with confirmed lower pole kidney stones enter the model in the 'Waiting for index intervention' health state and are assigned to the index intervention: either ESWL, FURS or PCNL. The model recreates the care pathways observed in the PUrE trial on an ITT basis incorporating the waiting times seen in the trial. Individuals undergoing treatment accrue the cost of the initial episode and move to the tunnel state 'Index intervention', where a reduction in QoL associated with the intervention is accounted for. Markov tunnel states are a series of temporary states – each one lasting one cycle – that must be visited in a fixed sequence.[68] Mirroring the PUrE trial definition for stone cleared, on exit from the tunnel Markov state, individuals enter either of 'Stone cleared/no fragments', 'Fragments ≤ 4 mm' or 'Fragments > 4 mm' Markov states. Furthermore, simulated individuals experiencing a recurrence and needing further interventions are moved to the 'Further intervention' Markov tunnel state (lasting three-monthly cycles as for the 'Index intervention' Markov state).

For the Markov process to end, Markov models need at least one absorbing state; that is, a state that cannot be left. This is represented by the 'Death' state in *Figure 10*. The Death state can be entered from any other state based on age-specific mortality rates reported in the UK life tables.[69] The state occupancy and estimated payoffs are updated on a constant monthly Markov cycle.

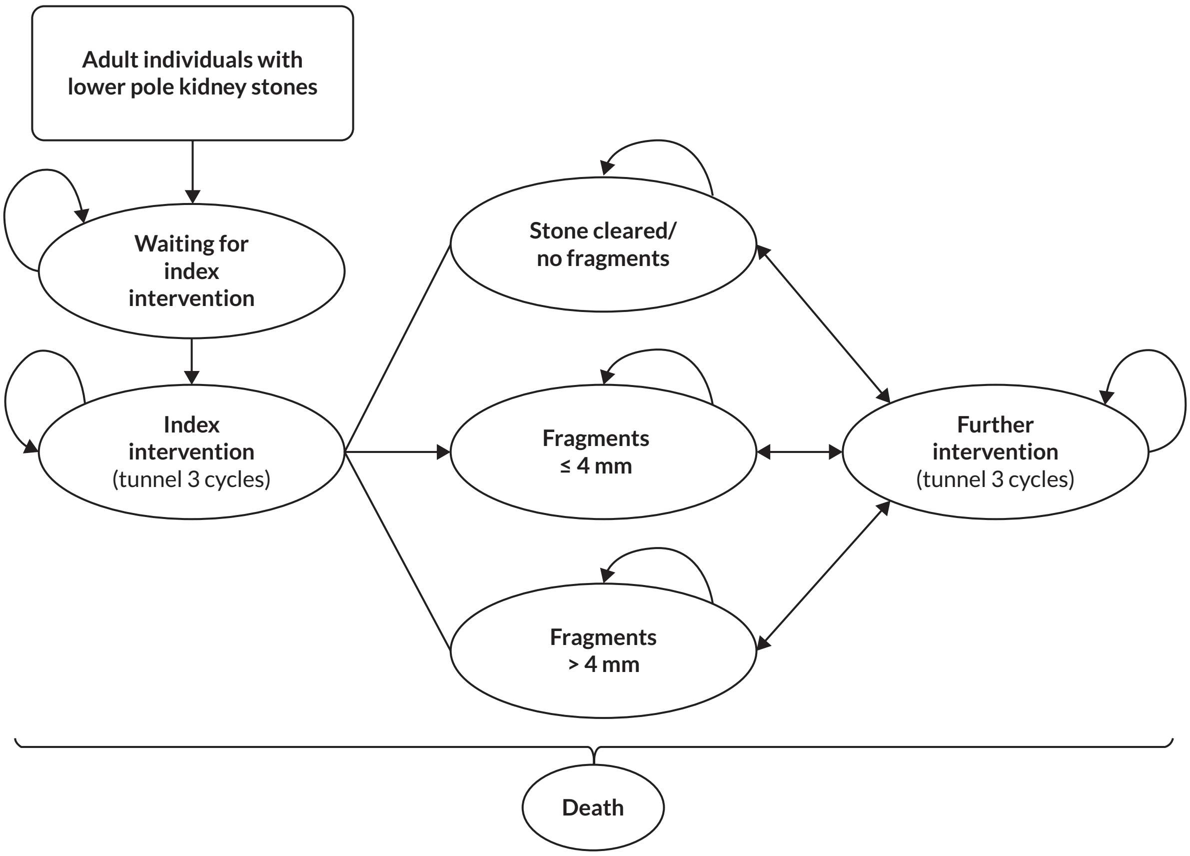


FIGURE 10 Simplified schematic for the PUrE Markov model.

Population

The characteristics of the cohort of individuals entering the model match those of the participants in the PUrE trials. The starting ages for the model were 54 and 59 years for RCT1 and RCT2, respectively, and the proportion of male individuals in the simulated cohort were 62% and 63% for RCT1 and RCT2, respectively.

Time horizon and discounting

A time horizon of 5 years was adopted based on discussions within the trial management group and on the expectation that any events beyond 5 years were likely to be related to a new stone and not to fragments left from the initial intervention. Following economic evaluation method guidelines for the UK, costs and QALYs accrued beyond 12 months were discounted at an annual rate of 3.5%.[65]

Clinical input parameters

While the PUrE trial follow-up was 12 months, data were collected up to 12 weeks post intervention for all trial participants, including those that had their index intervention after 12 months post randomisation. *Figures 11* and *12* show Kaplan–Meier curves estimated from the PUrE data for RCT1 and RCT2, respectively. Transition probabilities were estimated as $1 - (S(t_1)/S(t_0))$ straight from these Kaplan–Meier data, where *S* denotes the survival proportion and *ti*, the time (Markov cycle).

On a successful intervention, individuals can become stone-free. The RR for stone clearance at 12 weeks was obtained from the analysis of the PUrE data. Generalised linear models with binomial family and a log link, clustered by recruiting centre, were used to obtain the SFR risk ratios between FURS and ESWL for RCT1 (RR 1.99, 95% CI 1.61 to 0.2.45), and between FURS and PCNL for RCT2 (RR 0.65, 95% CI 0.53 to 0.81). The proportion of individuals with residual fragments of size ≤ 4 mm and above 4 mm, given the stone was not cleared, were obtained straight from the PUrE data (*Chapters 3* and *4*, *Tables 6* and *13*, and *Appendix 2*). Normal distributions were attached to risk ratios and beta distributions to the proportion of residual fragments of size ≤ 4 and > 4 mm (see *Appendix 9*, *Table 80*).

Data to estimate the proportion of individuals needing further treatment due to recurrence were retrieved from the retrospective study by Emmott *et al.* (2018)[70] and based on 263 individuals who received postoperative CTKUB with residual fragments of at least 1 mm. The authors report the Kaplan–Meier survival curves for time to the need

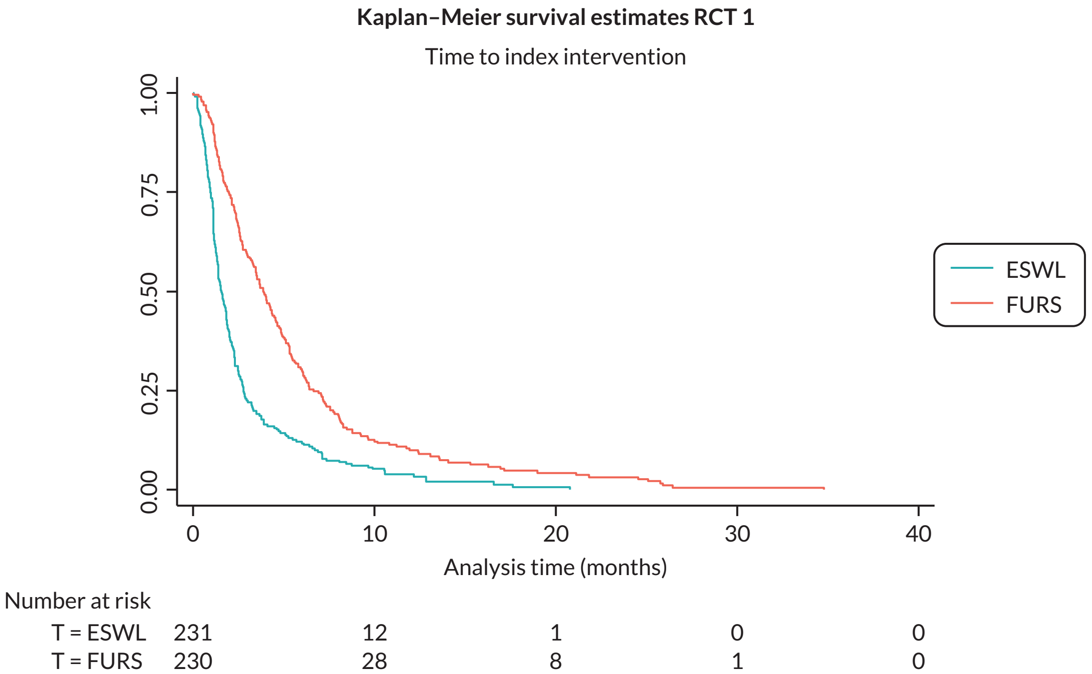
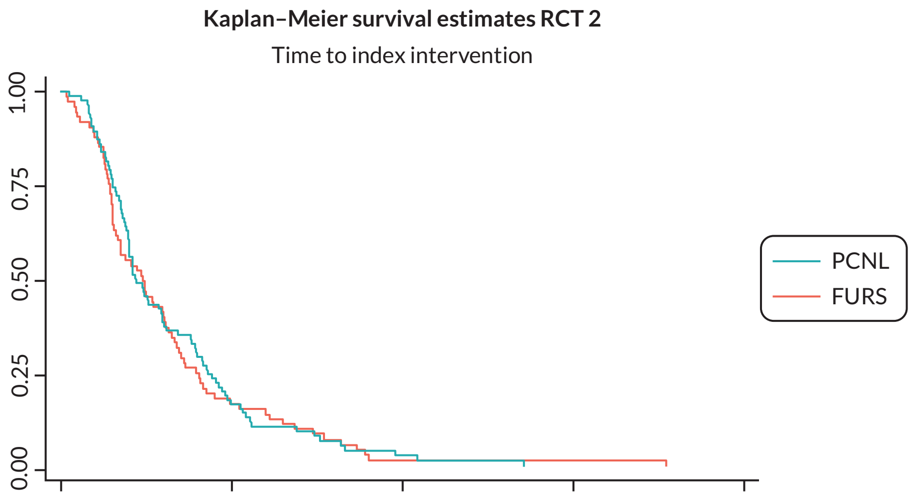


FIGURE 11 Kaplan–Meier plot of time to index operation by treatment allocation for RCT1.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.



Number at risk

	0	10	20	30	40
T = PCNL	86	14	2	0	0
T = FURS	73	12	1	1	0

FIGURE 12 Kaplan–Meier plot of time to index operation by treatment allocation for RCT2.

of PCNL surgery. Data from the original publication were extracted using a web plot digitiser (https://automeris.io/WebPlotDigitizer/) with individual data obtained following methods provided by Wei and Royston (2017) using the Stata command ipdfc.[71] Several parametric survival functions were fitted to the reconstructed data (exponential, Weibull, Gompertz and loglogistic). However, from visual inspection it was clear that none of the parametric adjustments satisfactorily fitted the extracted data. Therefore, a decision was taken to build up the transition probabilities for fragments ≤ 4 and > 4 mm to the 'Further intervention' Markov state, based on the extracted Kaplan–Meier data from Emmott *et al.* (2018)[70] (*Figure 13*) using the same formula used for the time to index intervention. Furthermore, the risk of recurrence for individuals with no fragments (stone cleared) was taken from D'Costa *et al.* (2019)[72] based on the population-based study by Vaughan *et al.* (2019).[73] The authors report a 17% 5-year recurrence rate for individuals that had experienced a first stone episode. This rate was converted to a monthly probability assuming a constant monthly rate.[74]

Health state utilities

PUrE EQ-5D-5L trial data were analysed to estimate health state utilities for the model (*Table 25*). Pooled baseline EQ-5D mean scores for each trial were used as utility weights for the 'Waiting for index intervention' Markov state. EQ-5D score mean and standard error by week were obtained for ESWL (RCT1) and PCNL (RCT2) groups. These results were combined with weekly treatment effects for FURS. Separate regression models were estimated for RCT1 and RCT2. Mixed-effects regression models for repeat measures with adjustment for baseline EQ-5D score, and including centre as a random effect, were defined to estimate the EQ-5D utility score difference between trial groups by data collection time point. Average utility and treatment effects for weeks 1–4, 5–8 and 8–12 were calculated and used as utility weights for tunnel cycles 1, 2 and 3, respectively. In addition, EQ-5D score data at 12 months post randomisation were classified according to the stone being completely cleared, and residuals fragments ≤ 4 and > 4 mm. Using all data pooled for RCT1 and 2, a mixed-effects regression analysis adjusted by baseline EQ-5D score was estimated with EQ-5D score at 12 months as dependent variable and stone clearance as categorical independent variable. This showed no difference in utility score between those with the stone completely cleared and those with reported fragments ≤ 4 mm. Therefore, the classification was dichotomised into completely clear/residual fragments ≤ 4 mm and residual fragments > 4 mm, and the regression model re-estimated. The results of this analysis were used to estimate the utility score difference between stone cleared/residual fragments ≤ 4 mm, and residual fragments > 4 mm Markov states (see *Table 25*).

Health service resource use and costs

Cost for ESWL, FURS and PCNL were obtained from the National Schedule of NHS Costs for year 2020–1 (*Table 26*). Unit costs for day cases were assumed for ESWL and FURS for HRG codes LB36Z and LB65E, respectively. Unit cost for

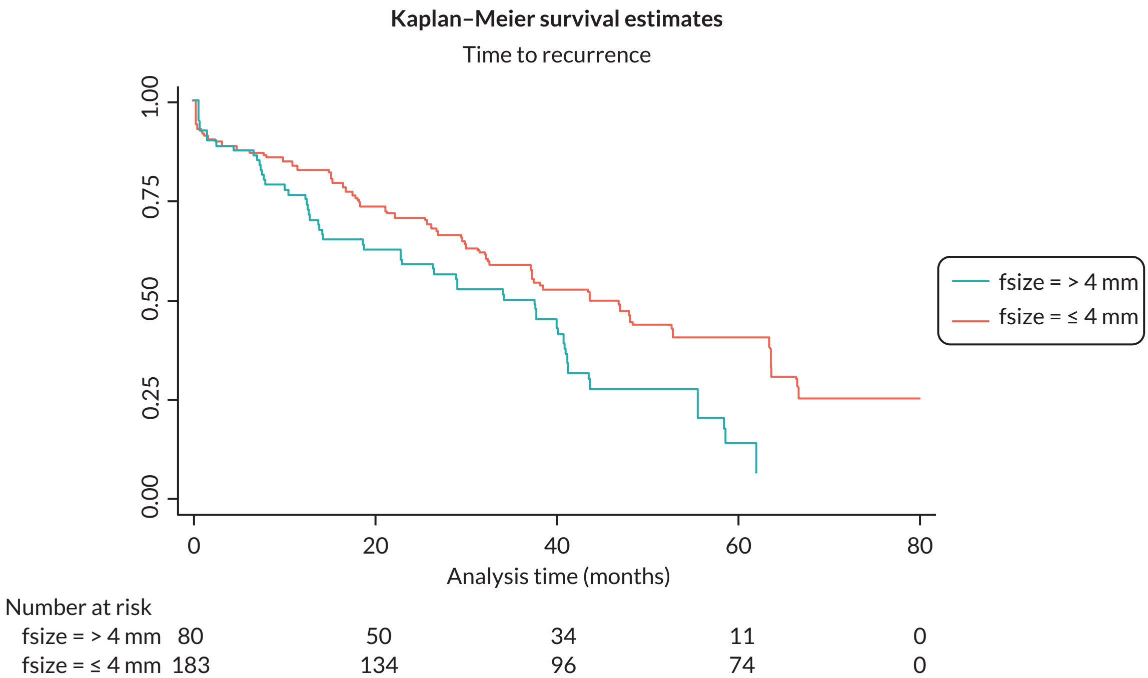


FIGURE 13 Kaplan–Meier plot of time to recurrence by size of residual fragment.

TABLE 25 EuroQol-5 Dimensions health state utility scores applied in the economic model[a]

	RCT1		RCT2	
	ESWL	FURS	PCNL	FURS
Markov state	Mean (SE)	Mean[b] (SE)	Mean (SE)	Mean[b] (SE)
Waiting for index intervention	***0.793 (0.01)***	***0.793 (0.01)***	***0.767 (0.022)***	***0.767 (0.022)***
Index/further intervention (average across relevant weeks)				
Week 1	0.789 (0.016)	−0.106 (0.023)	0.678 (0.029)	−0.114 (0.036)
Week 2	0.814 (0.016)	−0.008 (0.02)	0.763 (0.026)	−0.092 (0.04)
Week 3	0.816 (0.018)	0.028 (0.022)	0.818 (0.024)	−0.092 (0.032)
Week 4	0.819 (0.018)	0.034 (0.021)	0.839 (0.025)	−0.099 (0.04)
Week 5	0.84 (0.016)	0.026 (0.02)	0.878 (0.022)	−0.085 (0.039)
Week 6	0.839 (0.016)	0.045 (0.019)	0.836 (0.031)	−0.029 (0.032)
Week 7	0.831 (0.018)	0.05 (0.015)	0.807 (0.038)	−0.026 (0.038)
Week 8	0.837 (0.018)	0.057 (0.02)	0.843 (0.03)	−0.041 (0.044)
Week 9	0.826 (0.019)	0.06 (0.019)	0.872 (0.028)	−0.038 (0.039)
Week 10	0.83 (0.018)	0.052 (0.02)	0.864 (0.032)	−0.029 (0.037)
Week 11	0.84 (0.019)	0.047 (0.021)	0.853 (0.034)	−0.027 (0.038)
Week 12	0.829 (0.019)	0.05 (0.016)	0.851 (0.03)	−0.008 (0.04)
Stone cleared/no fragments	***0.965 (0.015)***			
Residual fragments				
≤ 4 mm	0.965 (0.015)			
> 4 mm	0.916 (0.021)			

SE, standard error.
a Source: analysis of PUrE trial data.
b Differences with respect to ESWL and PCNL for RCT1 and RCT2, respectively.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 26 Unit costs for ESWL, FURS and PCNL applied in the economic model

Intervention	Unit cost (£) mean (SE)	Source
ESWL	894 (89)	HRG LB36Z; extracorporeal lithotripsy – day case National Schedule of NHS Costs Year: 2020–1[62]
FURS	3350 (335)	HRG LB65E; major endoscopic, kidney or ureter procedures, 19 years and over, with CC score 0–2 – Day case. National Schedule of NHS Costs Year: 2020–1[62]
PCNL	8285 (829)	HRG LB75B; PCNL with CC score 0–1 – Elective inpatient. National Schedule of NHS Costs Year: 2020–1[62]

elective impatient using HRG code LB75B was assumed for PCNL. Gamma distributions were attached to all unit costs to explore parameter uncertainty by conducting probabilistic sensitivity analysis (PSA). Costs for the care pathways for the first year were obtained directly from the RCT within trial analysis; therefore, the HRG unit costs were applied for index and further interventions occurring after the first 12-monthly cycles.

Finally, no differences in the number of outpatient or GP contacts were observed by the study group in the within-trial analysis for PUrE, therefore, a simplifying assumption was made and no cost for these contacts were incorporated in the model.

Model verification

To assess the internal validity of the Markov model, model verification was conducted throughout the model implementation stage following the list proposed by Büyükkaramikli *et al.* (2019).[75] For example, model formulae were verified using external software [Microsoft Excel® (Microsoft Corporation, Redmond, WA, USA)]. The final model results were checked when varying the model input parameter values (e.g. defining all utility scores equal to 1 and 0 discount rate to check whether total QALYs equal total life years) to assess the consistency of the model performance given specific variation in the model input values (black-box testing). The validity of the projected estimates of surgical events was assessed by plotting the model Markov traces with the Kaplan–Maier curves for the PUrE trial with visual inspection showing Markov traces followed closely the Kaplan–Meier curves.

Model analysis

The Markov cohort model was run deterministically using expected costs and QALYs. Parameter uncertainty was characterised by running second-order Monte Carlo simulations using the probability distributions attached to the model mean parameter values. Beta distributions were defined for probabilities (risks) and utilities, gamma for costs and normal distributions for treatment effect for FURS. Details of the probability distributions are reported in *Appendix 9*, *Table 80*. PSA were conducted by running 10,000 random draws from the allocated distributions, producing 10,000 estimates of incremental costs and QALYs. Result tables also show the probability of ESWL, FURS, or PCNL being cost-effective at £13,000, £20,000 and £30,000 per QALY thresholds.[65] CEACs and cost-effectiveness scatterplots are reported in *Appendix 9*, *Figures 34* to *37*.

The base-case analysis for RCT1 and RCT2 assumed that individuals needing further treatment for recurrence after 12 months were treated with the intervention they were initially allocated to (e.g. ESWL and FURS for participants allocated to ESWL and FURS, respectively for RCT1). In addition, different utility weights were assumed for individuals with residual fragments above 4 mm (0.916) and those with residual fragments smaller or equal to 4 mm (0.965). Therefore, the difference in utility after 12 months results from the additional treatments received due to recurrence, the waiting time to receive the initial intervention, and the alternative utility weights according to residual fragment size after treatment.

The time horizon for the base-case analysis was 5 years. The analysis was conducted from the NHS perspective with cost and QALYs beyond 1 year discounted at a 3.5% rate.[65]

The within trial analysis for RCT2 using a micro-costing approach showed PCNL to be less costly and more effective than FURS. It is expected that a higher proportion of individuals treated initially with FURS will need further interventions compared with PCNL. It is also expected that the additional treatment costs and the reduction in QoL will worsen the cost-effectiveness of FURS versus PCNL. For this reason, the base case for the extrapolation beyond the trial follow-up was performed with the HRG costing that resulted in FURS being cost-effective at 12 months post randomisation.

Sensitivity analysis

Key sensitivity analyses were performed:

- assuming alternative time horizons from 12 months to 5 years;
- one-way sensitivity analysis (threshold analysis) varying the RR of stone free/stone clear rate;
- a scenario analysis incorporating equal utility weights for all individuals after treatment regardless of stone being cleared or not (i.e. utility weight of 0.965 for stone cleared, and fragments below or above 4 mm).

Results

RCT1

Base-case analysis

The results for the base-case cost–utility analysis are reported in *Table 27*. Understandably, the expected costs and QALYs are higher for the 5-year time horizon compared to the 12-month follow-up for the within-trial analysis. These higher costs are explained by the proportion of individuals needing treatment due to recurrence in the longer time period considered. The difference in costs and QALYs increase compared to the within-trial base-case analysis (see *Table 19*). The cost and QALY joint effect generate a decrease in the ICER below £30,000 but above £20,000 cost-effectiveness threshold values, indicating that FURS is still not cost-effective when the longer time horizon of 5 years is considered. The PSA shows FURS is unlikely to be considered cost-effective (26% at £20,000 per QALY threshold) unless a higher cost-effectiveness threshold is considered (i.e. 55% at £30,000). The cost-effectiveness scatterplot (see *Appendix 9*, *Figure 34*) shows FURS produces higher expected QALYs for 99% of the PSA iterations but is also more costly in all cases. Finally, the CEACs (see *Appendix 9*, *Figure 35*) show FURS has over 80% chance of being cost-effective at £50,000 per QALY threshold.

Alternative time horizon

The cost–utility analysis results for alternative time horizons applied to the base case are reported in *Table 28*. Cost and QALY increase the longer the time horizon. However, the ICER decreases the longer the time horizon assumed, and the probability of FURS being cost-effective rises. This is explained by the higher proportions of individuals requiring further treatment in the ESWL group. While the ICER between FURS and ESWL decreases the longer the time horizon considered, the fall is not enough to cross the £20,000 cost-effectiveness threshold.[76]

Scenario analyses

Same utility weight for individuals with residual fragments below or above 4 mm

The results for the scenario analysis are also reported in *Table 28*. Assuming equal utility for individuals with residual fragments sized > 4 mm increases QALYs in both ESWL and FURS model strategies. However, the increase in expected QALYs is proportionally higher for ESWL. This is explained by the higher proportion of individuals remaining in the 'Fragments > 4 mm' Markov state for ESWL. Therefore, the ICER increases well above the cost-effectiveness threshold used in the UK (e.g. £20,000 per QALY gained). Also, the probability of FURS being cost-effective further reduces to zero at £20,000 cost per QALY threshold.

Health Care Resource Group-based costing for randomised controlled trial initial interventions

Costing the initial intervention in PUrE with HRG-based unit costs increases the difference in cost between FURS and ESWL maintaining the base-case QALY difference unaltered. This results in a higher ICER just above £30,000 and a lower probability of FURS being cost-effective (just 17% at £20,000 per QALY threshold) (see *Table 28*).

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 27 Model cost–utility analysis results. Base case. RCT1: FURS vs. ESWL

Intervention	Mean total cost (£)	Δ Cost (£)	Mean total QALYs	Δ QALY	ICER (£/QALY)	Probability cost-effective £13,000	£20,000	£30,000
ESWL	3143		4.121			0.94	0.74	0.45
FURS	4811	1667	4.184	0.063	26,571	0.06	0.26	0.55

TABLE 28 One-way sensitivity analysis on model time horizon. RCT1: FURS vs. ESWL

Intervention	Mean total cost (£)	Δ Cost (£)	Mean total QALYs	Δ QALY	ICER	Probability cost-effective £13,000	£20,000	£30,000
Time horizon								
12 months (within-RCT analysis)								
ESWL	2223		0.787			1.00	0.99	0.91
FURS	3362	1138	0.804	0.017	65,163	0.00	0.01	0.09
24 months								
ESWL	2500		1.635			1.00	0.99	0.88
FURS	3889	1389	1.661	0.026	52,924	0.00	0.01	0.12
36 months								
ESWL	2739		2.492			1.00	0.95	0.75
FURS	4288	1549	2.531	0.039	40,150	0.00	0.05	0.25
48 months								
ESWL	2952		3.321			0.98	0.84	0.56
FURS	4567	1615	3.372	0.051	31,580	0.02	0.16	0.44
60 months								
ESWL	3143		4.121			0.94	0.74	0.45
FURS	4811	1667	4.184	0.063	26,571	0.06	0.26	0.55
No utility difference between those with RF size below or above 4 mm								
ESWL	3143		4.185			1.00	1.00	0.97
FURS	4811	1667	4.207	0.022	75,284	0.00	0.00	0.03
HRG-based costing for RCT initial interventions								
ESWL	3447		4.121			0.97	0.83	0.55
FURS	5348	1901	4.184	0.063	30,295	0.03	0.17	0.45

Relative risk of stone-free rate between flexible and extracorporeal shockwave lithotripsy

The results for the one-way sensitivity analysis where the SFR RR at 12 weeks after treatment was varied are reported in *Figure 14*. The stone-free RR is presented in the vertical axis while the ICER for FURS against ESWL is shown in the horizontal axis. ICERs above the cost-effectiveness threshold (e.g. £20,000) mean that, on average, FURS is not cost-effective. Conversely, ICERs below the cost-effectiveness threshold mean that FURS is, on average, cost-effective. In *Figure 14*, this is illustrated with a dashed vertical line representing the £20,000 per QALY threshold. The analysis shows that a stone-free RR above 2.2 is needed for FURS to be cost-effective.

RCT2

Base-case analysis

The results for the base-case analysis using HRG-based costing are reported in *Table 29*. Expected costs and QALYs are higher than the mean cost and QALYs for the within-trial analysis due to the extrapolation from 12 months to 5 years. The difference in costs of £1488 is higher compared to the cost difference for the within-trial analysis. While the proportion of individuals needing further interventions is lower for PCNL compared with FURS, the expected cost of the PCNL strategy rises due to the difference in the unit costs between PCNL and FURS. The difference in QALYs is much higher compared with the within-trial analysis (see *Table 23*). This is explained by the higher proportion of individuals moving to the stone-free Markov state in the PCNL group and the relatively lower QoL experienced for those going through FURS compared to those who experience PCNL. The cost and QALY differences result in an ICER of £24,382 just higher than the £20,000 threshold used in the UK.[65] The higher the cost per QALY willing to pay, the higher the probability of PCNL being cost-effective (i.e. 58% at £30,000 threshold – *Table 29*). The incremental cost-effectiveness scatterplot and the CEACs for PCNL and FURS are reported in *Appendix 9*, *Figures 36* and *37*. These show PCNL producing more expected QALYs for 97% of the iterations but at higher cost in 98% of the iterations.

Sensitivity analysis

The results for the one-way sensitivity analysis varying the model time horizon are reported in *Table 30*. The 12-month time horizon correspond to the within-trial analysis result when the HRG-based costing was used to cost the index interventions. The statistically significant difference in cost together with the lack of significant difference in QALYs resulted in an ICER of £883,375, well above the usual cost-effectiveness thresholds for the UK NHS. This means that at 12 months, ITT with FURS is less costly with no evidence of a substantial impact, on average, in QALYs. Cost and QALYs increase the longer the time horizon considered. However, the difference in QALY increased proportionally more, reducing the ICER below the £30,000 threshold when a time horizon of 48 months is considered.

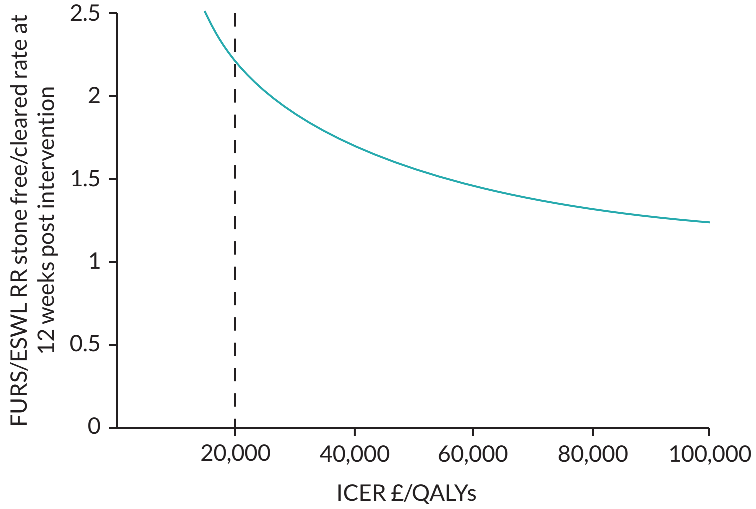


FIGURE 14 One-way sensitivity analysis varying RR of stone-clear between FURS and ESWL (> 1 FURS better).

TABLE 29 Model cost–utility analysis results. Base case. RCT2: FURS vs. PCNL

Intervention	Mean total cost (£)	Δ Cost (£)	Mean total QALYs	Δ QALY	ICER (£/QALY)	Probability cost-effective £13,000	Probability cost-effective £20,000	Probability cost-effective £30,000
FURS	8184		4.089			0.80	0.62	0.42
PCNL	9672	1488	4.150	0.061	24,382	0.20	0.38	0.58

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 30 One-way sensitivity analysis on model time horizon. RCT2: FURS vs. PCNL

Intervention	Mean total cost (£)	Δ Cost (£)	Mean total QALYs	ΔQALY	ICER (£/QALY)	£13,000	£20,000	£30,000
Time horizon 12 months								
FURS	5769		0.773			0.92	0.90	0.86
PCNL	6703	934	0.775	0.001	883,375	0.08	0.10	0.14
Time horizon 24 months								
FURS	6717		1.611			0.91	0.85	0.75
PCNL	7906	1189	1.632	0.021	57,298	0.09	0.15	0.25
Time horizon 36 months								
FURS	7253		2.466			0.89	0.79	0.63
PCNL	8613	1360	2.502	0.036	37,345	0.11	0.21	0.37
Time horizon 48 months								
FURS	7766		3.292			0.83	0.68	0.49
PCNL	9171	1405	3.342	0.050	28,179	0.17	0.32	0.51
Time horizon 60 months								
FURS	8184		4.089			0.80	0.62	0.42
PCNL	9672	1488	4.150	0.061	24,382	0.20	0.38	0.58

The results for the one-way sensitivity analysis varying the SFR RR are reported in *Figure 15*. The ICER between PCNL and FURS is presented in the horizontal axis and the SFR RR in the vertical axis (< 1 PCNL better). The £20,000 cost-effectiveness threshold is also shown and is represented by the vertical black dashed line. The vertical line divides cases where PCNL (left of the dashed line) or FURS (right of the dashed line) are considered cost-effective. The SFR RR has to be at least 0.6 for FURS to be considered cost-effective. For comparison, the RR obtained from PUrE data and used in the base-case analysis was 0.65 (95% CI 0.53 to 0.81). The analysis using micro-costing to cost the initial procedures is presented in *Appendix 9*, *Table 81* and *Figure 38*. This shows a RR over 0.95 is needed for FURS to be considered cost-effective.

The result for the scenario analysis assuming no difference in QoL between individuals with residual fragments above 4 mm is reported in *Table 31*. Expected costs do not vary with respect to those in the base case (see *Table 29*). However, the difference in utility reduces to a third of the difference seen in the base case (from 0.061 to 0.021). This increases the ICER over the usual cost-effectiveness threshold used in the NHS for decision-making.[76]

Summary and discussion

The within-trial cost utility analysis reported in this chapter indicates that ITT with FURS for kidney stones ≤ 10 mm in maximum dimension was more costly over a 12-month follow-up period than ESWL (the mean difference was £1138). This cost difference was driven by the high relative cost of the initial FURS intervention. While the mean costs of additional and subsequent hospital admissions were higher for the ESWL group, these higher costs did not overturn the initial difference in cost in favour of ESWL. A non-statistically significant QALY difference in of 0.017 QALYs in favour of FURS was observed for the 12 months after randomisation. This result is consistent with the main statistical analyses reported in *Chapter 3* where non-statistically significant differences were observed for the AUC at 12 weeks post treatment for all but one analyses. The significant cost difference coupled with a relatively small non-significant QALY difference favouring FURS (FURS produced marginally more QALYs) resulted in an ICER of £65,163 per additional QALY.

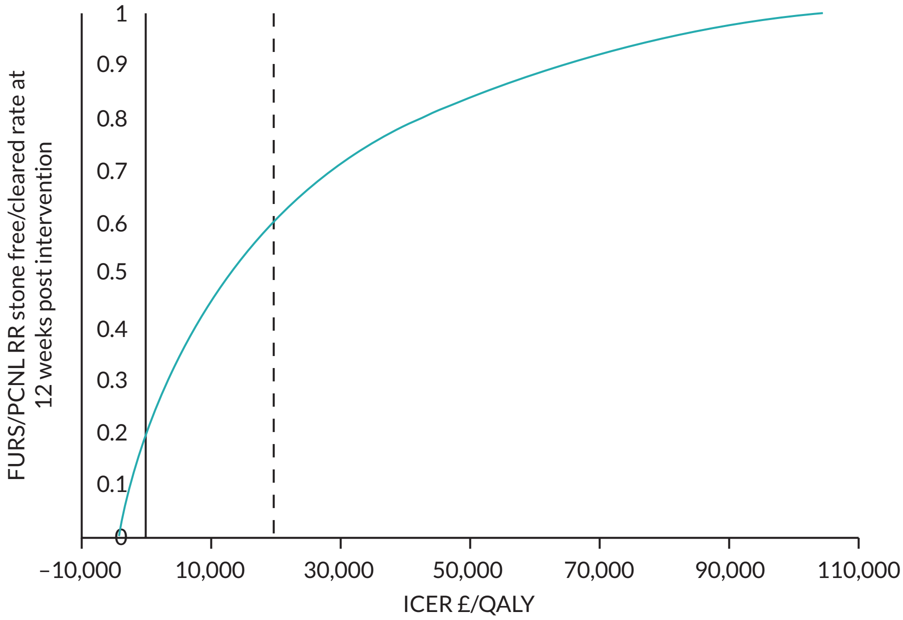


FIGURE 15 One-way sensitivity analysis varying RR of stone-clear between FURS and PCNL (< 1 PCNL better).

TABLE 31 Model cost–utility analysis results. Same utility weight for RF size smaller or bigger than 4 mm. RCT2: FURS vs. PCNL

Intervention	Mean total cost (£)	Δ Cost (£)	Mean total QALYs	Δ QALY	ICER (£/QALY)	Probability cost-effective £13,000	£20,000	£30,000
FURS	8184		4.140			0.94	0.90	0.81
PCNL	9672	1488	4.162	0.021	70,374	0.06	0.10	0.19

This is well above the usual cost-effectiveness threshold used in the UK NHS.[76] That is, moving from an ITT with ESWL to ITT with FURS could produce, on average, more QALYs but at a very high cost per QALY gained. The probabilistic analysis showed a high probability for ESWL to be cost-effective (i.e. over 99% at a £20,000 cost-effectiveness threshold). These results were robust to the sensitivity analysis using HRG-based costing instead of micro-costing.

Intention to treat with FURS is more costly and produces fewer QALYs than ITT with PCNL to treat stones size > 10 and ≤ 25 mm in maximum dimension, but these differences are not statistically significant: the mean difference was £733 (95% CI: −£508 to £1973) for costs and −0.001 QALY (95% CI: −0.044 to 0.042 QALYs). While PCNL showed higher costs associated with the initial procedure, the cost due to hospital readmissions and subsequent procedures were higher for FURS, resulting in an overall higher average cost for FURS. Therefore, mean cost difference against FURS was driven by the similar cost for the index procedures together with the higher cost of subsequent health care for the FURS group. This, together with the QALY difference means that, on average, FURS was more costly and produced fewer QALYs than PCNL, having a 98% chance of being cost-effective at a £20,000 cost-effectiveness threshold.

The sensitivity analysis using HRG-based costing for the initial interventions radically changed the base-case results. This analysis shows a mean cost difference of £934 (95% CI: −£2582 to £714) in favour of FURS for the 12-month follow-up. This is explained by the cost associated with the HRG codes for FURS (LB65) and PCNL (LB75) that give weighted average costs of £4259 and £8114; a difference of £3855 against PCNL that is not levelled by the higher costs of subsequent health care in the FURS trial group. It is open for discussion which of the two costing approaches (micro-costing or HRG-based) reflect better the valuation of the resource use difference between FURS and PCNL. As the QALY difference between the trial groups is very small and non-significant, the cost-effectiveness results mirror the

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

distribution of the difference in cost with the resulting ICER rising to £883,375. This means that ITT with FURS would save, on average, £883,375 per QALY forgone. This result, as well as the base-case analysis result, should be taken with caution given the non-significant and very small difference in QALYs.

We extended the 12-month trial follow-up analysis up to 5 years using a Markov model following best practice. The model reflects possible care pathways after the initial intervention and the potential need for further intervention due to recurrence. For stones ≤ 10 mm in maximum dimension, the model results showed that ITT with FURS generates more QALYs than ESWL but is also more costly, with incremental cost per QALY (£26,571) higher than the usual threshold used in the NHS and a low probability of being considered cost-effective (26% at £20K base case). These results were robust to a number of sensitivity analyses such as alternative time horizon, RR of SFR, and equal utility weights between those with residual fragments less than or equal to, and greater than 4 mm.

For kidney stones size > 10 and ≤ 25 mm in maximum dimension (RCT2), the cost-effectiveness of FURS was dependent on the method used to estimate costs, with PCNL and FURS being cost-effective when micro-costing or HRG-based costing were used, respectively. The extrapolation exercise showed PCNL is still cost-effective when micro-costing was used. Moreover, the results using HRG-based costing show PCNL cost-effectiveness improving for longer time horizons. However, while at 5 years PCNL generates more QALYs, it does this at an additional cost that is too high to be considered cost-effective. Finally, our analysis suggests the RR stone-free should be above 0.6 for FURS to be cost-effective when HRG-based costing was used (and above 0.95 when micro-costing was used).

Strengths and limitations

The main strength of the analysis reported in this chapter is related to the RCT design. RCTs provide the opportunity to collect data prospectively for individual participants and groups that are *a priori* comparable. The randomised treatment allocation increases the probability of unbiased cost and QALY calculations, and consequently a fairer estimate of the difference in cost and QALYs between the treatment groups. The PUrE study comprised two multicentre pragmatic trials facilitating the generalisability of the trial results to those individuals routinely treated in the UK NHS for kidney stones in the lower pole of the kidney. A further strength is that the economic analysis was conducted following a prespecified Health Economic Analysis Plan that followed established economic evaluation methods.[77]

An important limitation of the within-trial cost–utility analysis is the 12-month follow-up. One year might be a short period of time to allow for all relevant cost and consequences of the treatments under consideration. At 12 months post randomisation, a number of participants were still waiting for their index intervention and acceptable stone clearance rates differed between groups (see *Tables 6* and *13*). These issues could potentially trigger further differences in healthcare resource use and/or QoL between the study groups and provide a rationale for the extrapolation exercise. We used PUrE time to event data after 12 months to inform the proportion of the cohort needing their first surgical intervention, SFRs and RRs for FURS, and the proportion of the cohort with residual fragments of size ≤ 4 or > 4 mm after 12 weeks.

Our modelling extrapolation should inform the impact of using longer time horizons for the economic analysis.

As with any decision model, our modelling extrapolation is also subject to limitations. First, the data for the need for further treatment due to residual fragments (≤ 4 and > 4 mm dimension) were obtained from the Canadian study by Emmott *et al.* 2018, a retrospective hospital record review where the existence of residual fragments was assessed on day 1 after PCNL. These data might not accurately reflect the NHS care pathway after 12 months of the index intervention. However, the authors' risk of further intervention (Kaplan–Meier curves) was applied on exit of the tunnel Markov state and disregarding the first 90 days of the Kaplan–Meier data. Arguably, this should better reflect the need for further intervention and less the differing standard clinical practice in Canada and the UK. Secondly, a lower utility weight was assumed for those individuals with residual fragments larger than 4 mm based on the analysis of the PUrE data. While there is some evidence of lower QoL due to residual fragments above 4 mm dimension,[78] the evidence is mixed.[78] Also, the drop in QoL might come as a result of subsequent interventions with the risk of partially double counting the reduction in QoL in our model. We conducted sensitivity analysis assuming equal utility weights between those with fragments smaller and larger than 4 mm dimension. The base case results are robust to this change: for RCT1 FURS was even less likely to be cost-effective and for RCT2 cost-effectiveness is still dependent on the costing method used (i.e. FURS and PCNL being cost-effective when HRG or micro-costing methods are used, respectively).

Chapter 6 Discussion

Parts of this chapter are reproduced with permission from Wiseman *et al.*[1] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, even for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The text below includes additions and formatting changes to the original text.

When patients present with lower pole renal stones, if active treatment is pursued there are three possible technologies available within the NHS which might be used: ESWL, FURS with laser lithotripsy, and PCNL. The choice of treatment can be guided by stone size, likely stone composition, the anatomy of the drainage system of the affected kidney, clinician and patient preference, and availability of equipment and expertise.[12]

Guidance issued by the EAU[12] at the time of study design, in 2014, which was widely followed in UK clinical practice, recommended ESWL or flexible ureteroscopy (URS) as first line options for LPSs ≤ 10 mm. For stones > 20 mm, PCNL was recommended as the first-line treatment option. For stones between these two sizes, in the lower pole, the recommendation depended on whether there were unfavourable factors for ESWL, previously detailed. If unfavourable factors were present, then PCNL or URS were the first-line recommended options, with ESWL being second line. If unfavourable factors were not present then ESWL or PCNL or URS were all possible first-line treatment options.[12] The EAU Guideline also commented that there remained considerable uncertainty regarding the management of LPSs, with each treatment option having advantages and disadvantages.

A Cochrane review and meta-analysis in 2014 of RCTs compared ESWL with either FURS or PCNL for the treatment of renal stones[13] included three studies looking at LPSs and concluded that included trials were small and of low methodological quality. Another review (Donaldson, 2015)[31] included trials comparing PCNL with FURS and identified four further studies. Looking at the seven trials as a whole, the GRADE certainty of evidence ratings for the outcome of SFRs indicated they were of 'moderate' quality. The included trials reported few data on patient outcomes (such as QoL) or on resource use, and none on cost-effectiveness. The PUrE trial therefore set out to compare the three treatment options for lower pole renal stones, separated into two (RCT1 and RCT2) pragmatic, multicentre, patient-randomised, open-label superiority RCTs. RCT1 compared FURS with ESWL, as first treatments options in the pathway, recruiting patients with stones of maximum dimension ≤ 10 mm. RCT2 compared FURS with PCNL, as first treatments options in the pathway, recruiting patients with stones of maximum dimension > 10 and ≤ 25 mm. It aimed to provide robust evidence on health status, QoL, clinical outcomes, and resource use to both the NHS and society to close the gap in evidence. The guidance published by NICE in 2019[79] recommended that for stones anywhere in the kidney under 10 mm, ESWL should be offered, with FURS as an option only in specific situations. For stones between 10 and 20 mm FURS or ESWL could be considered, with PCNL offered only if URS or ESWL have failed. For stones over 20 mm, again not specifically LPSs, they recommended that PCNL should be offered, with FURS to be considered if PCNL is not an option.

Both RCT1 and RCT2 used a pragmatic outcome for stone clearance defined as 'Complete clearance' if the target stone had been treated such that no further action (including no need for observation of it) was required; 'Acceptable clearance' was chosen if observation was needed for a residual fragment but no further intervention was deemed necessary; and 'Unacceptable clearance' was recorded as the outcome for a residual stone that was deemed to require further intervention.

The remainder of the discussion chapter is split into a section on RCT1 then a section on RCT2, except for some general strengths and limitations which are considered for the whole PUrE study.

RCT1

Primary outcome

In the PURE study, RCT1 showed a small benefit in favour of FURS to 12 weeks post treatment compared to ESWL of 0.024 (−0.004, 0.053). When the model was adjusted for the differential waiting times, in effect removing the potential

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

impact of waiting times on the PO, the difference increased slightly to 0.04 (0.01, 0.07) in favour of FURS. However, this small improvement in health status did not reach the pre-determined threshold of clinical significance. There was a notable decrease in the weekly participant reported health status in both groups in the initial postoperative period, which might be expected when intervening for a stone that in many patients is relatively asymptomatic. This decrease was greater in the FURS group than the ESWL group and was no longer apparent at 2 weeks after treatment.

Secondary outcomes

At 12 weeks post intervention, RCT1 showed that patients who underwent FURS were significantly more likely to have complete stone clearance compared to ESWL (72% compared to 36%). The FURS group also required significantly fewer additional treatments compared to ESWL (9% compared to 27%). There was no difference in the complications recorded between the two groups.

Self-reported health status from randomisation up to 12 weeks was the same in both groups and the precision of estimate ruled out any benefit in favour of either intervention.

The pain measures recorded postoperatively inversely reflect the pattern of the health status (EQ-5D score) plots, being higher than baseline in both groups, and greater in the FURS group than the ESWL group, with a significant reduction in pain by week 2, and with levels continuing to decrease over the subsequent weeks. The worst level of pain since the intervention was consistently higher in the FURS group at each week.

Cost-effectiveness

This study has shown that the adjusted mean costs per participant were £2223 and £3362 for the ESWL and FURS groups, respectively, resulting in an adjusted cost difference of £1138 and a non-significantly higher adjusted QALY difference of 0.017 for the 12-month follow-up period. The ICER between FURS and ESWL was £65,163 per QALY gained by FURS, which is above the usual cost-effectiveness threshold considered in the UK decision-making (i.e. below £30,000 per QALY gained).[65] We found that at a threshold value of £20,000 per QALY, ESWL has a 99.9% chance of being cost-effective. The 5-year Markov model extrapolation showed that FURS generates more expected QALYs but is also more costly than ESWL with an incremental cost per QALY higher than the usual cost-effectiveness thresholds used for the NHS decision-making (i.e. £20,000 per QALY gained). The extrapolation analysis also showed the probability of ESWL being cost-effective reducing to 74% at £20,000 threshold value.

Comparison of findings with other studies

There are few studies reporting health status and QoL in kidney stone disease. The only directly comparable study is by Pearle *et al.*,[30] which showed that patients undergoing ESWL had a faster time to '100% recovered', as well as quicker return to work, driving and non-strenuous activity, in contrast to the patient-reported outcomes in the study. In their RCT of stones between 6 and 20 mm at any location in the kidney, Bosio *et al.* did not show any difference in patient satisfaction between ESWL and FURS, though it was not possible to break this down by location in the kidney or by stone size.[80] Another study of 70 symptomatic adults with larger LPSs, of 10–20 mm, who underwent FURS or ESWL, showed the average time to return to normal activity and voiding symptoms were significantly higher in the FURS group, and overall satisfaction score was significantly higher in the FURS compared to the ESWL group.[36] Our systematic review (see *Appendix 10*, *Tables 82–84*, *Figures 39–46*) shows that the evidence base on QoL remains inconclusive and of very low certainty due to risk of bias (RoB) and heterogeneity of measurement.

In contrast to PUrE, where there was a superior SFR in the FURS group compared to the ESWL group, the study by Pearle *et al.*, in 2005, did not show any difference in SFRs.[30] The meta-analysis by Donaldson *et al.* included three studies which are directly comparable to the study and showed a marginally superior SFR within 3 months for FURS compared to ESWL for LPSs under 10 mm.[31] The RCT undertaken by Bosio and colleagues showed higher SFRs at 1 and 6 months, but not at 1 year for FURS compared to ESWL, but this was for stones between 6 and 20 mm at any location in the kidney. The SFR was higher for stones 6–10 mm at 1 month for FURS, but not subsequently, and the SFR for only LPSs from 6 to 20 mm was not higher at any time point.[80] Our systematic review includes three studies reporting on SFR at 3 months, when restricting stone size to ≤ 10 mm, that is, the inclusion criteria for RCT1 (see *Appendix 10*, *Figure 41*) and the SFRs were not statistically different between ESWL and FURS, with moderate certainty of evidence – which was downgraded due to high RoB (see *Appendix 10*, *Figure 40*). The present study RCT1 found significantly higher SFRs at 3 months for FURS (72.1%) vs. ESWL (36.4%). The literature is inconsistent when reporting on SFRs, both in terms of imaging used and whether fragments up to a certain size are allowable or clinically significant. The SFRs in our study are likely higher as imaging was not mandated, but need for further treatment or observation of the stone was the determining factor of whether the patient was stone-free. This is a pragmatic and meaningful outcome to patients and clinicians, especially in terms of what happens in a health system when not all patients are followed up with imaging. Additionally, it is accurate and reproducible across studies. However, the use of this definition of stone free could explain the higher SFRs reported in the RCT1.

The present study also showed that the FURS group had a convincingly reduced number of required additional treatments 12 weeks after intervention compared to ESWL [RR: 0.32 (0.18, 0.58); < 0.001]. While some previous comparable studies also showed higher retreatment rates in the ESWL arm, these did not reach statistical significance in contrast to PUrE.[34,81] This is likely due to the size of the present study in comparison. This point was also examined in our systematic review, where a total of seven studies reported on retreatment rates (see *Appendix 10*, *Figure 43*). Problems with outcome definition and heterogeneity make meta-analysis inappropriate.

The complication rate was low and similar in the two arms of the present study, with seven complications in each arm of the study at Clavien 3 or greater. This is in line with previously published studies.[34,81] In our systematic review (see *Appendix 10*, *Figure 44*), it remains uncertain which intervention is more effective in terms of reducing the risk of complications, and the certainty of evidence remains very low. However, the majority of complications for both treatments in RCT1 were minor.

The results from RCT1 support guidance published in 2019 by NICE,[79] which recommended that ESWL should be offered in the first instance for stones < 10 mm, anywhere in the kidney. This is because of the evidence that ESWL offered a better balance of benefits and costs, and that FURS should be considered only if there are contraindications for ESWL, or anatomical reasons or multiple stones, or if a previous course of ESWL has failed.

Comparing the costs of ESWL and FURS in RCT1 to previous studies is challenging, mainly due to the complexity of the procedures and differing equipment used even within a single healthcare system, let alone the differences across healthcare systems outside the UK. These challenges are shown in the meta-analysis by Geraghty.[82] The authors attempted to tabulate measurements of URS compared with ESWL for ureteric stones and stones located anywhere in the kidney in studies from several countries and found that different costs of treatment were included in different studies. It should be noted, however, that URS is ureteroscopy with a rigid instrument, and this is usually undertaken for stones in the ureter, and differs from FURS which uses a flexible instrument, with FURS usually being costlier due to increased number of single-use items used.

On average, the FURS care pathway cost more than the ESWL care pathway in PUrE, and this cost difference was mainly driven by the unit cost of FURS. While the mean costs of additional and subsequent hospital admissions were higher for the ESWL group, these higher costs did not overturn the initial difference in cost in favour of ESWL. Koo *et al.*[83] reported the cost in a UK setting of the SWL and flexible ureteroscopic laser lithotripsy procedure for LPSs under 20 mm. The total mean costs of the interventions were £2602 for FURS and £426 for ESWL, at 2010 costs, which is similar to the costs in our PUrE study for FURS, which were £3362, when adjusting for inflation between 2010 and 2020. However, the costs of ESWL in the present study were significantly greater at £2223, likely due to the fact that they incorporate additional per-protocol treatments as well as any additional costs in the year after randomisation.

In the TISU study, which looked at the treatment of ureteric stones of any size,[84] the cost of ESWL was calculated at £1550, based on the costs of a six-month follow-up after randomisation. The costs of ESWL in PUrE are on average £672 higher, which can likely be accounted for by increased costs related to the longer duration of follow-up, together with an increased chance of requiring additional treatment due to the location of the stone in the lower pole, a position from where it is less likely to clear without intervention.

A study by Zhang *et al.* in 2018[85] undertaken in the Chinese healthcare system showed that the cost of ESWL was $535, and that for FURS was $4657 but for stones which were larger than in the present study, at between 10 and 20 mm. Again, the costs of treatments across healthcare systems are difficult to compare, and the greater size of the stones in this study also makes direct comparison of costs less valid.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

In summary, the current study shows that FURS is more expensive than ESWL, and while the absolute cost of ESWL is higher than that reported in many studies, the relatively long follow-up and inclusion of further interventions in the costs, account for increase in ESWL costs in RCT1.

RCT2

Primary outcome

In the PUrE RCT2 study, patients with larger kidney stones (i.e. from 10 mm up to 25 mm) were randomised to treatment with FURS or PCNL. It was difficult to recruit to this study and the final sample size was well short of the planned size. Despite the reduced sample size, the results are of the same order (mean, SD, and effect sizes) as in RCT1 and the results are powered sufficiently for significance and detect marginal difference. The difference in PO significantly favoured PCNL, the health status being 0.07 (0.02, 0.11) higher over the 12-week post-intervention period. These estimated effect sizes should be considered alongside the original MCID, which was originally deemed to be around 0.07. This suggests that although significant, the difference in health status is marginal.

It should be noted that RCT2 participants did not have a particularly low health status score, even to begin with, meaning large improvements would be difficult to achieve.

Waiting times in RCT2 were longer than anticipated but very similar in both treatment groups with a mean interval of approximately 6 months until their treatment but with large SD indicating a wide range in the distribution of waiting time, adjusting for waiting time gave similar results for the PO.

Secondary outcomes

Both RCT1 and RCT2 used a pragmatic outcome for stone clearance defined as 'Complete clearance' if the target stone had been treated such that no further action (including no need for observation of it) was required; 'Acceptable clearance' was chosen if observation was needed for a residual fragment but no further intervention was deemed necessary; and 'Unacceptable clearance' was recorded as the outcome for a residual stone that was deemed to require further intervention. At 12 weeks post intervention, the FURS care pathway had fewer patients with complete stone clearance than PCNL (47.9% vs. 70.9%) with the overall effect indicating that for the larger stones in this arm of the study PCNL was superior to FURS [OR: 0.27 (0.16, 0.43); $p < 0.001$] for stone clearance. Treatment-related complications occurred in 22% (16/73) of participants randomised to FURS and in 20% (17/86) randomised to PCNL. These were mostly minor (17 Clavien-Dindo grade 1 or 2 accounting for 21% of the FURS arm and 15% for PCNL), 3% of patients in the FURS arm (2/73) and 5% in the PCNL arm (4/86) had a complication with Clavien–Dindo grade 3 or above.

The level of stone-related pain that participants experienced on the day (using a scale of 0–10, with lower scores representing less pain) at baseline was low with a median score of zero (IQI 0–3). Pain was experienced in the previous week by just less than half of the patients (44% overall), and this did not usually require analgesia – the median for days requiring pain medication in the previous was 0 days (IQI 0–7). At the time of intervention, patients still had a low level of pain with a mean of 1.6 for FURS and 1.8 for PCNL preoperatively. The majority of patients had had no pain in the 7 days before treatment (56.2% FURS/55.8% PCNL), and accordingly few patients had required pain relief in the prior 7 days (1.3 requiring treatment for FURS vs. 1.5 days requiring treatment for PCNL). Following treatment, the 'worst pain' since intervention was worse for FURS than PCNL. QoL appears primarily affected by pain: the weekly pain measures mirrored health status for both FURS and PCNL. As is intuitively logical, pain was worst in the early postoperative period for both procedures, with pain was very slightly higher in the first weeks for those allocated to FURS compared to those who underwent PCNL but which then decreases. This is reflected by an initial worsening of QoL as already discussed.

Simple profiles of health status and the EQ-5D-VAS score both indicate slight improvement in both arms from baseline to the intervention date, possibly due to the reassurance that a plan was in place the anticipated treatment would make the patient better. There was a fall during the first week post intervention, particularly for patients in the FURS treatment arm as expected after such surgery. Thereafter, levels improved to a steady state similar to their original baseline level but remained lower for FURS patients than those in PCNL arm, whose steady state was slightly greater than their baseline before intervention.

Additional treatments

For the PUrE study, additional intervention was defined as any procedure required that was not part of the treatment pathways. The RCT2 treatments compared PCNL (which was expected to completely remove stones up to 25 mm in a single procedure) and FURS (which was also expected to be a single procedure although, an additional FURS procedure was accepted as part of the original treatment provided it took place with 6 weeks of the initial FURS procedure). We found that there were fewer unplanned procedures following PCNL, such that the FURS group had a larger risk of undergoing additional treatments compared to PCNL [RR: 3.96 (2.07, 7.57); $p < 0.001$]. Again though, as with the 12-week results, FURS had a lower rate of stone clearance compared to PCNL [OR: 0.24 (0.16, 0.35); $p < 0.001$] and an increased rate of requiring additional treatment and/or had additional treatment [RR: 2.78 (1.69, 4.57); $p < 0.001$].

Patient-reported pain, health status, and quality of life

The secondary outcomes of health, measured by the EQ-5-5L score and the VAS scale, and QoL as the SF-12 scores (physical and mental domains) are also reported *Table 15*. Both EQ-5D measures show similar same patterns with initial detriment seen in FURS followed by improvement after 2–5 weeks statistically comparable albeit a bit lower to the PCNL group. The SF-12 measures do not show any difference by 12 weeks.

Cost-effectiveness

The QoL results using the mean EQ-5D scores were slightly higher for the FURS group than PCNL at baseline and pre intervention, and then reduced for both FURS and PCNL 1 week after treatment, followed by a steady recovery up to week 6. Overall, the week-by-week post-treatment scores were slightly higher for PCNL. At 3 months post treatment, the EQ-5D scores were higher for both groups compared with their scores at baseline and pre treatment before converging back to their baseline values at 12 months. Our analysis showed mean QALYs of 0.775 and 0.773 for PCNL and FURS, respectively, equating to an adjusted mean difference of −0.001 QALYs (95% CI: −0.004 to 0.042).

Since there was a negligible difference in QoL improvement between the two treatments, the cost-effectiveness is driven by the cost of the procedure. The treatment costs for FURS and PCNL were established as 'micro-costings' using the details of the patient journey within the trial, as well as using more generalised HRG costs.

Using detailed micro-costing for the surgical intervention and postoperative treatment, the 12-month mean cost per participant was £5298 for ITT with FURS and £4565 for ITT with PCNL, resulting in an adjusted mean difference that treatment by PCNL was £733 (95% CI: −£508 to £1973) cheaper than treatment with FURS, with an ICER between FURS and PCNL of £692,724. Therefore, ITT with FURS is, on average, more costly and does not produce additional QALYs compared with ITT with PCNL. Further analysis to address uncertainty (e.g. 1000 bootstrapped iterations of the regression analysis nested multiple imputation) showed PCNL is highly likely to be cost-effective at the usual cost-effectiveness threshold values used for decision-making in the UK NHS (e.g. probability of 0.87 at £20,000 per QALY gained) up to 12 months post randomisation. A sensitivity analysis was conducted using HRG costs for both the index and additional interventions (e.g. repeat FURS). For this, the ITT with PCNL mean total cost was £6703 and with FURS was £5769, such that PCNL was now £934 more expensive than FURS. Since the mean difference in QALYs remained unchanged as this analysis only varied the calculation of costs, this change in cost attribution meant that FURS was highly likely to be considered cost-effective (i.e. 90% at £20,000 cost per QALY threshold) if HRG costs are used rather than the detailed micro-costings summarised above.

Since the more expensive intervention between FURS and PCNL were reversed when based on HRG rather than micro-costings, the cost-effectiveness depends on the method used for assessing the cost.

While data on the time in anaesthetic, theatre and recovery room were meticulously collected and related staff, equipment and disposable costs incorporated in the micro-costing, the approach might underestimate the cost of the procedures by failing to include cost categories such as staff being on call when the patient is in the recovery room after surgery. While the micro-costing could have underestimated the cost of, for example, PCNL, it is unclear why the HRG cost associated to PCNL (LB75 – Elective inpatient; £8285) is almost two and a half times the HRG cost allocated

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

to FURS (LB65 – Day case; £3350) when the time to complete the procedures and the length of hospital stay are not substantially different between the two procedures.

Since there was a negligible difference in QoL improvement between the two treatments, the cost-effectiveness is driven by the difference in cost between the trial groups, and this was different for the micro-costing approach (FURS £5298 and PCNL costing £4565) and HRG-based costing (where FURS costs £5769 and PCNL costs £6703).

The discrepancy between the micro-costings measured in the trial and the more general HRG costing is perhaps even more relevant at a hospital level as opposed to the wider system, where both represent a cost. For hospital providers, the micro-costings are an expense, whereas the HRG cost represents income for NHS trusts/hospitals. Considering these differences between the costs of performing PCNL for a hospital, and the relatively higher reimbursement via the HRG, PCNL is the intervention which should be chosen over FURS from a purely economic point of view.

Lessons: HRG for PCNL is 'good value' for trusts (unless the micro-costings have underestimated the cost, for example, with an interventional radiologist time, or more consumable use, for example, laser and/or basket for miniaturised PCNL track). Nevertheless, these data should drive a greater use of PCNL as both the more cost-effective and better-paid intervention.

Five-year extension

Since residual fragments are associated with a higher risk of recurrent symptoms and intervention[67] economic assessment was extrapolated to anticipate additional events over a 5-year time horizon using a cohort Markov model to reflect possible care pathways after the initial intervention and the potential need for further intervention due to recurrence.

The RR for stone clearance was established using PUrE trial participants who had undergone CT imaging at 12 weeks to provide SFRs and patients with residual fragments of size ≤ 4 or > 4 mm after 12 weeks follow-up. The model also incorporated the care pathways observed in the trial including the observed waiting times.

The extrapolation exercise showed the cost-effectiveness of FURS was dependent on the method used to estimate costs with PCNL still being cost-effective when micro-costing was used. Moreover, the results using HRG-based costing show PCNL cost-effectiveness improving for longer time horizons. However, while at 5 years, PCNL generates more QALYs, it does this at an additional cost that is too high to be considered cost-effective. Finally, our analysis suggests the RR SFR should be above 0.6 (FURS generates 60% of the SFR of PCNL at 12 weeks post surgery) for FURS to be cost-effective when HRG-based costing was used and above 0.95 when micro-costing was used.

Comparison of findings with other studies

Our up-to-date meta-analysis described in *Appendix 10*, *Tables 82*, *83* and *85*, *Figures 39–40* and *47–52* showed that the operative time for PCNL was shorter in all five out of the six studies that reported procedure time (see *Table 85* and *Figure 52*), ranging from 6 minutes longer to 38 minutes shorter than FURS. Bozzini *et al.* reported that PCNL showed a non-significant difference in operative time (72 vs. 56 minutes) but significantly longer fluoroscopy exposure time (176 vs. 32 seconds), and a longer hospital stay (3.7 vs. 1.3 days) than FURS.[86] Zeng also found no significant difference in operating time between Super mini PCNL (SMP) and FURS.[87]

However, Jiang showed that the 'all-seeing needle' micro-PCNL patients had a shorter operative time than FURS,[88] while Fayad *et al.* also found that operating time was significantly shorter for mini PCNL (72 minutes) than FURS (110 minutes).[89] Zhang also found that the operating time was shorter in Ultramini PCNL (UMP, 69 minutes) than FURS (93 minutes) although ESWL had the shortest procedure time (46 minutes).[85]

Zhang showed an increasing length of stay (see *Table 85* and *Figure 51*) from 1.1 days for ESWL to 3.2 days for FURS and 5.3 days for Ultra-mini PCNL for LPSs 1–2 cm.[85] Likewise, Yavuz *et al.* showed that hospital stay increased with increasing invasiveness of the intervention from 1 day for FURS, 1.5 days for micro PCNL, 2 days for both ultra-mini and mini PCNL and 3 days for standard PCNL.[90] In contrast, both Fayad *et al.* and Zeng *et al.* reported that the mini PCNLs performed in their respective studies had a comparable hospital stay to FURS.[87,88] Jiang *et al.* also showed no significant difference in length of stay between micro PCNL and FURS, although the LOS was longer than usual for UK practice with more than 3 days in hospital for both procedures.[88]

It is likely that these data reflect depend on national systems including reimbursement and are therefore not a particularly useful marker of early procedure complexity and outcome. Given this, a validated PROM might act as a more useful comparator for future research.

As reported in PUrE RCTs, the complications reported in the literature for FURS and PCNL are usually Clavien–Dindo grades I and II, with no clinically relevant difference between the two treatments. Although eight RCTs reported complications (see *Table 85* and *Figure 50*) and in our revised meta-analysis, there is a lack of standardised reporting which makes these difficult to compare against.

Bozzini *et al.* found an increasing complication rate from 6.7% for ESWL, 14.5 for FURS and 19.3% for PCNL.[86] Whereas Zhang *et al.* noted 6.7% in SWL and 8.3% in FURS and 16.7% in UMP.[85] These were mostly Clavien–Dindo Grade I and II, with FURS having a higher rate of postoperative fever than UMP. Jiang *et al.* found no significant difference between micro PCNL and FURS for LPSs < 20 mm for complications including haemoglobin drop.[88] In keeping with this, blood transfusion is uncommon with treatment of stones of this size and was not required in either supermini-PCNL or FURS, which had a comparable complication profile overall.[87]

The auxiliary procedure rate in the Zeng *et al.* study between SMP and FURS was lower in the SMP group,[87] whereas the auxiliary procedure rate was comparable for ESWL and FURS but lower for PCNL.[86] As stones become larger, the differences are likely to further favour PCNL over FURS, although a smaller track intervention compared with standard PCNL seems reasonable. For example, in patients with stones larger than 20 mm in the lower pole who had mini-PCNL had a significantly lower likelihood of having a postoperative nephrostomy (2.8% vs. 21.1%), a shorter duration of nephrostomy when one was inserted, and a shorter hospital stay than standard PCNL.[91]

The retreatment rate in ESWL was significantly higher at 30% than FURS 5% or UMP 1.6% and Bozzini *et al.* also showed significantly higher retreatment rate in ESWL (61.3%) compared to FURS (4.3%) and PCNL (2.2%).[85,86]

In their study comparing SWL versus FURS versus PCNL for single LPS 1–2 cm, Bozzini found that overall SFR (a composite between a < 3 mm residual fragment and negative urine culture) improved from 61.8% for ESWL to 82.1% for FURS and 87.3% for PCNL.[86] Their exclusion criteria differed from PUrE RCT2 by excluding patients with unfavourable lower pole anatomy (patients with a steep infundibular-pelvic angle < 30°, a longer calyx more than 10 mm, or a narrow infundibulum < 5 mm), which may contribute to their improved results. Furthermore, patients with uric acid stones had additional post-operative urinary alkalization with potassium citrate therapy.[86]

Similarly, Zhang *et al.* noted a significantly better SFR (defined as no residual stone or < 0.3 cm on computed tomography at 3 months postoperatively) from 73% for SWL to 92% for FURS and as high as 98% for UMP following the treatment of a single 10–20 mm radio-opaque lower caliceal stone.[85] Yavuz *et al.* also (see *Figure 47*) reported improved SFRs of 90.1% for ultra-mini PCNL and 94.1% and 94% for mini and standard PNL compared with 76% for FURS and 77% for Micro-PCNL.[90] Micro-PCNL was also shown to have similar outcomes to FURS by Jiang *et al.*, who reported FURS SFR of 84.5% and 79.3% for micro PCNL.[85]

In their prospective, multicentre trial including 160 patients with 10–20 mm lower pole renal calculi Zeng *et al.* evaluated stone-free status defined as residual fragments ≤ 3 mm 1 day postoperatively [on plain kidney ureter bladder (KUB) X-ray and ultrasound] and at 3 months (by CTKUB). The mean stone diameters were similar to those in the PUrE RCT2 trial at 15.0 mm for the SMP patients and 14.3 mm for the FURS patients. SMP achieved a significantly better clearance at day 1 (91.2% vs. 71.2) and at 3 months review (93.8% vs. 82.5%) compared with FURS.[87] In contrast, Fayad *et al.* found stone clearance (defined as ≤ 2 mm on CT at 12 weeks) for lower calyceal stones < 20 mm was not significantly different for mini PCNL (92.7%) and FURS (84.3%).[89]

The up-to-date meta-analysis showed that two studies also demonstrated that FURS was more expensive to perform than PCNL. Zhang found that FURS was more expensive than PCNL (US$ 4660 vs. 4090) which was related to extra

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

consumable/disposable costs with FURS, whereas much of the equipment for PCNL is reusable.[85] Similarly, Yavuz found that the cost of the procedure decreased with increasing invasiveness of treatment, such that FURS was more than twice as costly as ultra-mini, mini, and standard PCNL, with no significant differences between these different PCNL sizes. However, micro-PCNL cost was mid-way between the cost of FURS and the other three percutaneous interventions.[90] The cost analysis included in Zhang *et al.* also found that the total cost of SWL was significantly lower than that of UMP, but that the overall cost of FURS was the highest.[85]

Our original systematic review and meta-analysis for the PUrE trial design showed that QoL was seldom reported in the studies analysed. This review concluded that information regarding QoL should be an area for improvement in knowledge of treatment outcome and was therefore made a PO in the PUrE RCT. In the PUrE study, the AUC from treatment to 12 weeks and found that QoL for PCNL was marginally superior.

Bosio reported a patient-centred assessment of their procedure and found no significant difference in patient perception of the procedures that 88.2% of patients undergoing ESWL and 87.1% of those undergoing FURS were satisfied with the procedure itself, that 80.9% versus 88.6% were satisfied with the results and 86.8% and 77% of ESWL and FURS patients would choose the same procedure again.[80]

In our up-to-date systematic review/meta-analysis prepared for this monograph and included as *Appendix 10*, four RCTs provided some data related to QoL or satisfaction outcomes. In more detail, Bosio *et al.* (2022) reported no clinically or statistically significant difference in patient satisfaction with the procedures[80] that is, similar to the PUrE RCT2 findings. Yavuz *et al.* (2020) found that return to daily activities was significantly faster for FURS (3.9 days) than standard PCNL (13.5 days).[90] Interestingly the recovery time depended on track size, with a significant increase from 4.5 days for Micro PCNL, 6.5 days for ultra-mini PCNL, 9.3 days for Mini PCNL and 13.5 days standard PCNL to return to normal daily activity. Reporting findings that were opposite to our RCT2 data, both Jin and Zeng found that FURS had the advantage over miniaturised PCNL with less postoperative pain.[87,92]

For larger renal stones from 10 to 20 mm, NICE 2019 recommends FURS or SWL as equivalent options with PCNL the next step if FURS or SWL have failed. In renal stones > 20 mm the NICE 2019 guidelines concur with both the EAU and AUA guidelines that PCNL has the strongest evidence to support it.[79] PCNL is therefore the first-line intervention to offer for these larger stones, although FURS should be considered if PCNL is not appropriate.

Strengths of the PUrE study

- RCT1 was a large pragmatic multicentre randomised trial, adequately powered for a difference in health status, making the findings generalisable.
- Although RCT2 did not reach the target sample size, the results indicate sufficient power given the significant marginal differences in health status.
- Both RCTs had a PO which was focused on patient health status and other QoL outcomes, hitherto an under-researched area.
- These were pragmatic trials embedded within current urological practice across the UK; therefore, delivered findings that are relevant to the NHS.
- The randomised treatment allocation, the multicentre pragmatic trial design that facilitates the generalisability of the trial results to individuals routinely treated in the UK NHS, and the prespecified Health Economic Analysis Plan that followed established economic evaluation methods are strengths of the health economic evaluation.
- Provides granular detail on the costs of treatment throughout the patient pathway across both trials.
- Baseline characteristics show that the trial population was broadly similar to that in previously published studies.
- The definition used for SFR was pragmatic and related to additional hospital visits/treatment and was not always imaging-based. These patient-centred outcomes are clear, meaningful, and clinically relevant, based on whether additional treatment was needed ('unacceptable'), observation was needed ('acceptable') or the stone had been sufficiently well treated that no further intervention or observation of any small fragments was deemed necessary ('complete clearance').

Limitations of the PUrE study

- Treatment received by participants could not be blinded; however, this reflects the routine use of these interventions in an NHS setting.
- There may have been heterogeneity of clinical practice across the multiple centres which would affect both RCTs.
- The heterogeneity in clinical practice will also mean that the costs of the individual procedures will vary between units, a factor that will need to be considered when applying the outcomes of the study to decision-making in an individual unit basis.
- The trials were completed during the COVID-19 pandemic; while the majority of participants were recruited and completed follow-up prior to the pandemic, recruitment was curtailed, and the study timelines were extended to accommodate the longer waiting times to treatment.
- The differing waiting times for the different treatments in RCT1 may have affected patient and clinician willingness to take part in the study.
- Response rates to health status and QoL questionnaires became lower as their stone episode became less recent and patients recovered from their treatment.
- Variability in imaging used for follow-up, but again, this reflects how decisions are made in the NHS.
- The cost–utility within trial analysis was for 12 months of follow-up, which might be too short to observe all relevant cost and consequences of the treatments to be recorded and therefore a modelling extrapolation was attempted.

Waiting times

The waiting times to intervention, particularly for FURS and PCNL, were longer than expected and did impact on completion of treatment and outcomes within the 12-month post-randomisation period for a small proportion of participants, compounded by the COVID-19 pandemic for the small proportion who were waiting for treatment or were recruited during the pandemic. We think that the long waiting times for FURS and PCNL made it harder for some sites to capture the intervention date in a timely fashion to trigger collection of the PO, with the PO completion rates being lower for those randomised to FURS in RCT1, and for both FURS and PCNL in RCT2 than for ESWL. The longer time from recruitment to intervention may also have had an impact on participant motivation to engage and complete their study questionnaires. Health status and QoL from the time of the decision to treat is what the patient perceives whereas the doctor's perspective is more likely to be weighted by the apparent patient QoL from the time of treatment itself, it was important to consider both these starting points in our analyses.

Reporting equality, diversity and inclusion

The participant population in both RCT1 and RCT2 were similar in age and sex to the UK population with kidney stones requiring treatment. We optimised participation in the trial through the use of recordings of recruitment appointments in participating sites to feedback to the recruiters' potential conscious/unconscious biases. A 'top tip for recruitment' sheet was generated from this and other such initiatives in similar trials to enable a broad range of participants to be recruited into the trial.

The research team included PPI partners who were actively involved throughout the study contributing to the study design, development of study materials and contributing to discussions of the study results, as well as preparation of the plain English summary. The study team represented a broad range of expertise in quantitative and qualitative methodologies. Less experienced members of the team, including a summer student wanting to gain experience in the clinical trial arena, were encouraged to lead discrete components of the study (under senior supervision) and lead the study SWATs.

Implications for health care

The choice of health status as the PO measure provides important data for patient counselling, assisting joint decision-making. It also allowed for the cost-effectiveness of the treatments for LPSs to be assessed.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Consent/patient expectations and experience

The relatively high baseline health status and low preoperative pain scores for most patients meant that there was relatively little room for improvement following treatment. Indeed, there was a dip in patients' health status in the first week, seemingly driven by postoperative pain, and this is important for preoperative counselling. Furthermore, the health status and QoL scores returned towards baseline levels at 12 months post intervention, rather than showing a substantial improvement. The treatment of largely asymptomatic LPSs should therefore be considered as a preventative intervention for the future rather than in the expectation of an immediate QoL or healthcare benefit.

It is therefore important in preoperative counselling and patient consent to emphasise that, for patients in whom symptoms are slight, the goal of active stone treatment is to avoid renal/ureteric colic from < 10 mm stones treated by ESWL or FURS in RCT1, or to avoid a more extensive FURS or a PCNL if the stones were to enlarge in the future. For those patients with larger stones, such as those in RCT2, the goal of treatment is to prevent stone growth, with the requirement of a more extensive operation in the future, and/or to try and prevent some of the complications of renal stones, such as infections or loss of renal function. The benefits of these interventions may not be readily appreciated by the patient or the healthcare system, therefore.

The consistent drop in health status at week 1 in both trials, particularly for those in the FURS treatment arms, has important implications for practice during consent to treatment. Based on these data, it is important for clinicians to emphasise that patients are likely to feel worse before they feel better. While this is true of most surgical interventions, and therefore likely to be expected by most patients, the PUrE study provides detailed information regarding the anticipated post-treatment trajectory of QoL that will help make patients' consent to be more fully informed.

The results from RCT1 shows that there is no significant difference in health status based on treatment allocation. However, in RCT2, the health status over 12 weeks post intervention was better for the PCNL group than the FURS group. Furthermore, when considering secondary outcomes, patients should be counselled that for smaller LPSs < 10 mm, FURS leads to a higher SFR than ESWL, and for LPSs > 10 mm, PCNL leads to higher SFRs than FURS.

Resource allocation including theatre scheduling and inpatient beds

RCT1 demonstrated that ESWL was the more cost-effective treatment for LPSs < 10 mm. This complements the recommendation from NICE Renal Stone Guidelines 2019 that ESWL should be offered as the first-line intervention for renal stones of this size.[79]

The cost-effectiveness difference between FURS and PCNL is less clear-cut, because of the discrepancy between a micro-costing analysis (which showed PCNL was more cost-effective) or HRG-based analysis (which showed that FURS was more cost-effective).

While data on the time in anaesthetic, theatre and recovery room were meticulously collected and related staff, equipment and disposable costs incorporated in the micro-costing, the approach might have underestimated the cost of the procedures by failing to include cost categories, such as staff being on call when the patient is in the recovery room after surgery or theatre preparation time. While the micro-costing could have underestimated the cost of, for example, PCNL, it is unclear why the HRG cost associated to PCNL (LB75 – Elective inpatient; £8285) is almost two and a half times the HRG cost allocated to FURS (LB65 – Day case; £3350) when the time to complete the procedures and the length of hospital stay are not substantially different between the two procedures.

More transparency in the way the HRG costing are estimated and the reasons behind procedures being associated with corresponding HRG codes and cost would be beneficial for decision-making.

Implications for research

Further research is needed to answer the following research questions:

What is the patient preference in terms of therapy and what level of patient tolerability and acceptance does each treatment modality enjoy?

What effect will suction devices, improvements in laser technology, and intraoperative pressure monitoring have on postoperative pain, QoL, SFRs, complications and costs of FURS?

What is the effect of miniaturisation of PCNL on postoperative pain, length of stay, complications, SFRs and costs?

What is the clinical and cost-effectiveness of full metabolic assessment compared with standard advice alone, in people who have undergone treatment with LPSs?

Conclusion

Taken together, these trials have shown that ESWL, FURS and PCNL (including smaller track varieties) are all safe and effective in the treatment of < 10–25 mm lower pole kidney stones but have various attributes that require the 'pros and cons' to be evaluated for individual patients.

Extracorporeal shockwave lithotripsy is less invasive, and less expensive than FURS or PCNL, but has a lower SFR, and consequently higher retreatment rate than these surgical interventions. Operation time is shorter for PCNL than FURS but is associated with a longer hospital stay and more complications.

The PUrE study shows that ESWL for the treatment of lower renal pole stones under 10 mm was more cost-effective than FURS because it was cheaper and there was no meaningful difference in patient health status when comparing the two. SFRs were higher, however, with FURS. From an overall NHS perspective, the costs savings of treating all patients with these stones with ESWL as their first treatment would be substantial. For larger stones, 10–25 mm, PCNL gave better health status over 12 weeks post intervention and SFRs than FURS and was more cost-effective on a micro-costing basis.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Additional information

CRediT contribution statement

Oliver Wiseman (https://orcid.org/0000-0001-8039-6903) (Consultant Urologist): Conceptualisation, Methodology, Funding acquisition, Investigation, Supervision, Writing – original draft (equal lead), Writing – reviewing and editing.

Daron Smith (https://orcid.org/0000-0001-8314-4795) (Endoluminal Endourologist): Conceptualisation, Methodology, Funding acquisition, Investigation, Supervision, Writing – original draft (equal lead), Writing – reviewing and editing.

Kath Starr (https://orcid.org/0000-0003-3356-7751) (Senior Trial Manager): Conceptualisation, Methodology, Funding acquisition, Investigation, Project administration, Supervision, Writing – reviewing and editing.

Lorna Aucott (https://orcid.org/0000-0001-6277-7972) (Senior Statistician): Methodology, Formal analysis, Visualisation, Writing – original draft (equal lead), Writing – reviewing and editing.

Rodolfo Hernández (https://orcid.org/0000-0003-2619-8230) (Health Economist): Conceptualisation, Methodology, Funding acquisition, Formal analysis, Visualisation, Writing – original draft (equal lead), Writing – reviewing and editing.

Ruth Thomas (https://orcid.org/0000-0002-8316-2616) (Triallist, Research Manager): Methodology, Funding acquisition, Project administration, Writing – original draft (lead), Writing – reviewing and editing.

Steven MacLennan (https://orcid.org/0000-0002-2691-8421) (Senior Research Fellow): Conceptualisation, Methodology, Funding acquisition, Investigation, Writing – original draft (equal lead), Writing – reviewing and editing.

Charles Terry Clark (https://orcid.org/0000-0003-1813-9767) (PPI): Methodology, Funding acquisition, Writing – reviewing and editing, PPI.

Graeme MacLennan (https://orcid.org/0000-0002-1039-5646) (Professor/CHaRT Director/Triallist): Conceptualisation, Methodology, Funding acquisition, Supervision, Validation, Writing – reviewing and editing.

Dawn McRae (https://orcid.org/0009-0009-0553-1453) (Trial Manager): Methodology, Investigation, Project administration, Supervision, Writing – reviewing and editing.

Victoria Bell (https://orcid.org/0000-0001-9518-9296) (Trial Manager): Project administration, Investigation, Writing – reviewing and editing.

Seonaidh Cotton (https://orcid.org/0000-0002-7883-0608) (Senior Trial Manager): Project administration, Supervision, Writing – reviewing and editing.

Zara Gall (https://orcid.org/0000-0002-8494-672X) (Consultant Urologist): Investigation, Supervision, Writing – reviewing and editing.

Ben Turney (https://orcid.org/0000-0003-0885-6404) (Consultant Urologist): Concept, Funding acquisition, Investigation, Writing – reviewing and editing.

Samuel McClinton (https://orcid.org/0000-0002-0539-9570) (Professor/Consultant Urologist): Concept, Methodology, Funding acquisition (lead), Investigation, Supervision, Writing – reviewing and editing.

Other contributions:
Charles Terry Clark in addition has been involved as a PPI team member since conception of the trial and oversaw the conduct of the trial and the interpretation of results and the writing/editing of the report.

Acknowledgements

The Health Services Research Unit (HSRU) and the Health Economics Research Unit (HERU) are core funded by the Chief Scientist Office of the Scottish Government Health and Social Care Directorates. The views and opinions expressed by authors in this publication are those of the authors. The authors accept full responsibility for this publication.

The authors thank all the patients who participated in PUrE.

A special thanks goes to the trial managers and data coordinators who helped support the study: Katie Banister, Sarah Cameron, Karen Campbell, Tracey Davidson, Alison McDonald, Jess Wood (Trial Managers), Zoe Batham, Rebecca Bruce, Dianne Dejean, and Sharon Wren (Data Co-ordinators).

We thank CHaRT programming team led by Mark Forrest and other staff within CHaRT/HSRU for their assistance with PUrE: Paul Manson, Information Scientist, Samantha Wileman, Quality Assurance Manager.

Thanks to Mr Rodolfo Agustín Hernández and Mr Andrés G. Hernández for retrieving HRG based unit costs for hospital interventions based on the PUrE data and following NHS costing algorithm (NHS Digital: *Casemix Companion HRG+ 2020/21* and *HRG+ Reference Costs Code to Group*).

We thank the members of the PMG for their ongoing advice and support of the trial, plus the independent members of the TSC and DMC; and the staff at the recruitment sites who facilitated the recruitment, treatment and follow-up of trial participants (all listed below). Finally, we thank the National Institute for Health Research and the Health Technology Assessment Programme for funding the PUrE study.

Independent members of the Trial Steering Committee
Nigel Parr (Chair), Linda Harper, Jeff Hussey, Barbara Farrell, Tim Everitt (PPI).

Independent members of the Data Monitoring Committee
Steve Payne (Chair), Robert Hills, David Douglas.

Project Management Group
Grant Holders: Ken Anson, Terry Clark, Rodolfo Hernández, Thomas Lam, Graeme MacLennan, Steven MacLennan, Samuel McClinton, Robert Pickard, Daron Smith, Kath Starr, David Thomas, Ruth Thomas, Ben Turney, Oliver Wiseman, and the CHaRT Trial Office.

Clinical centres
Thanks to all staff at NHS sites responsible for recruitment in the clinical centres as follows:

Ashford and St Peter's Hospitals NHS Foundation Trust (Ashford and St Peter's):

Sachin Agrawal (PI), Nimalan Arumainayagam, Iain Burton, Maria Croft, Victoria Frost, Keshnie Govender, Catherine Gray, Stephanie Ivie, Megan McGee. **Belfast Health and Social Care Trust *(Belfast)*: David Connolly (PI),** Jay Atkinson, Karen Johnston, Ajay Pahuja, Trevor Thompson, Christina Wilson, Siobhan Woolsey. **Betsi Cadwaladr University Health Board *(Wrexham)*: Iqbal Shergill (PI),** Stacy Ackerley, Issah Ahmed, Mohamed Yehia Ahmed Abdallah, Annette Bolger, Sally Dolan, Emma Hall, Rachel Hughes, Tania Morgan, Mary Roberts, Jane Stockport, Heather Thomas, Lewis Waggett, Claire Watkins. **Blackpool Teaching Hospitals NHS Foundation Trust *(Blackpool)*: Chandrasekharan Badrakumar (PI), Jaginkere Chiran (PI),**Denise Bennett, Paul Billington, Gemma Brown, Janet Brown, Shamina Hankinson, Carol Jeffs, Alex

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Maley, Praveen Rao, Ella Riedel, Emma Stoddard, Mohamed Youssef. **Bradford Teaching Hospitals NHS Foundation Trust (*Bradford*): James Forster (PI),** Richard Benton, Kay Cockroft, Hayley Inman, Charlotte Johnson-Smith, Anne Kay, Gail Opio-te, Helen Robertshaw, Jane Sewell, Kelvin Stewart, Sarah Tinker. **Cambridge University Hospitals NHS Foundation Trust (Cambridge): Oliver Wiseman (PI),** Kelly Leonard, Lorraine Starling. **Dartford and Gravesham NHS Trust (Darent Valley):** Seshadri Sriprasad (PI), Jayasimha Abbaraju, Faqar Anjum, Protip Bose, Ryan Coe, Tom D'Angelo, Tracy Edmunds, Anca Gherman, Musaad Hamdoon, Charlotte Kamundi, M Melhem, Subrabrata Muknerjee, Jospeh Nariculum, Naomi Oakley, Olumide Olufuwa, Eleanor Ray, Christine Scott, Carmel Stuart, Sherma Turner, Danielle Whitting. **East Kent Hospitals University NHS Foundation Trust *(Canterbury)*: Jemma Hale (PI), Nitin Shrotri (PI),** Noor Gifford, Sashi Kommu, Angela Moon, Janine Musselwhite, Joshua O'Donnell, Pam Offord, Rachel Vernall. **East Lancashire Hospitals NHS Trust (*Blackburn*): Shalom Srirangam (PI),** Gaynor Bowen, Stephen Duberley, Jillian Fitchett, Karl Worrall. **East Sussex Healthcare NHS Trust *(Eastbourne)*: Simon Mackie (PI),** Anne Cowley, Emma Edmunds, Jenny Law, Kelly Mintrim, Claire Rutherford. **Epsom and St Helier University Hospitals NHS Trust *(Epsom)*: Stephen Gordon (PI),** Rosavic Chicano, Rachel Hayre, Ben Horsburgh, Lisheng Hou, Pareeta Patel, Neringa Vilimiene. **Gloucestershire Hospitals NHS Foundation Trust *(Cheltenham)*: Kim Davenport (PI),** Adnan Ahmad, Susan Beames, John Henderson, Joseph Jelski, Bashir Mohamed (Associate PI), Jonathan Ord, Paula Townsend, Amanda Tyler. **Imperial College Healthcare NHS Trust *(*Charing Cross*)*: Ranan Dasgupta (PI),** Zaynab Al-Saad, Sanela Andrijac, Tamer El-Husseiny, Lynsey Evison, Christine Gan, Milad Hanna, Marya Hussain, Stephanie Ivie, Ella Jameson, Angelina Karatziotou, Mitra Kondjin-Smith, Laura McLeavy, Rhea Strudwick, Thilipan Thaventhiran. **Kingston Hospital NHS Foundation Trust *(Kingston)*: Rashmi Singh (PI),** Isabel Bradley, Jennifer Crooks, Rita Fernandes, Jefferson Lisboa Santos, Tracey O'Brien, Andrew Swain. **Leeds Teaching Hospitals NHS Trust *(Leeds)*: Michael Kimuli (PI),** Irfan Jina, Lorraine Wiseman. **Liverpool University Hospitals NHS Trust** (formerly Royal Liverpool and Broadgreen University Hospitals NHS Trust) **(*Liverpool*): Ciaran Lynch (PI),** Patricia Kelly. **London North West University Healthcare NHS Trust (*Northwick Park*): Erik Havranek (PI),** Temi Adedoyin, Annamaria Harmathova, Asif Raza, Jeff Webster. **Manchester University NHS Foundation Trust (**formerly Central Manchester University Hospitals NHS Foundation Trust) ***(Central Manchester)*: Ben Grey (PI),** Houda Chea, Ann Crump, Laura Derbyshire, Babra Hanif, Natasha Jones, Iain McIntyre, Richard Napier-Hemy, Mohammed Nazir, Anne Palmer, Laura Perry, Angelique Quistin, David Triggs, Ian Venebles. **Manchester University NHS Foundation Trust** (formerly University Hospital of South Manchester NHS Foundation Trust) ***(Wythenshawe)*: Graham Young (PI),** Karyee Chow, Thiruendran Gunendran, Rebecca Ilyas. **Mid and South Essex NHS Foundation Trust** (formerly Mid Essex Hospital Services NHS Trust) **(Broomfield*)*: Daniel Swallow (PI),** Mandy Austin, Tracey Camburn, Lauren Fergery, Fiona McNeela, Sue Smolen, Petros Tsafrakidis. **Mid Yorkshire Hospitals Trust *(Wakefield)*: Abed Ali (PI), Anthony Browning (PI), Stephanie Symons (PI),** Jim Anderson, Hollie Brooke, Paul Hughes, Stephen Littler, Beverley Taylor. **Newcastle Upon Tyne Hospitals NHS Foundation Trust *(Newcastle)*: David Thomas (PI),** Nicola Brown, Alexandra Hall, Paul Hindmarch, Ann Hudson, Bernadette Kilbane, Lynn Langhorne, Peter Murphy, Wendy Robson, Dianne Wake. **NHS Fife (**formerly Fife Health Board) **(Kircaldy*)*: Robyn Webber (PI),** Laura Beveridge, Susan Pick. **NHS Grampian (**formerly Grampian Health Board) **(*Aberdeen*): Ismail El-Mokadem (PI), Samuel McClinton (PI),** Angela Allan, Kim Black, Maeve Dooher, Gayle Hutcheon, Sandra Mann, Deborah Price, Abdul Rauf, Lorna van Lierop, Lynne Walker, Nicholas Cohen, Thomas Lam. **NHS Lothian** (formerly Lothian Health Board) ***(Edinburgh)*: Mark Cutress (PI),** Carol Ann Stephenson, Rachael Campbell, Kareen Darnley, Siatos Dimitrios, James Donaldson, Daniel Good, Sarah KcKim, Katherine Lawrence, Simon Phipps, Hazel Smith, Catherine Somerville, Ben Thomas, Maggie Wishart. **Norfolk and Norwich University Hospitals NHS Foundation Trust (*Norfolk and Norwich*): Stuart Irving (PI),** Neil Burgess, Karen Convery, Patricia Dickens, William Finch, Deirdre Fottrell-Gould, Janette Guymer, Lisa Hudig, Jocelyn Keshet-Price, Martyn Oliver, Gill Pout, Emily Tropman. **North Bristol NHS Trust (Bristol Southmead*)*: Francis Keeley (PI),** Rebecca Cousins, Lyndsey Johnson, Leigh Morrison, Emily Perry, Joe Philip, Constance Shiridzinomwa, Sophie Short, Anthony Timoney. **Northern Care Alliance NHS Foundation Trust** (formerly Salford Royal NHS Foundation Trust) ***(Salford)*: David Ross (PI),** Nicholas Boxall, Euan Green, Victoria Thomas. **Northern Care Alliance NHS Foundation Trust** (formerly Pennine Acute Hospitals NHS Trust) **(Oldham*)*: Andreas Bourdomis (PI),** Dionne Adams, Mariya Barakova, Carolyn Corbett, Susan Dermody, Karen Hallett, Zahid Hussain, Joanne Johnson, Dann Johnstone, Simon Kaye, Susan Kernan, Louise Morby, Lyndsay Scarratt, Raveendra Surange, Sarah Winnard, Lynne Winstanley. **Oxford University Hospitals NHS Foundation Trust *(Oxford)*: Ben Turney (PI),** Emily Grout. **Rotherham NHS Foundation Trust (Rotherham*)*: Bo Parys (PI),** Dawn Collier, Katy Curtis, Kathryn Dixon, Victoria Murray, Rachel Walker. **Royal Berkshire NHS Foundation Trust *(Reading)*: Sunil Kumar (PI),** Caroline Hayden, Johanna Thomas, Karen Wilmott. **Royal Bournemouth and Christchurch Hospitals NHS Foundation Trust *(Bournemouth)*: Joshua Phillips (PI),** Andrew Bates, Norbert Boker, Emma Bromwich, Maria Letts, Tina Stoycheva, Luke Vamplew, Steven Williams. **Royal Cornwall Hospital NHS Trust *(Royal Cornwall)*: Christopher Blake (PI),** Benita

Adams, Nicki Devooght-Johnson, Eve Fletcher, Fiona Hammonds, Emma Lamarti, Mark Mantle, Catherine Miller, Anna Old, Hannah Osborn, Jessica Summers, Sue Webber, Lucy Whitbread, Belinda Wroath, Gabby Young. **Royal Devon University Healthcare NHS Foundation Trust** (formerly Royal Devon and Exeter NHS Foundation Trust) ***(Exeter):*** **Melanie Walton (PI),** Angela Cottrell, Thomas Dutton, Melanie Hutchings, Joseph John, Sinead Kelly, Victoriano Mariano III, Evanna McEvoy, Rachelle Meyer, Madeline Moore, Linda Park, John Pascoe, Pauline Sibley, Michelle Walter. **Salisbury NHS Foundation Trust (*Salisbury)*: James Brewin (PI),** Moyeen Bakar, Sa'id Dabah Aljamal, Sarah Diment, Victoria King, Holly Morgan, Mostafa Ragab, Abby Rand, Ben Wilson. **Sheffield Teaching Hospitals Foundation Trust *(Sheffield)*: Jake Patterson (PI),** Anne Frost, James Hall, Susannah Hulton, Sanjay Rajpal. **South Tees Hospitals NHS Foundation Trust *(Middlesborough)*: Ashok Sakthivel (PI),** Chandrasekharan Badrakumar (Associate PI), Basavaraj Gowda (Associate PI), Alaiyi West (Associate PI), Susan Walker. **South Tyneside and Sunderland NHS Foundation Trust (**formerly City Hospitals Sunderland NHS Foundation Trust) ***(Sunderland)*: David Bryant (PI),** Sue Asterling, Eleanor Dungca, Deborah Kemp, Jenny Koo Ng, Gareth Lewis, Stuart McCracken, Ameet Patel. **Southport and Ormskirk Hospitals NHS Trust (Southport*)*: Rahul Mistry (PI),** Anna Morris. **St George's University Hospitals NHS Foundation Trust *(St. George's)*: Kenneth Anson (PI),** Christina Armoogum, Marco Bolgeri, Elizabeth Cruddas, Deirdre Daly, Cherrelle De Souza, Mark Lynch, Kristoffer Ohlin, Sarah Rounding, Agne Sekmokaite, Rashmi Singh. **St Helens and Knowsley Teaching Hospitals NHS Trust *(Whiston)*: John McCabe (PI),** Karen Chadwick, Sharon Dealing, Susan Dowling, Sandra Greer, Clare Harrop. **Stockport NHS Foundation Trust *(Stockport)*: Zara Gall (PI),** Adebanji Adeyoju, Louise Brown, Patricia Clitheroe, Sarah Connolly, Patricia Coughlan, Janette Curtis, Emma Goodwin, Helen Haydock, Rion Healy, Sheila Hodgkinson, Susan Hopkins, Alissa Kent, Magda Lucia Kujawa, Sean Rezvani, Sarah Smallwood, Lara Smith, Clare Tibke. **University Hospital Southampton NHS Foundation Trust *(Southampton)*: Bhaskar Somani (PI),** Andrew Guy, Mark Harris, Kimberley Harris, Clare Hutchison, Shery Michael, Abbie Morley, McDonald Mupudzi, Amelia Pietropaolo, Winningtom Ruiz, Ana Salgueiro, Mariya Shaji, Sanchia Triggs, Hannah Wardall. **University Hospitals Coventry and Warwickshire NHS Trust *(Coventry)*: Anthony Blacker (PI),** Dannielle Burgess, Davina Hewitt, Samantha Hyndman, Sarah O'Toole, Maria Truslove. **University Hospitals Plymouth NHS Trust** (formerly Plymouth Hospitals NHS Trust) ***(Plymouth):*** **Andrew Dickinson (PI), Ivo Dukic (PI),** Maria Brennan, Lyn Cogley, Sue Freemantle, Salvatore Natale. **University Hospitals Sussex NHS Foundation Trust** (formerly Western Sussex NHS Foundation Trust) ***(Western Sussex)*: Dan Magrill (PI),** Isobel Amey, Yolanda Baird, Irvin Balagosa, Marian Flynn-Batham, Raquel Gomez, Susanna Greenslade, Sally Moore, Marian Nelmes, Matthew Smith, Tan Tsawayo, Paimaun Zakikhani.

University Hospitals Sussex NHS Foundation Trust (formerly Brighton and Sussex University Hospitals NHS Trust) ***(Brighton)*: Andrew Symes (PI),** Paul Frattaroli, Allison Leslie, Lorraine Shah-Goodwin, Mel Smith, Kate Trivedi. **Wirral University Teaching Hospital NHS Foundation Trust *(Wirral)*: Snehal Patel (PI),** Mandy Edwards, Richard Glendinning, Keith Morris, Andrea Young.

Sites opened but never recruited

NHS Highland (formerly Highland Health Board) ***(Inverness)*: Sara Ramsey (PI),** Fiona Barrett, Kathleen MacLeod, Debbie McDonald. **University Hospitals of Leicester NHS Trust *(Leicester)*: Masood Khan (PI),** Jane Douglas-Moore, Mandy Le-Butt, Shumikazi Mzazi.

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 PUrE study received favourable ethics opinion from North of Scotland Research Ethics Committee on the 12 November 2015 (REC reference number 15/NS/0113).

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Information governance statement

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

Disclosure of interests

Full disclosure of interests: Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at https://doi.org/10.3310/WFRE6844.

Primary conflicts of interest: Daron Smith, Kath Starr, Lorna Aucott, Rodolfo Hernández, Ruth Thomas, Steven MacLennan, Graeme MacLennan, Seonaidh Cotton and Samuel McClinton report: NIHR HTA funding award for the PUrE trial 13/152/02 during the conduct of the study. Oliver Wiseman reports: grants from the EMS and Coloplast institutions; consulting fees from Boston Scientific, EMS, Coloplast and Ambu organisations; and payments from Boston Scientific, EMS, Coloplast and Devicare, during the conduct of the study. Daron Smith reports: grant for stent design, Award ID: NIHR202935 (CASSETTE), invoices in progress for Olympus KeyMed for lecture/course delivery for SpR Education and for plenary lecture at BAUS during 2022; payments received for Webinar support from Cook and Coloplast organisations; honorary board membership of the Endourology Academy; and honorary committee membership for World of Endourology Society, during the conduct of this study. Lorna Aucott reports: membership of PHR Research Funding Board between 13 June 2017 and 12 June 2023 and COVID-19 Reviewing between 1 June 2020 and 30 September 2020 during the conduct of this study. Zara Gall reports: payments from Boston Scientific for speaking at Symposium on Ureteroscopy at BAUS in June 2022 and for organisation and delivery of PCNL course in April 2023; and committee membership of BAUS Endourology February 2000–February 2023. Ben Turney reports: multiple research grants to support PhD students and research projects from Boston Scientific, research grant to investigate the genetics of BPH from the Urology foundation and joint industry/ESPRC funding to support mathematics and engineering students. BT also reports consulting fees from Boston Scientific; payments or honoraria from Boston Scientific and from Devicare; participation on International Advisory Board for Boston Scientific; leadership/advisory role for Wellbeck Oxford; stock with Ambulatory Surgery International; and receipt of research project equipment from Boston Scientific. Patents reported by Ben Turney are: New Invention Disclosure #21-D0431(Multi-class segmentation in ureteroscopy and laser lithotripsy) (filed September 2022), New Invention Disclosure: #22-D0371 (Automated Estimation of Stone Size in Laser Lithotripsy) (filed September 2022), and New Invention Disclosure: #17-D0422 (Visualization through Non–Circular Sheath Via Optimized Fluid Dynamic in Channel Flow) during the conduct of this study.

Publications

Wiseman O, Smith D, Starr K, Aucott L, Hernández R, Thomas R, *et al*. The PUrE randomised controlled trial 1: Clinical and cost effectiveness of flexible ureterorenoscopy and extracorporeal shockwave lithotripsy for lower pole stones of ≤10 mm. *Eur Urol* 2025;**88**(2):179–89. https://doi.org/10.1016/j.eururo.2025.02.002

References

1. Wiseman O, Smith D, Starr K, Aucott L, Hernández R, Thomas R, *et al*. The PUrE randomised controlled trial 1: Clinical and cost effectiveness of flexible ureterorenoscopy and extracorporeal shockwave lithotripsy for lower pole stones of ≤10 mm. *Eur Urol* 2025;**88**(2):179–89. https://doi.org/10.1016/j.eururo.2025.02.002
2. Neisius A, Preminger GM. Stones in 2012: epidemiology, prevention and redefining therapeutic standards. *Nat Rev Urol* 2013;**10**:75–7. https://doi.org/10.1038/nrurol.2012.253
3. Jour I, Lam A, Turney B. Urological stone disease: a 5-year update of stone management using hospital episode statistics. *BJU Int* 2022;**130**:364–9. https://doi.org/10.1111/bju.15728
4. Pearle MS, Calhoun EA, Curhan GC; Urologic Diseases of America Project. Urologic diseases in America project: urolithiasis. *J Urol* 2005;**173**:848–57. https://doi.org/10.1097/01.ju.0000152082.14384.d7
5. Burgher A, Beman M, Holtzman JL, Monga M. Progression of nephrolithiasis: long-term outcomes with observation of asymptomatic calculi. *J Endourol* 2004;**18**:534–9. https://doi.org/10.1089/end.2004.18.534
6. Glowacki LS, Beecroft ML, Cook RJ, Pahl D, Churchill DN. The natural history of asymptomatic urolithiasis. *J Urol* 1992;**147**:319–21. https://doi.org/10.1016/s0022-5347(17)37225-7
7. Osman MM, Alfano Y, Kamp S, Haecker A, Alken P, Michel MS, Knoll T. 5-year-follow-up of patients with clinically insignificant residual fragments after extracorporeal shockwave lithotripsy. *Eur Urol* 2005;**47**:860–4. https://doi.org/10.1016/j.eururo.2005.01.005
8. Bihl G, Meyers A. Recurrent renal stone disease-advances in pathogenesis and clinical management. *Lancet* 2001;**358**:651–6. https://doi.org/10.1016/S0140-6736(01)05782-8
9. Turney BW, Reynard JM, Noble JG, Keoghane SR. Trends in urological stone disease. *BJU Int* 2012;**109**:1082–7. https://doi.org/10.1111/j.1464-410X.2011.10495.x
10. National Statistics. *Hospital Episode Statistics, Admitted Patient Care, England: 2013–14: Procedures and Interventions*. NHS Digital; 2015. URL: https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity/hospital-episode-statistics-admitted-patient-care-england-2013-14 (accessed 17 May 2023).
11. Cass AS, Grine WB, Jenkins JM, Jordan WR, Mobley TB, Myers DA. The incidence of lower-pole nephrolithiasis – increasing or not? *Br J Urol* 1998;**82**:12–5. https://doi.org/10.1046/j.1464-410x.1998.00684.x
12. Türk C, Petřík A, Sarica K, Seitz C, Skolarikos A, Straub M, *et al.* EAU guidelines on interventional treatment for urolithiasis. *Eur Urol* 2016;**69**:475–482.
13. Srisubat A, Potisat S, Lojanapiwat B, Setthawong V, Laopaiboon M. Extracorporeal shock wave lithotripsy (ESWL) versus percutaneous nephrolithotomy (PCNL) or retrograde intrarenal surgery (RIRS) for kidney stones. *Cochrane Database Syst Rev* 2014;**11**:1465–858. https://doi.org/10.1002/14651858.CD007044.pub3
14. Tiselius HG, Ackermann D, Alken P, Buck C, Conort P, Gallucci M; Working Party on Lithiasis, European Association of Urology. Guidelines on urolithiasis. *Eur Urol* 2001;**40**:362–71. https://doi.org/10.1159/000049803
15. Albala DM, Assimos DG, Clayman RV, Denstedt JD, Grasso M, Gutierrez-Aceves J, *et al.* Lower pole I: a prospective randomized trial of extracorporeal shock wave lithotripsy and percutaneous nephrostolithotomy for lower pole nephrolithiasis-initial results. *J Urol* 2001;**166**:2072–80. https://doi.org/10.1016/s0022-5347(05)65508-5
16. Candau C, Saussine C, Lang H, Roy C, Faure F, Jacqmin D. Natural history of residual renal stone fragments after ESWL. *Eur Urol* 2000;**37**:18–22. https://doi.org/10.1159/000020093
17. Cass AS. Extracorporeal shockwave lithotripsy or percutaneous nephrolithotomy for lower pole nephrolithiasis? *J Endourol* 1996;**10**:17–20. https://doi.org/10.1089/end.1996.10.17

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

18. Havel D, Saussine C, Fath C, Lang H, Faure F, Jacqmin D. Single stones of the lower pole of the kidney. Comparative results of extracorporeal shock wave lithotripsy and percutaneous nephrolithotomy. *Eur Urol* 1998;**33**:396–400. https://doi.org/10.1159/000019623
19. Pearle MS, Lotan Y. Urinary Lithiasis: Etiology, Epidemiology, and Pathogenesis. In Campbell MF, Wein AJ, Kavoussi LR, editors. *Campbell-Walsh Urology*. Philadelphia, PA: W. B. Saunders; 2007.
20. Auge BK, Dahm P, Wu NZ, Preminger GM. Ureteroscopic management of lower-pole renal calculi: technique of calculus displacement. *J Endourol* 2001;**15**:835–8. https://doi.org/10.1089/089277901753205852
21. Chung SY, Chon CH, Ng CS, Fuchs GJ. Simultaneous bilateral retrograde intrarenal surgery for stone disease in patients with significant comorbidities. *J Endourol* 2006;**20**:761–5. https://doi.org/10.1089/end.2006.20.761
22. Grasso M, Ficazzola M. Retrograde ureteropyeloscopy for lower pole caliceal calculi. *J Urol* 1999;**162**:1904–8. https://doi.org/10.1016/S0022-5347(05)68065-2
23. Holland R, Margel D, Livne PM, Lask DM, Lifshitz DA. Retrograde intrarenal surgery as second-line therapy yields a lower success rate. *J Endourol* 2006;**20**:556–9. https://doi.org/10.1089/end.2006.20.556
24. Kourambas J, Delvecchio FC, Munver R, Preminger GM. Nitinol stone retrieval-assisted ureteroscopic management of lower pole renal calculi. *Urology* 2000;**56**:935–9. https://doi.org/10.1016/s0090-4295(00)00821-9
25. Preminger GM. Management of lower pole renal calculi: shock wave lithotripsy versus percutaneous nephrolithotomy versus flexible ureteroscopy. *Urol Res* 2006;**34**:108–11. https://doi.org/10.1007/s00240-005-0020-6
26. Stav K, Cooper A, Zisman A, Leibovici D, Lindner A, Siegel YI. Retrograde intrarenal lithotripsy outcome after failure of shock wave lithotripsy. *J Urol* 2003;**170**:2198–201. https://doi.org/10.1097/01.ju.0000096421.56225.79
27. Skolarikos A, Jung H, Neisius A, Petřík A, Somani B, Tailly T, *et al. EAU Guidelines on Urolithiasis*. 2023. URL: https://d56bochluxqnz.cloudfront.net/documents/full-guideline/EAU-Guidelines-on-Urolithiasis-2023.pdf (accessed 5 May 2023).
28. National Institute for Health and Care Excellence. *Renal and Ureteric Stones: Assessment and Management*. NICE National Institute for Health and Care Excellence; 2019. URL: www.nice.org.uk/guidance/ng118 (accessed 5 May 2023).
29. Yuruk E, Binbay M, Sari E, Akman T, Altinyay E, Baykal M, *et al.* A prospective, randomized trial of management for asymptomatic lower pole calculi. *J Urol* 2010;**183**:1424–8. https://doi.org/10.1016/j.juro.2009.12.022
30. Pearle MS, Lingeman JE, Leveillee R, Kuo R, Preminger GM, Nadler RB, *et al.* Prospective, randomized trial comparing shock wave lithotripsy and ureteroscopy for lower pole caliceal calculi 1 cm or less. *J Urol* 2005;**173**:2005–9. https://doi.org/10.1097/01.ju.0000158458.51706.56
31. Donaldson JF, Lardas M, Scrimgeour D, Stewart F, MacLennan S, Lam TBL, McClinton S. Systematic review and meta-analysis of the clinical effectiveness of shock wave lithotripsy, retrograde intrarenal surgery, and percutaneous nephrolithotomy for lower-pole renal stones. *Eur Urol* 2015;**67**:612–6. https://doi.org/10.1016/j.eururo.2014.09.054
32. Kumar A, Nanda B, Kumar N. 1827 a prospective randomized comparison between shock wave lithotripsy and flexible ureterorenoscopy for lower calyceal stones less than 2 cm: a single center experience. *J Urol* 2013;**189**:e750.
33. Salem A, Saad I, Emran A, Abdelhakiem M, Abdelrazzak O, Abdelkader M. 1829 laser lithotripsy versus ESWL for lower calyceal renal stones. *J Urol* 2013;**189**:e751. https://doi.org/10.1016/ j.juro.2013.02.2192
34. Sener NC, Imamoglu MA, Bas O, Ozturk U, Goktug HNG, Tuygun C, Bakirtas H. Prospective randomized trial comparing shock wave lithotripsy and flexible ureterorenoscopy for lower pole stones smaller than 1 cm. *Urolithiasis* 2014;**42**:127–31. https://doi.org/10.1007/s00240-013-0618-z

35. Kuo R, Lingeman JE, Leveillee RJ, Pearle MS, Watkins S, Fineberg NS; Lower Pole Study Group. Lower pole II: Initial results from a comparison of shock wave lithotripsy (SWL), ureteroscopy (URS), and percutaneous nephrostolithotomy (PNL) for lower pole nephrolithiasis. *J Urol* 2003;**169**:486.

36. Singh BP, Prakash J, Sankhwar SN, Dhakad U, Sankhwar PL, Goel A, Kumar M. Retrograde intrarenal surgery vs extracorporeal shock wave lithotripsy for intermediate size inferior pole calculi: a prospective assessment of objective and subjective outcomes. *Urology* 2014;**83**:1016–22. https://doi.org/10.1016/j.urology.2013.12.026

37. McClinton S, Starr K, Thomas R, MacLennan G, Lam T, Hernandez R, *et al.* The clinical and cost effectiveness of surgical interventions for stones in the lower pole of the kidney: the percutaneous nephrolithotomy, flexible ureterorenoscopy and extracorporeal shockwave lithotripsy for lower pole kidney stones randomised controlled trial (PUrE RCT) protocol. *Trials* 2020;**21**:479. https://doi.org/10.1186/s13063-020-04326-x

38. The EuroQol Group. EuroQol: a new facility for the measurement of health-related quality of life. *Health Policy* 1990;**16**:199–208.

39. Brooks R. EuroQol: the current state of play. *Health Policy* 1996;**37**:53–72.

40. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol group. *Ann Med* 2001;**33**:337–43. https://doi.org/10.3109/07853890109002087

41. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, *et al.* Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;**20**:1727–36. https://doi.org/10.1007/s11136-011-9903-x

42. Janssen MF, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, *et al.* Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Qual Life Res* 2013;**22**:1717–27. https://doi.org/10.1007/s11136-012-0322-4

43. Buchholz I, Janssen MF, Kohlmann T, Feng YS. A systematic review of studies comparing the measurement properties of the three-level and five-level versions of the EQ-5D. *PharmacoEconomics* 2018;**36**:645–61. https://doi.org/10.1007/s40273-018-0642-5

44. Janssen MF, Bonsel GJ, Luo N. Is EQ-5D-5L better than EQ-5D-3L? A head-to-head comparison of descriptive systems and value sets from seven countries. *PharmacoEconomics* 2018;**36**:675–97. https://doi.org/10.1007/s40273-018-0623-8

45. Ware JE, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;**34**:220–33.

46. Ware JE, Kosinski M, Turner-Bowker DM, Gandek B. *User's Manual for the SF-12v2 Health Survey*. Lincoln, RI: QualityMetric; 2009.

47. Hernández-Alava M, Pudney S, Wailoo A. *Estimating the Relationship between EQ-5D-5L and EQ-5D-3L: Results from an English Population Study*. The University of Sheffield; 2020. URL: https://eepru.sites.sheffield.ac.uk/reports-publications/reports (accessed 22 November 2023).

48. Farivar SS, Liu H, Hays RD. Half standard deviation estimate of the minimally important difference in HRQOL scores? *Exp Rev Pharmacoecon Outcomes Res* 2004;**4**:515–23. https://doi.org/10.1586/14737167.4.5.515

49. Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Qual Life Res* 2005;**14**:1523–32. https://doi.org/10.1007/s11136-004-7713-0

50. Treweek S, Innes K, MacLennan G. Sending Christmas cards to trial participants to improve retention. *Trials* 2017;**18**:406.

51. Coleman E, Arundel C, Clark L, Doherty L, Gillies K, Hewitt C, *et al.* Bah humbug! association between sending Christmas cards to trial participants and trial retention: randomised study within a trial conducted simultaneously across eight host trials. *BMJ* 2021;**375**:e067742. https://doi.org/10.1136/bmj-2021-067742

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

52. Ragab M, Baldin N, Collie J, Tran MGB, Al-Hayek S, S Parsy K, *et al.* Qualitative exploration of the renal stone patients' experience and development of the renal stone-specific patient-reported outcome measure. *BJU Int* 2020;**125**:123–32. https://doi.org/10.1111/bju.14873

53. Zou G. A modified Poisson regression approach to prospective studies with binary data. *Am J Epidemiol* 2004;**159**:702–6. https://doi.org/10.1093/aje/kwh090

54. Cooper K, Breeman S, Scott NW, Scotland G, Hernández R, Clark TJ, *et al.* Laparoscopic supracervical hysterectomy compared with second-generation endometrial ablation for heavy menstrual bleeding: the HEALTH RCT. *Health Technol Assess* 2019;**23**:1–108. https://doi.org/10.3310/hta23530

55. Ahmed I, Innes K, Brazzelli M, Gillies K, Newlands R, Avenell A, *et al.* Protocol for a randomised controlled trial comparing laparoscopic cholecystectomy with observation/conservative management for preventing recurrent symptoms and complications in adults with uncomplicated symptomatic gallstones (C-gall trial). *BMJ Open* 2021;**11**:e039781. https://doi.org/10.1136/bmjopen-2020-039781

56. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. 4th edn. Oxford: OUP; 2015.

57. Glick HA, Doshi JA, Sonnad SS, Polsky D. *Economic Evaluation in Clinical Trials*. Oxford: OUP; 2014.

58. Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs AH, Carswell C, *et al.* Consolidated health economic evaluation reporting standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. *Appl Health Econ Health Policy* 2022;**20**:213–21. https://doi.org/10.1007/s40258-021-00704-x

59. Jones K, Burns A. *Unit Costs of Health and Social Care 2021*. Kent: Personal Social Services Research Unit; 2021. URL: https://kar.kent.ac.uk/id/eprint/92342 (accessed 22 November 2023).

60. Curtis LA, Burns A. *Unit Costs of Health and Social Care 2019*. Canterbury: Personal Social Services Research Unit, University of Kent; 2019.

61. Data Intelligence PHS. *Scottish Health Service Costs 2019*. Public Health Scotland. URL: www.isdscotland.org/health-topics/finance/costs/ (accessed 1 November 2020).

62. NHS England. *2020/21 National Cost Collection Data Publication*. NHS England; 2022. URL: www.england.nhs.uk/publication/2020-21-national-cost-collection-data-publication/ (accessed 17 May 2023).

63. Office for National Statistics. *Earnings and Hours Worked, Age Group: ASHE Table 6*. Office of National Statistics; 2023. URL: www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/datasets/agegroupashetable6 (accessed 8 November 2023).

64. UK Government. *Expenses and Benefits: Business Travel Mileage for Employees' Own Vehicles*. GOV.UK; 2023. URL: www.gov.uk/expenses-and-benefits-business-travel-mileage/rules-for-tax (accessed 8 November 2023).

65. National Institute for Health and Care Excellence. *NICE Health Technology Evaluations: The Manual (PMG36)*; 2022. URL: www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation (accessed 22 November 2023).

66. Hernández-Alava M, Pudney S, Wailoo A. *Estimating the Relationship between EQ-5D-5L and EQ-5D-3L: Results from an English Population Study*. The University of Sheffield; 2020. URL: https://eepru.sites.sheffield.ac.uk/reports-publications/reports (accessed 22 November 2023).

67. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal 2013 (PMG9)*. National Institute for Health and Care Excellence (NICE); 2013. URL: www.nice.org.uk/process/pmg9 (accessed 22 November 2023).

68. Gray AM, Clarke PM, Wolstenholme JL, Wordsworth S. *Applied Methods of Cost-Effectiveness Analysis in Healthcare*. Oxford: OUP; 2010.

69. Buxton J. *National Life Tables*: UK. Office for National Statistics; 2021. URL: www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables (accessed 1 June 2023).

70. Emmott AS, Brotherhood HL, Paterson RF, Lange D, Chew BH. Complications, re-intervention rates, and natural history of residual stone fragments after percutaneous nephrolithotomy. *J Endourol* 2018;**32**:28–32. https://doi.org/10.1089/end.2017.0618

71. Wei Y, Royston P. Reconstructing time-to-event data from published Kaplan–Meier curves. *Stata J* 2017;**17**:786–802.

72. D'Costa MR, Pais VM, Rule AD. Leave no stone unturned: defining recurrence in kidney stone formers. *Curr Opin Nephrol Hypertens* 2019;**28**:148–53. https://doi.org/10.1097/MNH.0000000000000478

73. Vaughan LE, Enders FT, Lieske JC, Pais VM, Rivera ME, Mehta RA, *et al.* Predictors of symptomatic kidney stone recurrence after the first and subsequent episodes. *Mayo Clin Proc* 2019;**94**:202–10. https://doi.org/10.1016/j.mayocp.2018.09.016

74. Briggs A, Sculpher M, Claxton K. *Decision Modelling for Health Economic Evaluation*. Oxford: OUP; 2006.

75. Buyukkaramikli NC, Rutten-van Molken MPMH, Severens JL, Al M. TECH-VER: a verification checklist to reduce errors in models and improve their credibility. *PharmacoEconomics* 2019;**37**:1391–408. https://doi.org/10.1007/s40273-019-00844-y

76. National Institute for Health and Care Excellence. *Technology Appraisal Processes*. National Institute for Health and Care Excellence; 2023. URL: www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-technology-appraisal-guidance/process (accessed June 2023).

77. Thorn JC, Davies CF, Brookes ST, Noble SM, Dritsaki M, Gray E, *et al.* Content of health economics analysis plans (HEAPs) for trial-based economic evaluations: expert Delphi consensus survey. *Value Health* 2021;**24**:539–47. https://doi.org/10.1016/j.jval.2020.10.002

78. Sahin C, Tuncer M, Yazıcı O, Horuz R, Çetinel AC, Eryıldırım B, *et al.* Do the residual fragments after shock wave lithotripsy affect the quality of life? *Urology* 2014;**84**:549–54. https://doi.org/10.1016/j.urology.2014.05.012

79. National Institute for Health and Care Excellence. *Renal and Ureteric Stones: Assessment and Management (NICE guideline 118)*. National Institute for Health and Care Excellence; 2019. URL: www.nice.org.uk/guidance/ng118/evidence (accessed 4 March 2023).

80. Bosio A, Alessandria E, Dalmasso E, Agosti S, Vitiello F, Vercelli E, *et al.* Flexible ureterorenoscopy versus shockwave lithotripsy for kidney stones. *Eur Urol Focus* 2022;**8**:1816–22. https://doi.org/10.1016/j.euf.2022.04.004

81. Pearle MS, Lingeman JE, Leveillee R, Kuo R, Preminger GM, Nadler RB, *et al.* Prospective randomized trial comparing shock wave lithotripsy and ureteroscopy for lower pole caliceal calculi 1 cm or less. *J Urol* 2008;**179**:69. https://doi.org/10.1016/j.juro.2008.03.140

82. Geraghty RM, Jones P, Herrmann TRW, Aboumarzouk O, Somani BK. Ureteroscopy is more cost effective than shock wave lithotripsy for stone treatment: systematic review and meta-analysis. *World J Urol* 2018;**36**:1783–93. https://doi.org/10.1007/s00345-018-2320-9

83. Koo V, Young M, Thompson T, Duggan B. Cost-effectiveness and efficiency of shockwave lithotripsy vs flexible ureteroscopic holmium: yttrium-aluminium-garnet laser lithotripsy in the treatment of lower pole renal calculi. *BJU Int* 2011;**108**:1913–6. https://doi.org/10.1111/j.1464-410X.2011.10172.x

84. Dasgupta R, Cameron S, Aucott L, MacLennan G, Kilonzo MM, Lam TB, *et al.* Shockwave lithotripsy compared with ureteroscopic stone treatment for adults with ureteric stones: the TISU non-inferiority RCT. *Health Technol Assess* 2022;**26**:1–70. https://doi.org/10.3310/WUZW9042

85. Zhang H, Hong TY, Li G, Jiang N, Hu C, Cui X, *et al.* Comparison of the efficacy of ultra-mini PCNL, flexible ureteroscopy, and shock wave lithotripsy on the treatment of 1–2 cm lower pole renal calculi. *Urol Int* 2019;**102**:153–9.

86. Bozzini G, Verze P, Arcaniolo D, Dal Piaz O, Buffi NM, Guazzoni G, *et al.* A prospective randomized comparison among SWL, PCNL and RIRS for lower calyceal stones less than 2 cm: a multicenter experience: a better

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

understanding on the treatment options for lower pole stones. *World J Urol* 2017;**35**:1967–75. https://doi.org/10.1007/s00345-017-2084-7

87. Zeng G, Zhang T, Agrawal M, He X, Zhang W, Xiao K, *et al.* Super-mini percutaneous nephrolithotomy (SMP) vs retrograde intrarenal surgery for the treatment of 1–2 cm lower-pole renal calculi: an international multicentre randomised controlled trial. *BJU Int* 2018;**122**:1034–40. https://doi.org/10.1111/bju.14427
88. Jiang K, Chen H, Yu X, Chen Z, Ye Z, Yuan H. The 'all-seeing needle' micro-PCNL versus flexible ureterorenoscopy for lower calyceal stones of ≤ 2 cm. *Urolithiasis* 2019;**47**:201–6. https://doi.org/10.1007/s00240-018-1049-7
89. Fayad AS, Elsheikh MG, Ghoneima W. Tubeless mini-percutaneous nephrolithotomy versus retrograde intrarenal surgery for lower calyceal stones of ≤2 cm: a prospective randomised controlled study. *Arab J Urol* 2017;**15**:36–41. https://doi.org/10.1016/j.aju.2016.10.002
90. Yavuz A, Kilinc MF, Bayar G. Outcomes of different minimally invasive techniques in lower calyceal stones of 1 to 2 centimeters: a prospective, randomized study. *Arch Esp Urol* 2020;**73**:307–15.
91. Kandemir E, Savun M, Sezer A, Erbin A, Akbulut MF, Sarilar O. Comparison of miniaturized percutaneous nephrolithotomy and standard percutaneous nephrolithotomy in secondary patients: a randomized prospective study. *J Endourol* 2020;**34**:26–32. https://doi.org/10.1089/end.2019.0538
92. Jin L, Yang B, Zhou Z, Li N. Comparative efficacy on flexible ureteroscopy lithotripsy and miniaturized percutaneous nephrolithotomy for the treatment of medium-sized lower-pole renal calculi. *J Endourol* 2019;**33**:914–9. https://doi.org/10.1089/end.2019.0504
93. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. *BMJ* 2011;**342**:d40. https://doi.org/10.1136/bmj.d40
94. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. Wiley Classics Library ed. United States of America: John Wiley & Sons Inc; 2004.
95. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, *et al*. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;**343**:d5928.
96. Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the journal of clinical epidemiology. *J Clin Epidemiol* 2011;**64**:380–2.
97. Kumar A, Kumar N, Vasudeva P, Kumar Jha S, Kumar R, Singh H. A prospective, randomized comparison of shock wave lithotripsy, retrograde intrarenal surgery and miniperc for treatment of 1 to 2 cm radiolucent lower calyceal renal calculi: a single center experience. *J Urol* 2015;**193**:160–4. https://doi.org/10.1016/j.juro.2014.07.088
98. Sener NC, Bas O, Sener E, Zengin K, Ozturk U, Altunkol A, Evliyaoglu Y. Asymptomatic lower pole small renal stones: shock wave lithotripsy, flexible ureteroscopy, or observation? A prospective randomized trial. *Urology* 2015;**85**:33–7.
99. Vilches RM, Aliaga A, Reyes D, Sepulveda F, Mercado A, Moya F, *et al.* Comparison between retrograde intrarenal surgery and extracorporeal shock wave lithotripsy in the treatment of lower pole kidney stones up to 15 mm. prospective, randomized study. *Actas Urol Esp* 2015;**39**:236–42.
100. Kandemir A, Guven S, Balasar M, Sonmez MG, Taskapu H, Gurbuz R. A prospective randomized comparison of micropercutaneous nephrolithotomy (microperc) and retrograde intrarenal surgery (RIRS) for the management of lower pole kidney stones. *World J Urol* 2017;**35**:1771–6.

Appendix 1 Summary of breaches in the PUrE trials

TABLE 32 Summary of breaches

Breach no., date	Summary of breach
1, 27 July 2017	Breach to consent procedure. Participant randomised without signing written consent on the basis of documented verbal consent Breach assessed as non-serious and closed out on 25 August 2017
2, 15 February 2019	Breach to consent procedure. Participant randomised prior to signing written consent on the basis of documented verbal consent and data collected before written consent was obtained Breach assessed as non-serious and closed out on 25 February 2019
3, 20 June 2018	Breach to consent procedure. Participant randomised prior to signing written consent on the basis of documented verbal consent and data collected before written consent was obtained Breach assessed as non-serious and closed out on 25 February 2019
4, 26 July 2017	Breach to consent procedure. Participant randomised prior to signing written consent on the basis of documented verbal consent Breach assessed as non-serious and closed out on 4 March 2019
5, 28 March 2018	Breach to consent procedure. Participant randomised prior to signing written consent on the basis of documented verbal consent Breach assessed as non-serious and closed out on 4 March 2019

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Appendix 2 Outcome and variable derivations and definitions

Area under the curve

The PO is health status as measured by the EQ-5D score AUC reported from an anchor point up to and including 12 weeks post intervention.

Using the trapezoidal area formula, a section of the AUC of the EQ-5D between 2 consecutive follow-up observations can be calculated as the mean of the 2 EQ-5D values multiplied by the time difference in between. Thus, for each time epoch, an element of the AUC will be determined using the formula:

$a_f = 0.5\ (e_f + e_{f-1})(t_f - [t_{f-1}])$, over

where t is the time of a follow-up observation, f since randomisation, and e_f is the EQ-5D collected at that fth 'follow-up' point.

To find the total AUC, all these time epoch elements will be cumulated using Σa_f.

Time epochs. Definition of the required specified time point for the primary outcome

EQ-5D scores were derived using current NICE (UK) recommendations with validated Stata code from EuroQol crosswalk to the 3L tool[46] requiring all five EQ-5D-5L questionnaire domains to be completed. The code also accounts for age and sex.

For the PO to be valid, we specified each participate should have responses at three specific time zones:

1. An EQ-5D score at baseline and/or pre intervention (see Notes 1 and 2 below on missing data).
2. At least one EQ-5D score between 1 and 3 weeks post intervention.

And

3. At least one EQ-5D score from weeks 4 to 12 post intervention.

Notes:

1. A negative EQ-5D score would remain 'negative', and thus decrease the total calculated AUC reflecting the worse health status overall.
2. Missing baseline values were imputed for the EQ-5D score using the average within each gender-specific treatment arm (imputing mean baseline to intervention time). These were then slotted in with the actual EQ-5D in our database and the closest 'real' value selected. Although this missingness was expected to be very low, this approach ensured that a participant who had made sufficient contribution otherwise was not excluded from the main model which required the baseline as a covariate.
3. Pre-intervention EQ-5D score had more missingness. Provided sufficient repeated measures for criteria 2 and 3 above, we produced an estimate of the pre intervention based on a linear model of all the responses received prior to intervention – often this was just the baseline. Again, these were slotted within actual EQ-5D score observations in our data base and closest 'real' value selected.

4. If the last valid observed value was > 6 weeks post intervention but ≤ 12 weeks, then this last value carried forward for the intervening weeks up to the 12th week.
5. If the last observation is > 3 weeks post intervention but < 6 weeks, then other outcomes (e.g. death, complications, additional interventions other than those in the accepted treatment pathway) were examined. If considered stable, the last observed value was carried forward to the 12th week.
6. Any other missing EQ-5D scores for post-intervention follow-up remained missing for the main ITT analysis.

Various models to best reflect the protocol primary outcome

At the original study plan, the time between baseline and 12 weeks post intervention was not considered to be very different between treatment arms. However, for surgical treatments there are longer waiting times which have increased over time, exacerbated by the COVID-19 pandemic. As such, there is no common time period between participants. To still reflect the protocol PO, we decided to capture its essence by:

Three PO measures:

- Main PO – the main PO:
  - AUC of EQ-5D scores from pre intervention to 12 weeks, averaged over 12 weeks with an adjustment for baseline EQ-5D scores.
- Sensitivity 1:
  - The same averaged AUC as above but with two covariates – baseline EQ-5D score and 'waiting times'.
- Sensitivity 2:
  - The AUC as main PO, adjusted for clusters but without other covariates. Post hoc we also adjusted for the pre-intervention measure to account for the observed different baseline values.

Secondary outcome:

- AUC from baseline to 12 weeks – this is an AUC of all EQ-5D scores anchored from baseline right up to 12 weeks post intervention.

This final model includes any reported data at variable times. So that all these models may be compared for interpretation, all the AUC outcomes were standardised as weekly averages.

Description of different scenarios for multiple imputation

We used multiple imputation to assess the impact of missing data on effect estimates for the PO. The approach used was multiple imputation using chained equations (MICE) under a missing at random (MAR) assumption.

The general approach followed was: The auxiliary model included baseline and pre-intervention EQ-5D scores, age, sex, BMI, stone density and waiting time. MICE was used to impute any missing covariates. The percentage level of missingness (about 25%) implied 30 imputation data sets was sufficient. Outcome data were separately imputed for each treatment arm (within each RCT). Within each RCT, the treatment-arm imputed data sets were then combined and the imputed data sets analysed Rubin's rules.[93,94] This was done for each of the four analyses of the PO.

Participants with missing PO data arose in two ways in both RCTs: those completely missing any follow-up information or those partially missing follow-up data (i.e. too few EQ-5D-5L measurements to generate an AUC). The differential missingness type meant two approaches for imputation were considered:

i. *Basic*: Missing AUC data were imputed at the participant level. In this scenario, all partially observed EQ-5D follow-up data were ignored, and completed AUC outcome was imputed.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

ii. *Block*: We used the partially observed AUC and only imputed sections of the AUC that were missing. We did this by splitting the follow-up time into blocks:
   - Block 0: observations from randomisation up to the intervention [note most were sufficiently complete];
   - Block 1: 1–2 weeks post intervention;
   - Block 2: 3–5 weeks post intervention;
   - Block 3: 6–12 weeks post intervention.
     If a participant had some information within a block, this was utilised using their individual block average over the time of the block. If a block was completely missing for a participant that block was imputed using MICE. This mapped that individual to those with a comparable baseline EQ-5D scores, pre-intervention EQ-5D scores, age and gender for that specific block (the auxiliary model).

The analysis models were the same as the complete-case models detailed in the SAP www.abdn.ac.uk/hsru/what-we-do/trials-unit/portfolio/pure-123.

Secondary outcome measures: definitions and derivations

Time was divided into weekly epochs. If there were multiple observations used within each epoch, the average was used. Epochs were later used to define the three phases: baseline/pre intervention; weeks 1 to 3, and finally week 4 to 12.

TABLE 33 Stone clearance categories – definitions

Stone clearance	At 12 weeks post intervention
Complete clearance of the target stone from the urinary tract defined as no further action or observation required for that stone; acceptable clearance, where observation is required but no intervention planned; and unacceptable clearance, where further intervention will be required	*Complete* • Image done No, Stone present No, and Additional intervention No • Image done Yes, Stone present No, and Additional intervention No • Image done Yes, Stone present No, Fragments No, and Additional intervention No ***Acceptable*** • Image Yes, Stone present Yes and Fragments Yes (largest < 2 mm) and Additional intervention No • Image done No, and Additional intervention No • Image Yes, Stone present Yes, Fragments Yes (largest > 2 mm but < 4 mm), and Additional intervention No *Unacceptable* • Image Yes, Stone present Yes, Fragments No, and Additional intervention No • Image Yes, Stone present No, Fragments Yes (largest > 4 mm), and Additional intervention No • Image No and Additional intervention Yes • Image Yes and Additional intervention Yes If this results in being missing, we class it as missing

TABLE 33 Stone clearance categories – definitions (*continued*)

Stone clearance	At 12 months post randomisation
	NB: At 12 months the stone question was phrased as 'Is the stone cleared?' and 'Is there need for additional intervention?' Imaging was not prompted
	Complete
	Stone cleared without any additional intervention (had or planned)
	Acceptable
	Stone cleared without any additional intervention (had or planned)
	Unacceptable
	Stone cleared but with additional intervention (had or planned) Stone not clear with or without additional intervention If missing, we class it as missing
	Note: • Twelve months post randomisation can be before the actual intervention – these are mostly considered as 'unacceptable' unless it is known the stone has passed without any intervention • Twelve months post randomisation can be between intervention and a later 12-week post-intervention period – these were classified as above initially and then a 'higher' view taken to assess if any planned/had other treatment would be part of their care pathway given their continued treatment to the 12-week time point

Additional intervention – definitions

Additional intervention was any procedure outside of the treatment pathways defined for this trial. Classification of the treatment pathway varied slightly for each of the interventions proposed for each treatment arm (of both RCT1 and RCT2).

TABLE 34 Permitted treatment pathways for each treatment arm in both RCT1 and RCT2

Treatment arm	Treatment pathway (defined for the PUrE trial)
FURS is a surgical procedure where a flexible endoscope is passed into the kidney through which laser energy is used to fragment the stone	Expect there to be a single procedure, although an additional FURs procedure will be considered provided it takes place with 6 weeks of the initial FURS procedure
	NB: • Placement or removal of a stent or catheter at time of surgery is accepted • Insertion of stent/catheter/nephrostomy tube post randomisation but pre intervention will also be classed as additional intervention.
ESWL involves a (sound) shockwave, outside the body focused onto the kidney stone through the patient's flank skin, causing it to fragment	Two separate ESWL treatments will be considered as part of the initial ESWL intervention strategy if within an 8-week period
	NB: • Placement or removal of a stent or catheter at time of surgery is seen as additional intervention • Insertion of stent/catheter/nephrostomy tube post randomisation but pre intervention will classed as additional intervention

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 34 Permitted treatment pathways for each treatment arm in both RCT1 and RCT2 (*continued*)

Treatment arm	Treatment pathway (defined for the PUrE trial)
PCNL is a surgical procedure to remove stones via a small (10 mm) incision in the patient's flank. It uses a needle, contrast dye, the use of a hollow rigid access sheath, and a rigid metal telescope (nephroscope) to eventually see the stone and either retrieve it whole or to fragment it using a variety of energy delivery devices. In addition, a urinary catheter may be inserted to drain the bladder for a short period after the procedure	A single PCNL treatment is expected to be required to completely remove stones up to 25 mm
	NB: • Placement or removal of a stent or catheter at time of surgery is accepted • The nephrostomy tube is also part of the treatment plan if done at time of surgery – both insertion and removal • Insertion of stent/catheter/nephrostomy tube post randomisation but pre intervention will also be classed as additional intervention
No treatment – agreed stone (may) have passed/asymptomatic/other reasons.	If some intervention treatment is carried out, will be classed as additional intervention
	NB: Insertion of stent/catheter/nephrostomy tube post randomisation but before the nominated 'intervention date', will be classed as additional intervention

TABLE 35 Description of participant response outcome measured details of missing values protocols used

EQ-5D-5L Index all time points and supplementary/additional	We used the validated Stata code (package called eq5dmap) approved by NICE for use in the UK[47] This is required currently by NICE Missing elements then the index score is classed as missing
NRS all time points and supplementary/additional	The NRS is a segmented numeric version of the VAS in which a respondent selects a whole number (0–10 integers) that best reflects the intensity of their pain
SF-12 PCS and MCS (baseline, 12 weeks, and 12 months)	In-house validated code[45]

VAS, visual analogue scale.

Subgroup analyses and technical factors – definitions

Subgroup analyses

Subgroup analyses explored possible modification of treatment outcome effect by important factors on the main PO model, its two sensitivity models and for the two secondary outcomes: AUC from baseline to 12 weeks and stone clearance by 12 weeks (complete/acceptable/unacceptable). The subgroups identified were:

- Participant BMI: This was not always completed as it was mainly measured at time of the intervention not all participants had an intervention, for example the stone passed, conservative management was agreed, etc.). Where possible, patient records were examined if observed at other times. Hence, we compared the possible impact on treatment of a participant BMI being normal/overweight/obese/extreme, as a casewise analysis.

- Stone size (maximum dimension and volume): Since RCT1 only includes those with initial stones < 10 mm we consider only for RCT2. In addition, only the maximum dimension was assessed, comparing 10–15 versus > 15 mm.
- Stone density on CTKUB (Hounsfield units) to compared ≤ 1000 versus > 1000.

All subgroups included treatment-by-factor interactions in the model. A stricter level of statistical significance (two-sided 1% significance level) was applied to these analyses given their exploratory nature. Corresponding 99% CIs will therefore be calculated.

Technical factors impact

We described within each allocated group how technical factors impact the main PO (and its associated sensitivity outcomes) and for the two secondary outcomes: AUC from baseline to 12 weeks and stone clearance by 12 weeks (complete/acceptable/unacceptable). This was only for the initial intervention.

- access sheath versus no access sheath and digital versus non-digital instrument (FURS)
- fixed site versus mobile device (ESWL)
- skin-to-stone distance: establishing the cut-off categories was a post hoc process in collaboration with our project management team. Categories decided were: < 120 versus ≥ 120 mm
- calibre of access tract (PCNL).

Note: Because each of these were only applicable to a the specific treatments, the comparison is not between treatments – we have reported only between each category of each technical factor and only as summaries level not as any effect estimate.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Appendix 3 Summary of protocol amendments to the PUrE trials

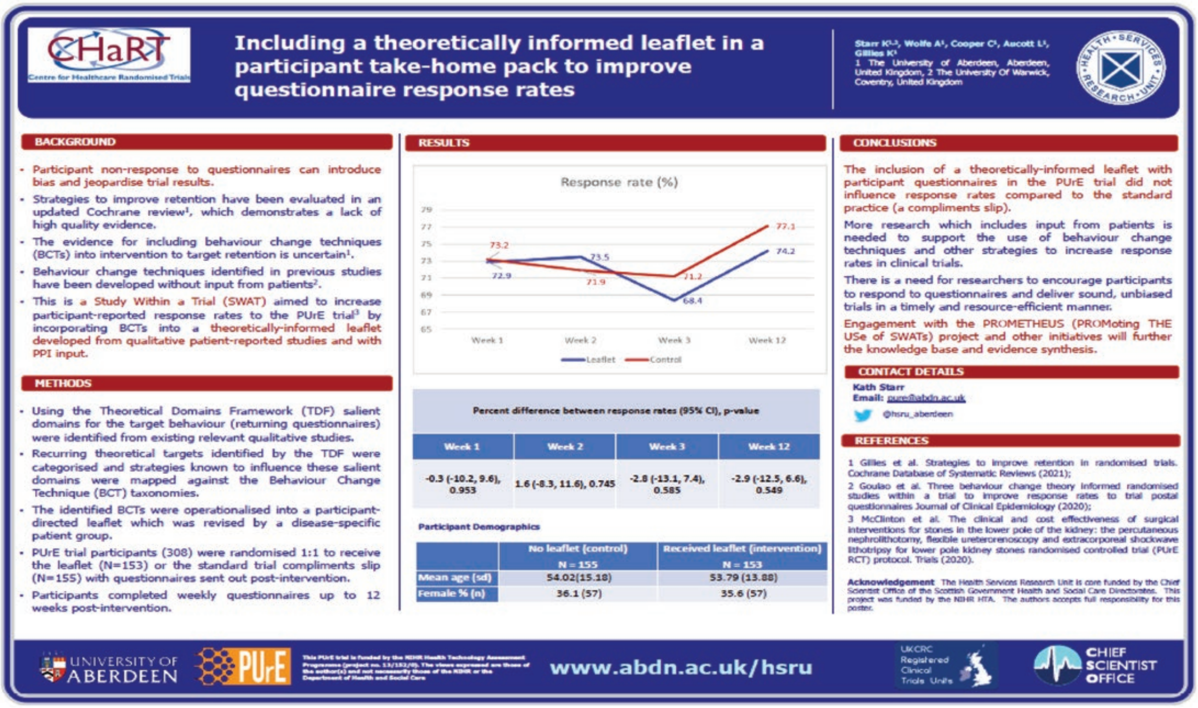
TABLE 36 Protocol amendment summary

Protocol version, date	Summary of revisions
1, 25 January 2016	Amendments to: • Introduction of weekly participant questionnaires for 12 weeks post intervention • Clarification that imaging should be part of standard care and not a trial process • Introduction of tokens of appreciation when participants return their 12-week and 12-month questionnaire

2, 2 September 2019	Amendments to:
	• Additional questionnaire for a subset of participants (CReSP)
	• Administrative changes:
	• Extension of the trial timelines
	• Removal of errors in the text:
	  • 'stratification of randomisation by site'
	  • 'stone mineral composition'
	• Clarification of the end of trial for the participant to account for cases when 12 weeks post intervention may occur after 12 months post randomisation

3, 8 October 2020	Amendments to: • Extension of recruitment end date to September 2020
4, 17 November 2021	Amendments to: • Last patient last visit now extended to August 2022 and end of study funding is February 2023, to align with the extended recruitment period due to COVID-19 delays • Funder has awarded a 9-month extension to allow for follow-up to be completed on remaining participants. No resource implication locally, this just means we have sufficient time to now complete follow-ups (as per protocol) on those participants recruited after the COVID-19 recruitment pause and for those participants for whom treatment was delayed due to the pandemic • Trial Steering Group membership also updated to reflect the current independent members

Appendix 4 Theoretically informed leaflet to improve response rates: presented at the 6th International Clinical Trials Methodology Conference, October 2022



Copyright © 2025 Wiseman *et al*. This work was produced by Wiseman *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Appendix 5 Recruitment and detailed stone measurements

Recruitment – by centre and over time

TABLE 37 Numbers recruited from each centre – RCT1

Centre number	FURS total = 232	ESWL total = 234	Overall total = 466	Months recruiting[a]
11	9	14	23	54.1
12	21	13	34	53.9
13[b]	0	2	2	40.0
14	5	4	9	51.7
15	4	2	6	51.5
16[b]	2	1	3	39.4
17[b]	15	16	31	38.5
18	2	6	8	50.3
19[b]	6	5	11	38.7
20[b]	3	1	4	37.6
21	2	1	3	51.2
23[b]	2	0	2	38.1
24[b]	4	3	7	38.5
25	1	4	5	49.2
26	18	9	27	50.2
27[b]	5	4	9	35.4
28[b]	3	4	7	32.3
29[b]	2	2	4	35.9
30[b]	1	0	1	28.1
31[b]	7	9	16	35.5
32[b]	6	10	16	34.3
34	4	11	15	44.9
35[b]	2	5	7	27.7
36[b]	9	8	17	32.0
37	17	29	46	44.0
38	20	14	34	43.7
40[b]	1	1	2	27.4
41	0	4	4	40.1
42	5	4	9	43.6
43[b]	1	2	3	29.9

TABLE 37 Numbers recruited from each centre – RCT1 (*continued*)

Centre number	FURS total = 232	ESWL total = 234	Overall total = 466	Months recruiting[a]
44[b]	1	1	2	23.9
45[b]	1	2	3	27.9
46[b]	7	4	11	28.4
47[b]	4	5	9	27.0
48[b]	4	0	4	27.1
49[b]	5	1	6	26.5
50	4	2	6	38.3
51[b]	0	1	1	23.9
52[b]	2	0	2	24.7
53[b]	3	1	4	25.3
54	4	7	11	36.9
55[b]	2	6	8	20.1
56	7	2	9	34.7
57	0	2	2	30.5
58	5	2	7	29.7
59	1	4	5	31.4
60[b]	0	2	2	16.5
62	5	4	9	21.9
TOTAL	232	234	466	.

a Excludes 4-month enforced closure due to COVID-19.
b Did not reopen after the pause in recruitment due to COVID-19.

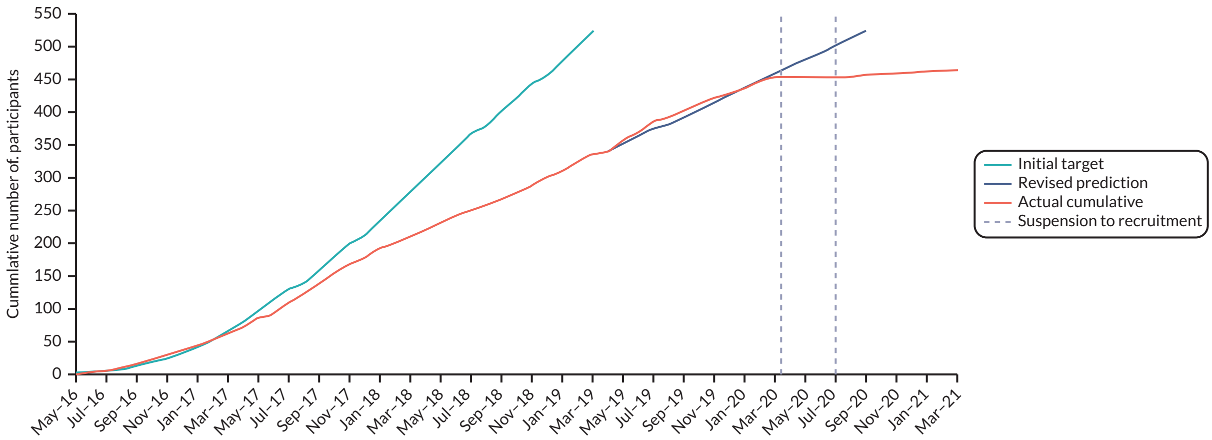


FIGURE 16 Cumulative numbers recruited by month – RCT1. Note: Initial target was at funding: revised prediction was for the extension to recruitment.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 38 Numbers recruited from each centre – RCT2

Centre number	FURS total = 73	PNCL total = 86	Overall total = 159	Months recruiting[a]
11	9	5	14	54.1
12	1	2	3	53.9
14	1	2	3	51.7
15	0	1	1	51.5
16[b]	0	1	1	39.4
17[b]	2	3	5	38.5
18	2	0	2	50.3
19[b]	1	2	3	38.7
21	1	2	3	51.2
23[b]	2	0	2	38.1
24[b]	1	4	5	38.5
25	1	0	1	49.2
26	10	13	23	50.2
27[b]	1	0	1	35.4
28[b]	1	0	1	32.3
30[b]	0	3	3	28.1
31[b]	1	1	2	35.5
32[b]	2	1	3	34.3
33	0	1	1	44.9
34	2	2	4	44.9
35[b]	0	1	1	27.7
36[b]	1	2	3	32.0
37	13	12	25	44.0
38	5	8	13	43.7
40[b]	1	0	1	27.4
41	2	1	3	40.1
42	1	1	2	43.6
43[b]	1	1	2	29.9
46[b]	2	3	5	28.4
47[b]	0	1	1	27.0
48[b]	0	1	1	27.1
49[b]	2	4	6	26.5
50	2	0	2	38.3
52[b]	0	1	1	24.7
54	0	3	3	36.9
55[b]	1	1	2	20.1

TABLE 38 Numbers recruited from each centre – RCT2 (*continued*)

Centre number	FURS total = 73	PNCL total = 86	Overall total = 159	Months recruiting[a]
56	2	1	3	34.7
57	1	1	2	30.5
62	1	1	2	21.9
Total	73	86	159	.

a Excludes 4-month enforced closure due to COVID-19.
b Did not reopen post COVID-19.

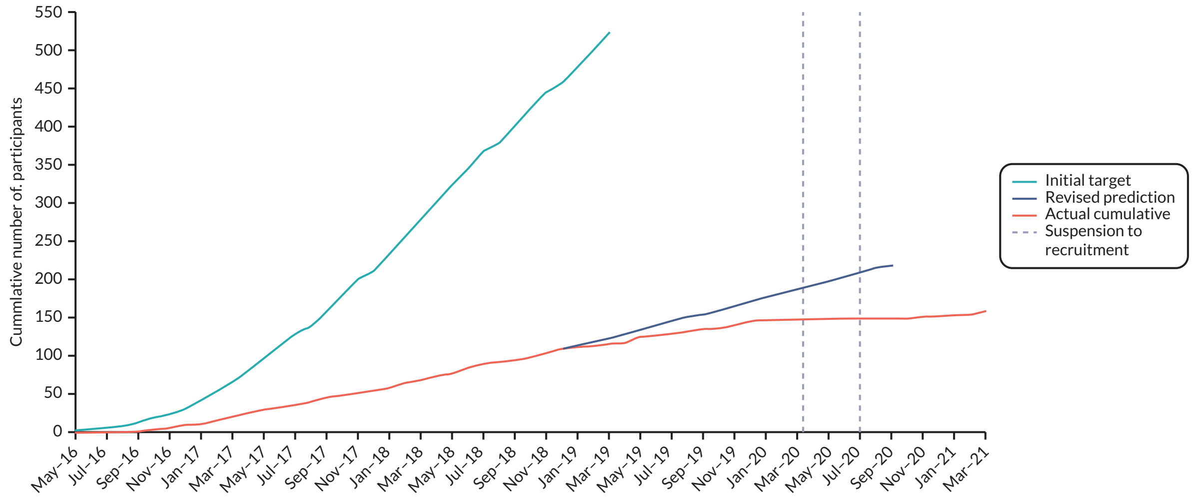


FIGURE 17 Cumulative numbers recruited by month time – RCT2. Note: Initial target was at funding: revised prediction was for the extension to recruitment.

Detailed stone measurements

TABLE 39 RCT1: detailed stone measurement

Detailed stone measurements	FURS *N* = 230	ESWL *N* = 231
Stone dimension:[a]		
Dimension 1[b]	*n* = 229	*n* = 231
Mean (SD)	6.8 (1.9)	6.7 (1.9)
Median (IQI)	6.5 (5.1–8.0)	6.9 (5.0–8.0)
Minimum, maximum	(3.0, 15.3)	(1.0, 14.0)
Dimension 2[b]	*n* = 229	*n* = 231
Mean (SD)	5.6 (1.9)	5.6 (1.9)
Median (IQI)	5.0 (4.0–7.0)	5.0 (4.0–7.0)
Minimum, maximum	(1.6, 10.0)	(2.0, 13.0)

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 39 RCT1: detailed stone measurement (*continued*)

Detailed stone measurements	FURS *N* = 230	ESWL *N* = 231
Dimension 3 Craniocaudal[c]	*n* = 229	*n* = 230
Mean (SD)	6.9 (2.1)	6.9 (2.1)
Median (IQI)	7.0 (5.3–8.0)	7.0 (5.0–8.3)
Minimum, maximum	(2.0, 14.6)	(1.6, 13.5)
Calculated stone volume[a]	*n* = 229	*n* = 230
Mean (SD)	162.0 (133.1)	157.0 (117.0)
Median (IQI)	126.1 (65.4–226.8)	131.9 (63.4–230.9)
Minimum, maximum	(7.7, 1068.7)	(1.8, 967.6)
Stone density (Hounsfield unit)[a]	*n* = 229	*n* = 230
Mean (SD)	821.4 (364.9)	746.5 (347.4)
Median (IQI)	797.0 (539.0–1070.0)	687.5 (494.0–976.0)
Minimum, maximum	(105.0, 1826.0)	(108.0, 1625.0)
Skin-to-stone distance:[a]		
at 0°	*n* = 230	*n* = 230
Mean (SD)	102.9 (21.9)	103.6 (20.2)
Median (IQI)	101.5 (90.0–116.0)	103.7 (90.0–115.0)
Minimum, maximum	(42.0, 179.7)	(54.4, 177.2)
at 45°	*n* = 229	*n* = 230
Mean (SD)	99.8 (23.7)	101.4 (22.9)
Median (IQI)	97.7 (85.5–113.0)	99.7 (86.0–116.0)
Minimum, maximum	(41.0, 200.0)	(52.8, 189.0)
at 90°	*n* = 229	*n* = 231
Mean (SD)	104.4 (24.4)	107.3 (25.2)
Median (IQI)	101.0 (88.0–120.0)	105.0 (92.0–120.0)
Minimum, maximum	(49.0, 200.0)	(46.0, 246.8)

a Continuous (or pseudo continuous) variable summaries *n*: mean (SD), median (IQI), (minimum, maximum): (IQI): interquartile interval.
b D1 and D2 – axial slice – maximal diameter in two planes perpendicular to each other.
c D3 – coronal slice/reconstruction – craniocaudal length.

TABLE 40 RCT2: detailed stone measurement

Detailed stone measurements	FURS *N* = 73	PCNL *N* = 86
Stone dimension:[a]		
Dimension 1[b]	*n* = 73	*n* = 86
Mean (SD)	12.4 (3.1)	12.9 (3.8)
Median (IQI)	12.0 (11.0–14.0)	12.2 (11.0–15.0)
Minimum, maximum	(4.6, 22.0)	(4.7, 24.0)

TABLE 40 RCT2: detailed stone measurement (*continued*)

Detailed stone measurements	FURS *N* = 73	PCNL *N* = 86
Dimension 2[b]	*n* = 73	*n* = 86
Mean (SD)	9.6 (2.7)	10.1 (2.9)
Median (IQI)	9.0 (8.0–11.0)	10.0 (8.0–12.0)
Minimum, maximum	(5.0, 17.0)	(4.0, 16.5)
Dimension 3 Craniocaudal[c]	*n* = 73	*n* = 86
Mean (SD)	13.1 (3.4)	13.8 (3.4)
Median (IQI)	12.8 (11.0–15.0)	13.8 (11.0–15.3)
Minimum, maximum	(6.0, 21.0)	(7.5, 24.0)
Calculated stone volume[a]	*n* = 73	*n* = 86
Mean (SD)	867.4 (503.4)	1009.1 (616.3)
Median (IQI)	707.9 (518.4–1114.2)	830.2 (505.2–1402.7)
Minimum, maximum	(173.5, 2430.6)	(181.2, 2925.3)
Hounsfield unit[a]	*n* = 73	*n* = 86
Mean (SD)	989.6 (347.7)	1008.3 (335.5)
Median (IQI)	995.0 (756.0–1328.0)	1059.5 (727.0–1267.0)
Minimum, maximum	(328.0, 1647.0)	(107.0, 1596.0)
Skin-to-stone distance:[a]		
at 0°	*n* = 73	*n* = 86
Mean (SD)	104.1 (23.5)	106.6 (20.6)
Median (IQI)	100.0 (91.0–115.1)	106.5 (90.0–119.9)
Minimum, maximum	(68.0, 182.0)	(66.9, 168.0)
at 45°	*n* = 73	*n* = 86
Mean (SD)	101.1 (24.8)	105.1 (23.5)
Median (IQI)	97.0 (84.0–113.7)	103.8 (88.0–120.0)
Minimum, maximum	(52.9, 182.3)	(63.5, 187.0)
at 90°	*n* = 73	*n* = 86
Mean (SD)	104.8 (24.6)	111.9 (23.9)
Median (IQI)	102.8 (85.0–120.0)	107.2 (96.0–125.0)
Minimum, maximum	(59.0, 173.2)	(72.0, 205.0)

a Continuous (or pseudo continuous) variable summaries *n*: mean (SD), median (IQI), (minimum, maximum): (IQI): interquartile interval.
b D1 and D2 – axial slice – maximal diameter in two planes perpendicular to each other.
c D3 – coronal slice/reconstruction – craniocaudal length.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Appendix 6 Health status: EQ-5D-VAS score plots over time

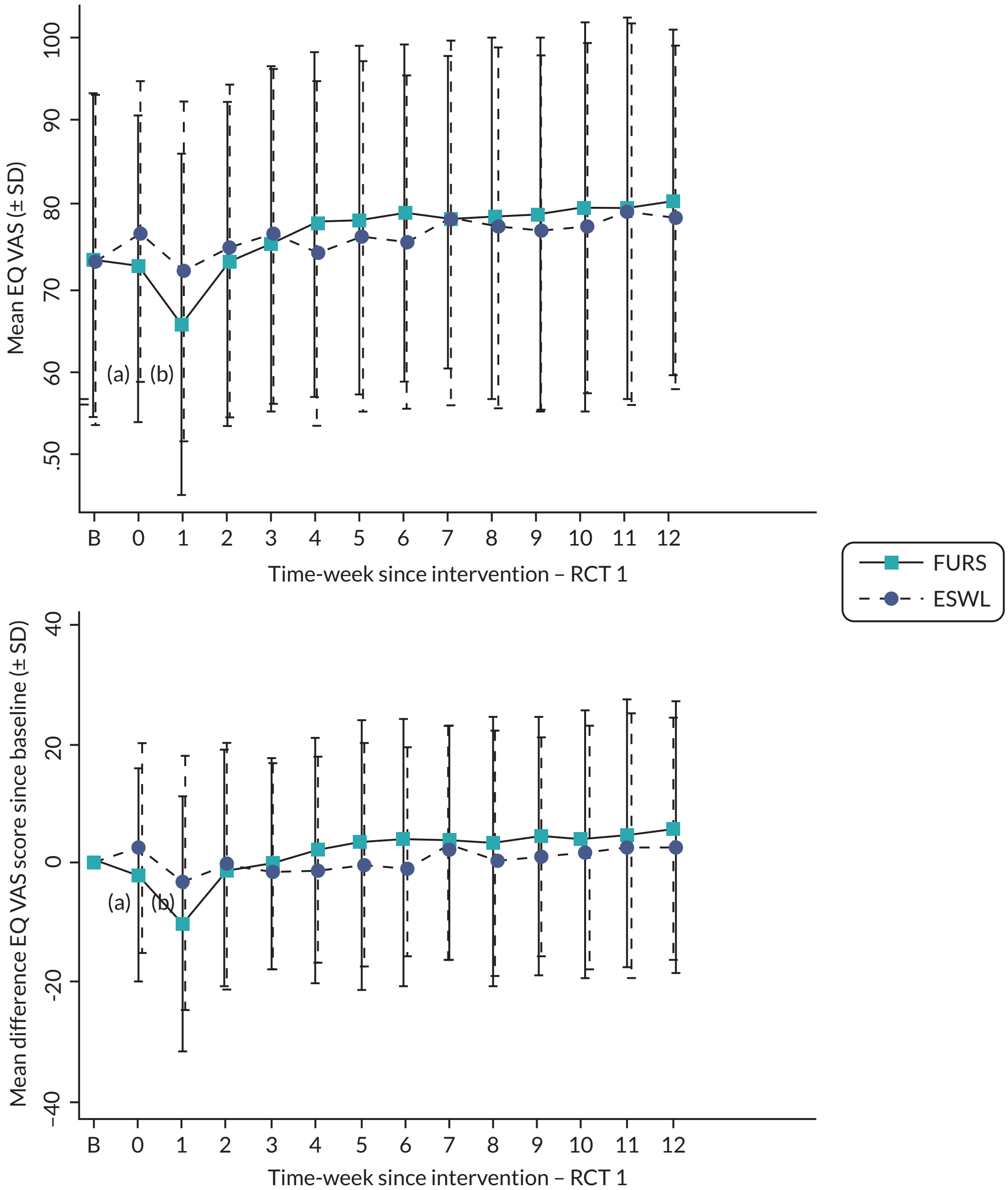


FIGURE 18 RCT1: health status plots over time – EQ-5D-VAS. Note: Time points B and 0 are baseline and pre intervention. Areas (a) and (b) do not necessarily represent 1 week from week 1 to 12, all are 1-week intervals.

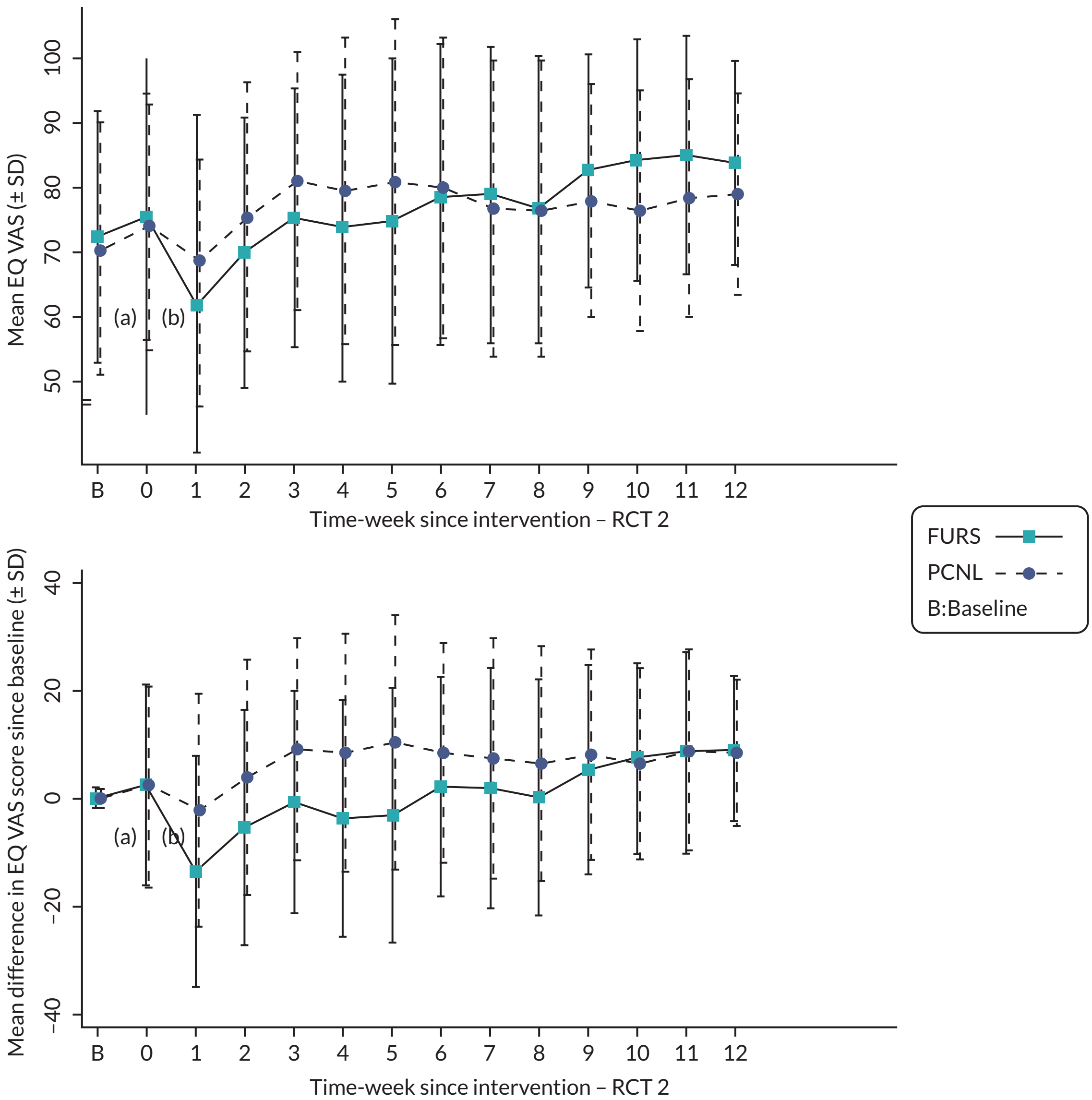


FIGURE 19 RCT2: health status plots over time – EQ-5D-VAS. Time points B and 0 are baseline and pre intervention. Areas (a) and (b) do not necessarily represent 1 week from week 1 to 12, all are 1-week intervals.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Appendix 7 Additional results tables and subgroup analyses

RCT1: per-protocol analyses

TABLE 41 RCT1: per-protocol analyses

Per protocol	FURS *N* = 189 Mean (SD): *n* = 150	ESWL *N* = 208 Mean (SD): *n* = 179	Effect Estimate	95% CI	*p*-value
Primary: main outcome					
EQ-5D score – weekly average AUC[a]	0.815 (0.199)	0.824 (0.210)	0.03	(0.00 to 0.06)	0.035
Adjusting covariate: baseline EQ-5D score	0.778 (0.217)	0.828 (0.193)			
Primary: sensitivity 1					
EQ-5D score – weekly average AUC[a]	0.815 (0.199)	0.824 (0.210)	0.05	(0.02 to 0.08)	0.001
Adjusting covariate: baseline EQ-5D score	0.778 (0.217)	0.828 (0.193)			
Adjusting covariate: waiting times to intervention in weeks	18.641 (15.142)	9.383 (9.313)			
EQ-5D score pre intervention	0.773 (0.210)	0.817 (0.209)			
Primary: sensitivity 2					
EQ-5D score – weekly average AUC[a]	0.815 (0.199)	0.824 (0.210)	0.03	(−0.01 to 0.06)	0.107
Adjusting covariate: Pre-intervention EQ-5D	0.773 (0.210)	0.817 (0.209)			
Secondary: AUC from baseline – 12 weeks					
EQ-5D score – weekly average AUC[a]	0.790 (0.197)	0.824 (0.194)	0.01	(−0.01 to 0.03)	0.572
Adjusting covariate: baseline EQ-5D score	0.778 (0.217)	0.828 (0.193)			
Stone clearance by 12 weeks	*x* (%): *n* = 187	*x* (%): *n* = 207			
Complete	152 (81.3)	74 (35.7)	7.31	(4.24 to 12.58)	< 0.001
Acceptable	8 (4.3)	27 (13.0)			
Unacceptable	27 (14.4)	106 (51.2)			
Had additional treatment by 12 weeks	*x* (%): *n* = 189	*x* (%): *n* = 208			
Yes	16 (8.5)	59 (28.4)	0.30	(0.14 to 0.64)	0.002
No	173 (91.5)	149 (71.6)			

a Time period (number of weeks) is 12, that is from intervention to 12 weeks post intervention.
b The time period (number of weeks) varies for each participant, because the time period is from randomisation/baseline to 12 weeks post intervention and the number of weeks from randomisation/baseline to intervention varies.

Note
Weekly average AUC is the area under the curve of the EQ-5D scores over time (in number of weeks). Effect estimate: positive favours FURS.

RCT1: multiple imputation analyses

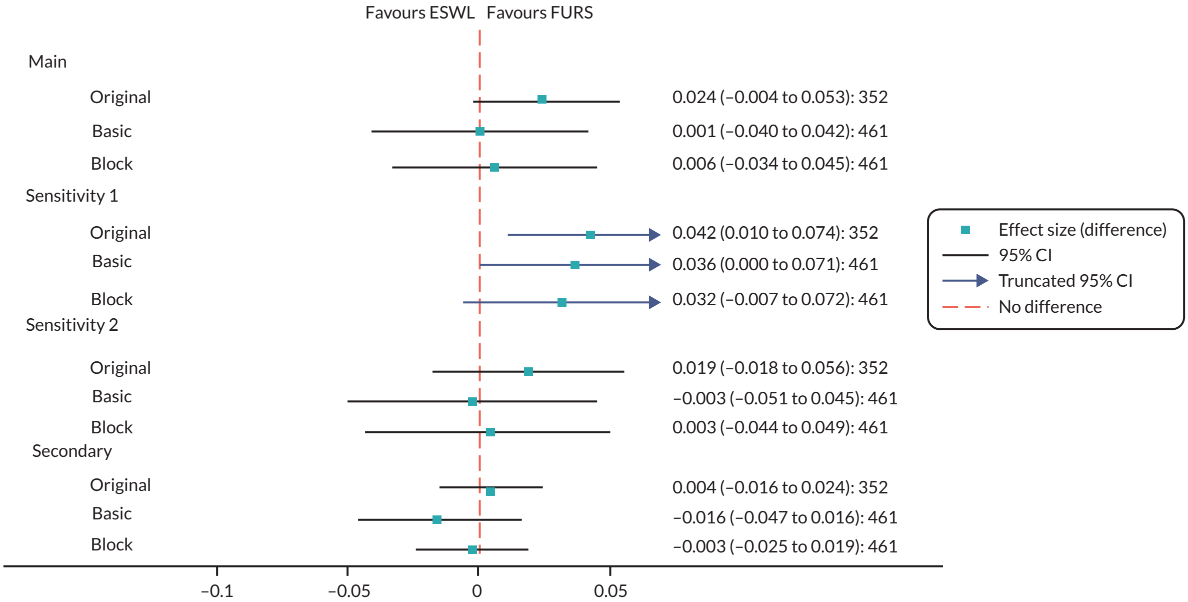


FIGURE 20 RCT1 PO – multiple imputation (MI) difference between FURS and ESWL.

RCT1: pain, health status, quality of life; pre-intervention to 12-week post-intervention and 12-month post-randomisation summaries

TABLE 42 RCT1: pre-intervention summaries for pain, health status and QoL

Variable	FURS *N* = 230	ESWL *N* = 231
Level of pain today[a,b]	*n* = 173	*n* = 185
Mean (SD)	2.5 (2.5)	2.0 (2.5)
Median (IQI)	2.0 (0.0–5.0)	1.0 (0.0–3.0)
Minimum, maximum	(0, 9)	(0, 10)
No. of weekly forms: 1	173	185
Stone pain in last 7 days[c]	*n* = 171	*n* = 184
Yes	100 (57.1)	87 (46.8)
No	71 (40.6)	97 (52.2)
No. of weekly forms: 1	171	184
missing	4 (2.3)	2 (1.1)
No. days needed pain relief in last 7 days[a]	*n* = 169	*n* = 181
Mean (SD)	2.5 (2.7)	1.7 (2.5)
Median (IQI)	1.0 (0.0–5.0)	0.0 (0.0–3.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	169	181

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 42 RCT1: pre-intervention summaries for pain, health status and QoL (*continued*)

Variable	FURS *N* = 230	ESWL *N* = 231
EQ-5D score[a,d]	*n* = 174	*n* = 184
Mean (SD)	0.759 (0.229)	0.802 (0.224)
Median (IQI)	0.806 (0.653–0.985)	0.868 (0.724–0.986)
Minimum, maximum	(–0.141, 0.989)	(–0.111, 0.989)
No. of weekly forms: 1	174	184
EQ-5D-VAS[a,e]	*n* = 174	*n* = 185
Mean (SD)	72.8 (17.8)	76.7 (17.1)
Median (IQI)	77.0 (60.0–86.0)	80.0 (70.0–90.0)
Minimum, maximum	(20, 100)	(10, 100)
No. of weekly forms: 1	174	185
SF-12 scores[a,f]		
Physical component score	*n* = 174	*n* = 185
Mean (SD)	45.7 (11.2)	47.9 (9.6)
Median (IQI)	47.9 (37.2–54.5)	50.1 (43.3–55.0)
Minimum, maximum	(18.2, 68.5)	(20.0, 62.3)
No. of weekly forms: 1	160	176
Mental component score	*n* = 160	*n* = 176
Mean (SD)	47.8 (10.8)	49.0 (10.4)
Median (IQI)	49.2 (41.2–56.3)	51.3 (42.2–57.4)
Minimum, maximum	(18.7, 68.1)	(17.5, 66.5)
No. of weekly forms: 1	160	176

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).
f SF-12 scores range from 0 to 100 (higher score is better).

TABLE 43 RCT1: week 1 post-intervention summaries for pain and health status

Variable	FURS *N* = 230	ESWL *N* = 231
Level of pain today[a,b]	*n* = 152	*n* = 183
Mean (SD)	3.8 (2.7)	2.1 (2.5)
Median (IQI)	4.0 (2.0–5.5)	1.0 (0.0–4.0)
Minimum, maximum	(0, 10)	(0, 9)
No. of weekly forms: 1	143	179
2	9	4

TABLE 43 RCT1: week 1 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 230	ESWL *N* = 231
Worst level of pain since intervention[a]	*n* = 151	*n* = 184
Mean (SD)	7.2 (2.7)	3.9 (3.1)
Median (IQI)	8.0 (6.0–9.0)	4.0 (1.0–7.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	142	182
2	9	2
Stone pain in last 7 days[c]	*n* = 150	*n* = 183
Yes	124 (81.0)	116 (63.0)
No	26 (17.0)	67 (36.4)
No. of weekly forms: 1	142	180
2	8	3
Missing	3 (2.0)	1 (0.5)
No. days needed pain relief in last 7 days[a]	*n* = 150	*n* = 183
Mean (SD)	5.1 (2.3)	2.3 (2.5)
Median (IQI)	6.0 (4.0–7.0)	2.0 (0.0–4.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	141	179
2	9	4
EQ-5D score[a,d]	*n* = 149	*n* = 180
Mean (SD)	0.633 (0.250)	0.784 (0.222)
Median (IQI)	0.687 (0.541–0.804)	0.825 (0.699–0.985)
Minimum, maximum	(−0.138, 0.989)	(−0.187, 0.989)
No. of weekly forms: 1	136	173
2	13	7
EQ-5D-VAS[a,e]	*n* = 146	*n* = 180
Mean (SD)	65.7 (20.1)	72.0 (20.5)
Median (IQI)	70.0 (50.0–80.0)	79.0 (60.0–89.0)
Minimum, maximum	(10, 100)	(10, 100)
No. of weekly forms: 1	132	171
2	14	9

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

Copyright © 2025 Wiseman *et al*. This work was produced by Wiseman *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 44 RCT1: week 2 post-intervention summaries for pain and health status

Variable	FURS *N* = 230	ESWL *N* = 231
Level of pain today[a,b]	*n* = 153	*n* = 189
Mean (SD)	1.9 (2.2)	1.7 (2.3)
Median (IQI)	1.0 (0.0–3.0)	0.0 (0.0–3.0)
Minimum, maximum	(0, 8)	(0, 10)
No. of weekly forms: 1	146	184
2	7	5
Worst level of pain since intervention[a]	*n* = 152	*n* = 189
Mean (SD)	6.9 (2.9)	3.7 (3.2)
Median (IQI)	8.0 (5.0–9.0)	3.0 (1.0–6.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	147	186
2	5	3
Stone pain in last 7 days[c]	*n* = 151	*n* = 185
Yes	88 (56.8)	90 (47.4)
No	63 (40.6)	95 (50.0)
No. of weekly forms: 1	144	180
2	7	5
Missing	4 (2.6)	5 (2.6)
No. days needed pain relief in last 7 days[a]	*n* = 150	*n* = 188
Mean (SD)	2.8 (2.6)	1.5 (2.2)
Median (IQI)	2.0 (0.0–5.0)	0.0 (0.0–2.8)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	143	183
2	7	5
EQ-5D score[a,d]	*n* = 152	*n* = 185
Mean (SD)	0.768 (0.216)	0.811 (0.218)
Median (IQI)	0.806 (0.692–0.985)	0.864 (0.741–0.987)
Minimum, maximum	(–0.241, 0.989)	(–0.092, 0.989)
No. of weekly forms: 1	142	180
2	9	5
3	1	
EQ-5D-VAS[a,e]	*n* = 146	*n* = 182
Mean (SD)	73.1 (19.8)	74.8 (20.3)
Median (IQI)	79.5 (60.0–90.0)	80.0 (69.0–90.0)
Minimum, maximum	(10, 100)	(10, 100)

TABLE 44 RCT1: week 2 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 230	ESWL *N* = 231
No. of weekly forms: 1	136	173
2	9	9
3	1	

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

TABLE 45 RCT1: week 3 post-intervention summaries for pain and health status

Variable	FURS *N* = 230	ESWL *N* = 231
Level of pain today[a,b]	*n* = 154	*n* = 169
Mean (SD)	1.5 (2.2)	1.5 (2.3)
Median (IQI)	0.0 (0.0–2.0)	0.0 (0.0–3.0)
Minimum, maximum	(0, 9)	(0, 10)
No. of weekly forms: 1	152	164
2	2	5
Worst level of pain since intervention[a]	*n* = 153	*n* = 169
Mean (SD)	6.1 (3.5)	3.6 (3.2)
Median (IQI)	7.0 (4.0–9.0)	3.0 (0.0–7.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	152	165
2	1	4
Stone pain in last 7 days[c]	*n* = 153	*n* = 166
Yes	54 (34.6)	73 (42.9)
No	99 (63.5)	93 (54.7)
No. of weekly forms: 1	151	161
2	2	5
Missing	3 (1.9)	4 (2.4)
No. days needed pain relief in last 7 days[a]	*n* = 153	*n* = 167
Mean (SD)	1.9 (2.6)	1.4 (2.2)
Median (IQI)	0.0 (0.0–3.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	151	162
2	2	5

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 45 RCT1: week 3 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 230	ESWL *N* = 231
EQ-5D score[a,d]	*n* = 149	*n* = 158
Mean (SD)	0.797 (0.245)	0.814 (0.219)
Median (IQI)	0.868 (0.697–0.987)	0.864 (0.748–0.987)
Minimum, maximum	(–0.234, 0.989)	(–0.237, 0.989)
No. of weekly forms: 1	143	153
2	6	4
3		1
EQ-5D-VAS[a,e]	*n* = 140	*n* = 158
Mean (SD)	75.2 (20.0)	76.2 (20.1)
Median (IQI)	80.0 (69.0–90.0)	80.0 (70.0–90.0)
Minimum, maximum	(19, 100)	(10, 100)
No. of weekly forms: 1	134	152
2	6	4
3		2

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

TABLE 46 RCT1: week 4 post-intervention summaries for pain and health status

Variable	FURS *N* = 230	ESWL *N* = 231
Level of pain today[a,b]	*n* = 158	*n* = 169
Mean (SD)	1.3 (2.0)	1.6 (2.3)
Median (IQI)	0.0 (0.0–2.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	157	158
2	1	10
3		1
Worst level of pain since intervention[a]	*n* = 158	*n* = 166
Mean (SD)	6.1 (3.5)	4.1 (3.3)
Median (IQI)	7.0 (3.0–9.0)	4.0 (1.0–7.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	157	156
2	1	9
3		1

TABLE 46 RCT1: week 4 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 230	ESWL *N* = 231
Stone pain in last 7 days[c]	*n* = 153	*n* = 166
Yes	46 (28.9)	72 (41.9)
No	107 (67.3)	94 (54.7)
No. of weekly forms: 1	152	156
2	1	9
3		1
Missing	6 (3.8)	6 (3.5)
No. days needed pain relief in last 7 days[a]	*n* = 155	*n* = 168
Mean (SD)	1.8 (2.6)	1.3 (2.1)
Median (IQI)	0.0 (0.0–3.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	154	158
2	1	9
3		1
EQ-5D score[a,d]	*n* = 147	*n* = 165
Mean (SD)	0.809 (0.253)	0.813 (0.233)
Median (IQI)	0.889 (0.726–0.987)	0.868 (0.750–0.987)
Minimum, maximum	(−0.240, 0.989)	(−0.237, 0.989)
No. of weekly forms: 1	144	160
2	3	5
EQ-5D-VAS[a,e]	*n* = 140	*n* = 163
Mean (SD)	77.8 (20.3)	74.2 (22.4)
Median (IQI)	84.8 (70.0–93.5)	80.0 (65.0–90.0)
Minimum, maximum	(10, 100)	(10, 100)
No. of weekly forms: 1	137	153
2	3	8
3		2

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 47 RCT1: week 5 post-intervention summaries for pain and health status

Variable	FURS *N* = 230	ESWL *N* = 231
Level of pain today[a,b]	*n* = 156	*n* = 174
Mean (SD)	1.2 (2.0)	1.3 (2.2)
Median (IQI)	0.0 (0.0–2.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 8)	(0, 10)
No. of weekly forms: 1	155	166
2	1	8
Worst level of pain since intervention[a]	*n* = 155	*n* = 174
Mean (SD)	5.9 (3.5)	3.9 (3.3)
Median (IQI)	7.0 (3.0–9.0)	4.0 (0.0–7.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	154	167
2	1	7
Stone pain in last 7 days[c]	*n* = 152	*n* = 172
Yes	41 (26.3)	61 (35.1)
No	111 (71.2)	111 (63.8)
No. of weekly forms: 1	151	167
2	1	5
Missing	4 (2.6)	2 (1.1)
No. days needed pain relief in last 7 days[a]	*n* = 154	*n* = 173
Mean (SD)	1.6 (2.6)	1.1 (2.0)
Median (IQI)	0.0 (0.0–3.0)	0.0 (0.0–1.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	153	166
2	1	7
EQ-5D score[a,d]	*n* = 149	*n* = 158
Mean (SD)	0.814 (0.239)	0.837 (0.209)
Median (IQI)	0.891 (0.729–0.987)	0.893 (0.754–0.987)
Minimum, maximum	(–0.191, 0.989)	(–0.117, 0.989)
No. of weekly forms: 1	147	152
2	2	5
3		1
EQ-5D-VAS[a,e]	*n* = 140	*n* = 156
Mean (SD)	78.1 (20.9)	76.2 (20.9)
Median (IQI)	85.0 (70.0–93.5)	80.0 (70.0–90.0)
Minimum, maximum	(10, 100)	(10, 100)

TABLE 47 RCT1: week 5 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 230	ESWL *N* = 231
No. of weekly forms: 1	138	146
2	2	9
3		1

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

TABLE 48 RCT1: week 6 post-intervention summaries for pain and health status

Variable	FURS *N* = 230	ESWL *N* = 231
Level of pain today[a,b]	*n* = 161	*n* = 172
Mean (SD)	1.1 (1.9)	1.2 (2.1)
Median (IQI)	0.0 (0.0–2.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 8)	(0, 10)
No. of weekly forms: 1	159	162
2	2	10
Worst level of pain since intervention[a]	*n* = 160	*n* = 171
Mean (SD)	6.0 (3.5)	4.1 (3.5)
Median (IQI)	7.0 (3.5–9.0)	4.0 (0.0–7.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	160	165
2		6
Stone pain in last 7 days[c]	*n* = 158	*n* = 167
Yes	41 (25.5)	61 (35.5)
No	117 (72.7)	106 (61.6)
No. of weekly forms: 1	156	157
2	2	10
Missing	3 (1.9)	5 (2.9)
No. days needed pain relief in last 7 days[a]	*n* = 156	*n* = 171
Mean (SD)	1.4 (2.3)	1.2 (2.0)
Median (IQI)	0.0 (0.0–2.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	154	161
2	2	10

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 48 RCT1: week 6 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 230	ESWL *N* = 231
EQ-5D score[a,d]	*n* = 152	*n* = 160
Mean (SD)	0.832 (0.226)	0.837 (0.207)
Median (IQI)	0.909 (0.751–0.987)	0.893 (0.754–0.987)
Minimum, maximum	(−0.130, 0.989)	(−0.117, 0.989)
No. of weekly forms: 1	150	155
2	2	5
EQ-5D-VAS[a,e]	*n* = 144	*n* = 158
Mean (SD)	78.9 (20.0)	75.5 (21.2)
Median (IQI)	85.0 (70.0–95.0)	80.0 (70.0–90.0)
Minimum, maximum	(20, 100)	(10, 100)
No. of weekly forms: 1	142	148
2	2	10

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

TABLE 49 RCT1: week 7 post-intervention summaries for pain and health status

Variable	FURS *N* = 230	ESWL *N* = 231
Level of pain today[a,b]	*n* = 153	*n* = 169
Mean (SD)	1.1 (2.1)	1.2 (2.0)
Median (IQI)	0.0 (0.0–1.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 8)	(0, 10)
No. of weekly forms: 1	153	167
2		2
Worst level of pain since intervention[a]	*n* = 154	*n* = 169
Mean (SD)	6.0 (3.4)	4.1 (3.5)
Median (IQI)	7.0 (4.0–9.0)	4.0 (0.0–7.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	154	168
2		1
Stone pain in last 7 days[c]	*n* = 147	*n* = 163
Yes	31 (20.1)	53 (31.4)
No	116 (75.3)	110 (65.1)

TABLE 49 RCT1: week 7 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 230	ESWL *N* = 231
No. of weekly forms: 1	147	161
2		2
Missing	7 (4.5)	6 (3.6)
No. days needed pain relief in last 7 days[a]	*n* = 151	*n* = 167
Mean (SD)	1.2 (2.3)	1.1 (1.9)
Median (IQI)	0.0 (0.0–2.0)	0.0 (0.0–1.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	151	165
2		2
EQ-5D score[a,d]	*n* = 149	*n* = 160
Mean (SD)	0.837 (0.230)	0.831 (0.229)
Median (IQI)	0.985 (0.786–0.987)	0.921 (0.751–0.987)
Minimum, maximum	(−0.130, 0.989)	(−0.522, 0.989)
No. of weekly forms: 1	148	158
2	1	2
EQ-5D-VAS[a,e]	*n* = 141	*n* = 156
Mean (SD)	78.1 (21.8)	78.0 (19.1)
Median (IQI)	85.0 (69.0–95.0)	80.0 (70.0–90.0)
Minimum, maximum	(10, 100)	(10, 100)
No. of weekly forms: 1	140	154
2	1	2

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

TABLE 50 RCT1: week 8 post-intervention summaries for pain and health status

Variable	FURS *N* = 230	ESWL *N* = 231
Level of pain today[a,b]	*n* = 153	*n* = 163
Mean (SD)	0.9 (1.9)	1.2 (2.2)
Median (IQI)	0.0 (0.0–1.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 9)	(0, 10)
No. of weekly forms: 1	152	154
2	1	9

continued

Copyright © 2025 Wiseman *et al*. This work was produced by Wiseman *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 50 RCT1: week 8 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 230	ESWL *N* = 231
Worst level of pain since intervention[a]	*n* = 152	*n* = 162
Mean (SD)	5.8 (3.6)	4.3 (3.5)
Median (IQI)	7.0 (3.0–9.0)	4.0 (1.0–8.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	151	154
2	1	8
Stone pain in last 7 days[c]	*n* = 150	*n* = 161
Yes	27 (17.5)	54 (32.9)
No	123 (79.9)	107 (65.2)
No. of weekly forms: 1	149	153
2	1	8
Missing	4 (2.6)	3 (1.8)
No. days needed pain relief in last 7 days[a]	*n* = 150	*n* = 162
Mean (SD)	1.1 (2.1)	1.1 (1.9)
Median (IQI)	0.0 (0.0–1.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	149	154
2	1	8
EQ-5D score[a,d]	*n* = 151	*n* = 156
Mean (SD)	0.836 (0.224)	0.836 (0.225)
Median (IQI)	0.893 (0.776–0.987)	0.920 (0.753–0.987)
Minimum, maximum	(–0.130, 0.989)	(–0.522, 0.989)
No. of weekly forms: 1	148	152
2	3	4
EQ-5D-VAS[a,e]	*n* = 142	*n* = 151
Mean (SD)	78.5 (21.7)	77.3 (20.9)
Median (IQI)	89.0 (70.0–95.0)	84.0 (68.0–90.0)
Minimum, maximum	(10, 100)	(10, 100)
No. of weekly forms: 1	137	141
2	5	10

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

TABLE 51 RCT1: week 9 post-intervention summaries for pain and health status

Variable	FURS *N* = 230	ESWL *N* = 231
Level of pain today[a,b]	*n* = 151	*n* = 168
Mean (SD)	0.9 (1.9)	1.1 (1.8)
Median (IQI)	0.0 (0.0–1.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 8)	(0, 10)
No. of weekly forms: 1	151	162
2		6
Worst level of pain since intervention[a]	*n* = 150	*n* = 167
Mean (SD)	5.9 (3.5)	4.1 (3.5)
Median (IQI)	7.0 (3.0–9.0)	4.0 (0.0–7.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	150	161
2		6
Stone pain in last 7 days[c]	*n* = 147	*n* = 162
Yes	29 (19.2)	49 (29.0)
No	118 (78.1)	113 (66.9)
No. of weekly forms: 1	147	156
2		6
Missing	4 (2.6)	7 (4.1)
No. days needed pain relief in last 7 days[a]	*n* = 144	*n* = 167
Mean (SD)	1.3 (2.3)	1.1 (2.0)
Median (IQI)	0.0 (0.0–1.5)	0.0 (0.0–2.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	144	161
2		6
EQ-5D score[a,d]	*n* = 145	*n* = 160
Mean (SD)	0.837 (0.233)	0.828 (0.232)
Median (IQI)	0.985 (0.759–0.987)	0.893 (0.740–0.987)
Minimum, maximum	(−0.284, 0.989)	(−0.522, 0.989)
No. of weekly forms: 1	143	156
2	2	4
EQ-5D-VAS[a,e]	*n* = 138	*n* = 158
Mean (SD)	78.9 (21.2)	76.8 (21.1)
Median (IQI)	86.0 (70.0–96.0)	80.0 (70.0–90.0)
Minimum, maximum	(10, 100)	(10, 100)

continued

Copyright © 2025 Wiseman *et al*. This work was produced by Wiseman *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 51 RCT1: week 9 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 230	ESWL *N* = 231
No. of weekly forms: 1	136	151
2	2	7

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

TABLE 52 RCT1: week 10 post-intervention summaries for pain and health status

Variable	FURS *N* = 230	ESWL *N* = 231
Level of pain today[a,b]	*n* = 144	*n* = 170
Mean (SD)	1.0 (1.9)	1.2 (2.2)
Median (IQI)	0.0 (0.0–1.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 9)	(0, 9)
No. of weekly forms: 1	144	167
2		3
Worst level of pain since intervention[a]	*n* = 143	*n* = 169
Mean (SD)	5.9 (3.4)	4.2 (3.6)
Median (IQI)	7.0 (3.0–9.0)	4.0 (0.0–8.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	143	166
2		3
Stone pain in last 7 days[c]	*n* = 137	*n* = 165
Yes	28 (19.3)	45 (26.5)
No	109 (75.2)	120 (70.6)
No. of weekly forms: 1	137	163
2		2
Missing	8 (5.5)	5 (2.9)
No. days needed pain relief in last 7 days[a]	*n* = 139	*n* = 169
Mean (SD)	1.2 (2.2)	1.0 (2.0)
Median (IQI)	0.0 (0.0–1.0)	0.0 (0.0–1.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	139	166
2		3

TABLE 52 RCT1: week 10 post-intervention summaries for pain and health status (*continued*)

Variable	FURS N = 230	ESWL N = 231
EQ-5D score[a,d]	*n* = 142	*n* = 166
Mean (SD)	0.845 (0.233)	0.833 (0.230)
Median (IQI)	0.985 (0.786–0.987)	0.965 (0.748–0.987)
Minimum, maximum	(−0.284, 0.989)	(−0.223, 0.989)
No. of weekly forms: 1	141	165
2	1	1
EQ5D VAS[a,e]	*n* = 132	*n* = 163
Mean (SD)	79.7 (22.0)	77.4 (20.0)
Median (IQI)	89.0 (70.0–97.0)	80.0 (70.0–90.0)
Minimum, maximum	(10, 100)	(10, 100)
No. of weekly forms: 1	131	159
2	1	4

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

TABLE 53 RCT1: week 11 post-intervention summaries for pain and health status

Variable	FURS N = 230	ESWL N = 231
Level of pain today[a,b]	*n* = 145	*n* = 168
Mean (SD)	0.8 (1.8)	1.1 (2.1)
Median (IQI)	0.0 (0.0–1.0)	0.0 (0.0–1.0)
Minimum, maximum	(0, 9)	(0, 9)
No. of weekly forms: 1	145	166
2		2
Worst level of pain since intervention[a]	*n* = 145	*n* = 167
Mean (SD)	5.6 (3.7)	4.3 (3.5)
Median (IQI)	7.0 (2.0–9.0)	4.0 (0.0–8.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	145	165
2		2
Stone pain in last 7 days[c]	*n* = 134	*n* = 166
Yes	28 (19.0)	45 (26.8)
Noo	106 (72.1)	121 (72.0)
No. of weekly forms: 1	134	164
2		2
Missing	13 (8.8)	2 (1.2)

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 53 RCT1: week 11 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 230	ESWL *N* = 231
No. days needed pain relief in last 7 days[a]	*n* = 143	*n* = 168
Mean (SD)	1.2 (2.2)	1.0 (1.9)
Median (IQI)	0.0 (0.0–1.0)	0.0 (0.0–1.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	143	166
2		2
EQ-5D score[a,d]	*n* = 143	*n* = 162
Mean (SD)	0.838 (0.241)	0.840 (0.233)
Median (IQI)	0.985 (0.786–0.987)	0.985 (0.753–0.987)
Minimum, maximum	(–0.284, 0.989)	(–0.184, 0.989)
No. of weekly forms: 1	143	156
2		6
EQ-5D-VAS[a,e]	*n* = 135	*n* = 160
Mean (SD)	79.6 (22.7)	78.9 (19.7)
Median (IQI)	90.0 (70.0–97.0)	85.0 (70.0–90.0)
Minimum, maximum	(10, 100)	(10, 100)
No. of weekly forms: 1	135	154
2		6

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

TABLE 54 RCT1: week 12 post-intervention summaries for pain, health status and QoL

Variable	FURS *N* = 230	ESWL *N* = 231
Level of pain today[a,b]	*n* = 155	*n* = 169
Mean (SD)	0.9 (1.9)	1.0 (1.9)
Median (IQI)	0.0 (0.0–1.0)	0.0 (0.0–1.0)
Minimum, maximum	(0, 10)	(0, 9)
No. of weekly forms: 1	155	165
2		4
Stone pain in last 7 days[c]	*n* = 151	*n* = 167
Yes	31 (19.7)	54 (31.6)
No	120 (76.4)	113 (66.1)

TABLE 54 RCT1: week 12 post-intervention summaries for pain, health status and QoL (*continued*)

Variable	FURS *N* = 230	ESWL *N* = 231
No. of weekly forms: 1	151	163
2		4
Missing	6 (3.8)	4 (2.3)
No. days needed pain relief in last 7 days[a]	*n* = 148	*n* = 163
Mean (SD)	1.2 (2.1)	1.0 (1.9)
Median (IQI)	0.0 (0.0–2.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	148	159
2		4
EQ-5D score[a,d]	*n* = 154	*n* = 167
Mean (SD)	0.837 (0.246)	0.832 (0.239)
Median (IQI)	0.985 (0.776–0.987)	0.985 (0.750–0.987)
Minimum, maximum	(–0.442, 0.989)	(–0.087, 0.989)
No. of weekly forms: 1	153	166
2	1	1
EQ-5D-VAS[a,e]	*n* = 144	*n* = 165
Mean (SD)	80.4 (20.6)	78.6 (20.2)
Median (IQI)	85.0 (70.0–97.5)	85.0 (70.0–90.0)
Minimum, maximum	(10, 100)	(10, 100)
No. of weekly forms: 1	142	160
2	2	5
SF-12 Scores[a,f]		
Physical component score	*n* = 144	*n* = 165
Mean (SD)	47.8 (11.6)	48.4 (10.7)
Median (IQI)	52.7 (40.0–56.1)	53.0 (40.9–56.1)
Minimum, maximum	(10.9, 67.4)	(15.2, 65.3)
No. of weekly forms: 1	145	165
Mental component score	*n* = 145	*n* = 165
Mean (SD)	48.6 (11.0)	48.1 (11.5)
Median (IQI)	50.6 (42.1–57.4)	51.8 (40.4–57.2)
Minimum, maximum	(16.9, 63.6)	(7.9, 62.9)
No. of weekly forms: 1	145	165

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD) median, (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).
f SF-12 scores range from 0 to 100 (higher score is better).

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 55 RCT1: 12-month post-randomisation summaries for pain, health status and QoL

Variable	FURS *N* = 230	ESWL *N* = 231
Level of pain today[a,b]	*n* = 146	*n* = 138
Mean (SD)	1.0 (2.1)	1.0 (2.0)
Median (IQI)	0.0 (0.0–0.0)	0.0 (0.0–1.0)
Minimum, maximum	(0, 9)	(0, 10)
No. of weekly forms: 1	146	138
Stone pain in last 7 days[c]	*n* = 142	*n* = 136
Yes	31 (20.7)	37 (26.1)
No	111 (74.0)	99 (69.7)
No. of weekly forms: 1	142	136
Missing	8 (5.3)	6 (4.2)
No. days needed pain relief in last 7 days[a]	*n* = 144	*n* = 135
Mean (SD)	0.8 (2.0)	0.8 (1.8)
Median (IQI)	0.0 (0.0–0.0)	0.0 (0.0–0.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	144	135
EQ-5D score[a,d]	*n* = 146	*n* = 134
Mean (SD)	0.814 (0.236)	0.810 (0.227)
Median (IQI)	0.891 (0.681–0.987)	0.889 (0.720–0.987)
Minimum, maximum	(–0.221, 0.989)	(–0.185, 0.989)
No. of weekly forms: 1	146	134
EQ-5D-VAS[a,e]	*n* = 134	*n* = 127
Mean (SD)	76.5 (21.9)	76.2 (21.2)
Median (IQI)	83.0 (68.0–90.0)	80.0 (70.0–90.0)
Minimum, maximum	(10, 100)	(10, 100)
No. of weekly forms: 1	134	127
SF-12 scores[a,f]		
Physical component score	*n* = 136	*n* = 130
Mean (SD)	47.2 (12.2)	47.6 (10.6)
Median (IQI)	51.7 (39.5–56.3)	51.2 (39.8,–56.1)
Minimum, maximum	(13.5, 68.0)	(24.2, 66.2)
No. of weekly forms: 1	136	130
Mental component score	*n* = 136	*n* = 130
Mean (SD)	49.3 (11.6)	47.9 (11.5)
Median (IQI)	51.4 (41.7–57.9)	50.5 (40.7–57.0)
Minimum, maximum	(11.9, 70.6)	(11.1, 64.2)
No. of weekly forms: 1	136	130

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).
f SF-12 scores range from 0 to 100 (higher score is better).

RCT1: subgroup analyses plots

Subgroup primary and main secondary outcome treatment comparison estimates: plots – RCT1

Graphical representations of the impact of each prespecified subgroup on the treatment effect on the five main outcome measures.

Three PO measures:

- Main PO: The main PO is the AUC of EQ-5D scores from pre intervention to 12 weeks, averaged over 12 weeks and adjusted for 'baseline EQ-5D scores' and where cluster effects are accounted for using a robust variance at centre level.
- Sensitivity 1: The same averaged AUC as above but with covariates of 'baseline EQ-5D scores' and 'waiting times'; clusters adjusted as above.
- Sensitivity 2: The AUC as main PO, adjusted for clusters but without other covariates.

Two secondary outcomes:

- AUC from baseline to 12 weeks – this is an AUC of all EQ-5D scores anchored from baseline right up to 12 weeks post intervention, standardised over the total number of weeks and adjusted for clusters.
- Stone clearance (complete/acceptable/unacceptable) by 12 weeks post intervention also adjusted for clusters.

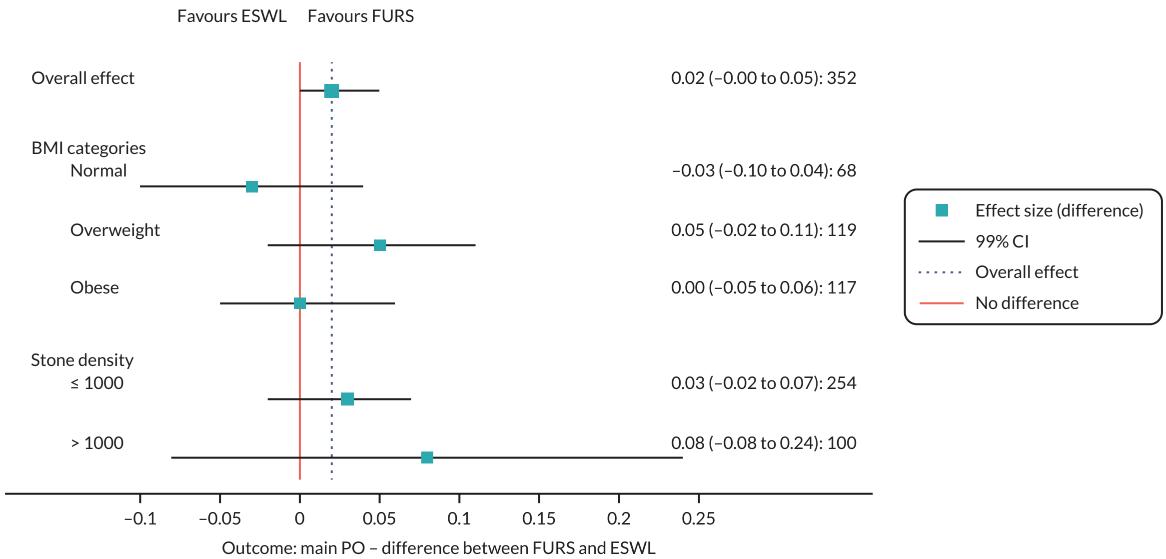


FIGURE 21 RCT1 subgroups – main PO: forest plot showing subgroup model effects.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

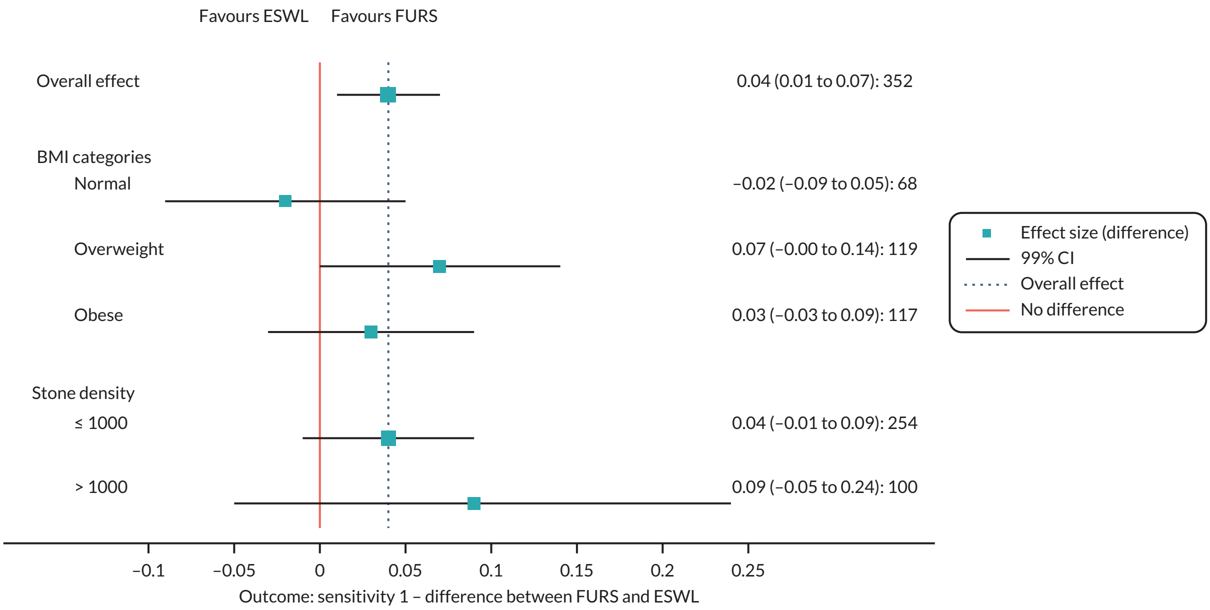


FIGURE 22 RCT1 subgroups – sensitivity 1: forest plot showing subgroup model effects.

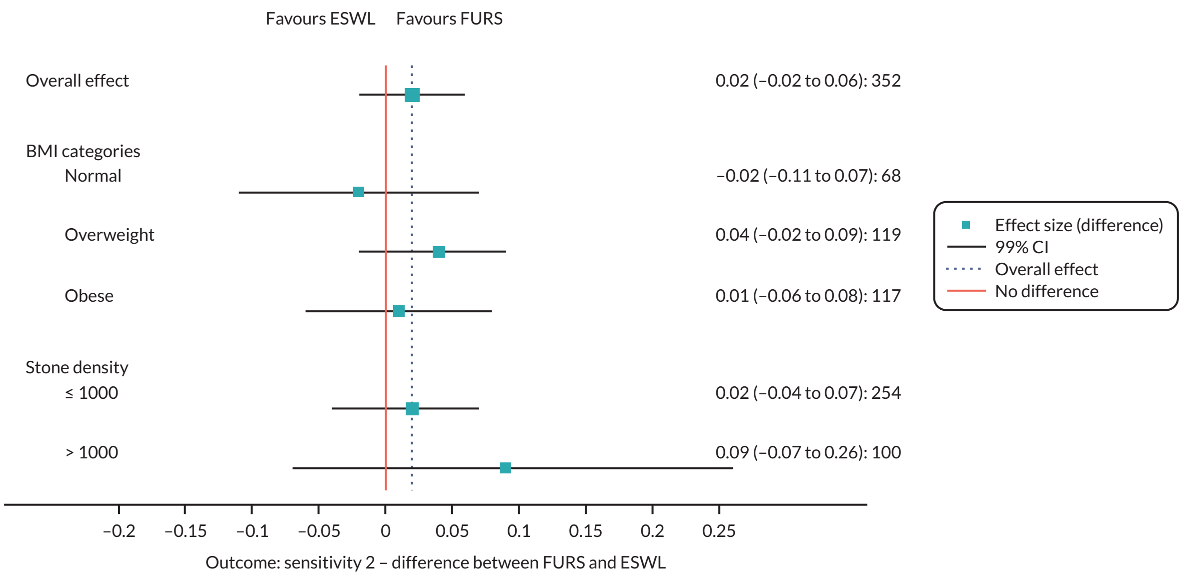


FIGURE 23 RCT1 subgroups – sensitivity 2: forest plot showing subgroup model effects.

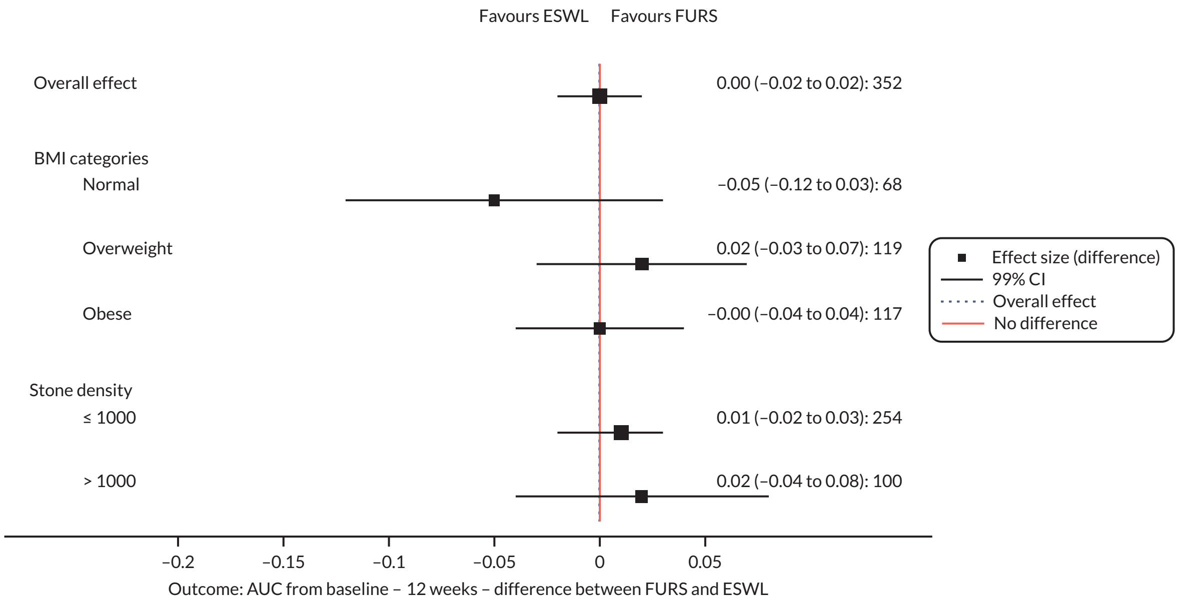


FIGURE 24 RCT1 subgroups – AUC from baseline – 12 weeks: forest plot showing subgroup model effects.

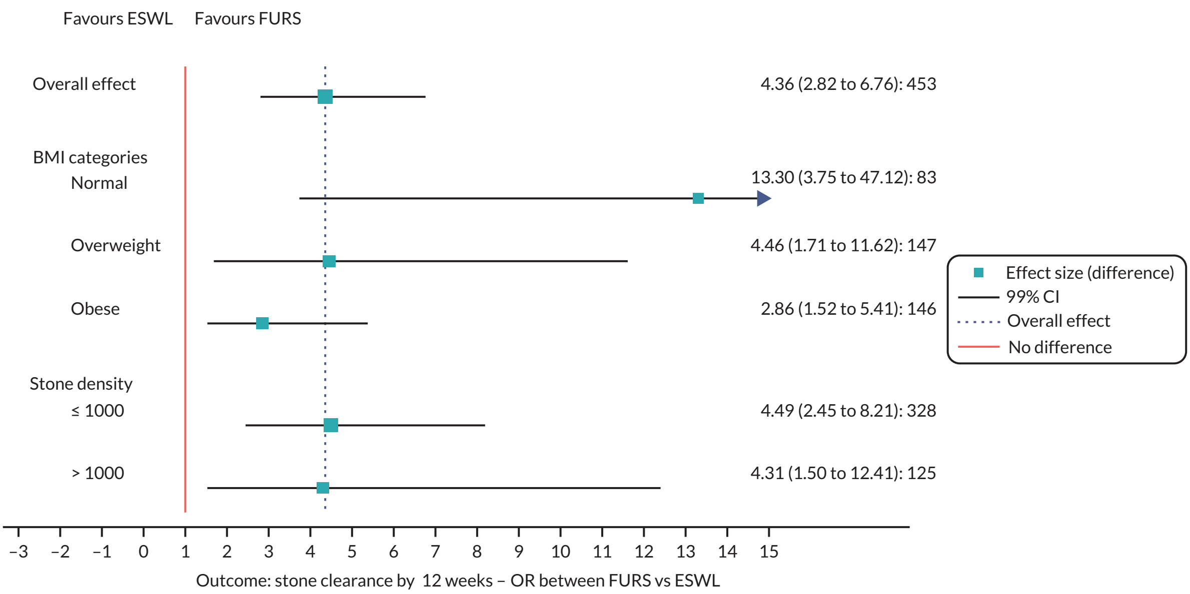


FIGURE 25 RCT1 subgroups – stone clearance at 12 weeks: forest plot showing subgroup model effect.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

RCT2: per-protocol analyses

TABLE 56 RCT2: per-protocol analyses

Per protocol	FURS *N* = 61 Mean (SD): *n* = 49	PCNL *N* = 64 Mean (SD): *n* = 54	Effect estimate	95% CI	*p*-value
Primary: main outcome					
EQ-5D score – weekly average AUC[a]	0.794 (0.209)	0.825 (0.192)	−0.06	(−0.11 to −0.01)	0.025
Adjusting covariate: baseline EQ-5D score	0.820 (0.221)	0.779 (0.229)			
Primary: sensitivity 1					
EQ-5D scores – weekly average AUC[a]	0.794 (0.209)	0.825 (0.192)	−0.05	(−0.10 to −0.01)	0.031
Adjusting covariate: baseline EQ-5D score	0.820 (0.221)	0.779 (0.229)			
Adjusting covariate: waiting times to intervention	21.263 (15.727)	22.396 (15.981)			
EQ-5D score – pre intervention	0.836 (0.144)	0.774 (0.224)			
Primary: sensitivity 2					
EQ-5D score – weekly average AUC[a]	0.794 (0.209)	0.825 (0.192)	−0.07	(−0.13 to −0.00)	0.040
Adjusting covariate: pre-intervention EQ-5D score	0.836 (0.144)	0.774 (0.224)			
Secondary: AUC from baseline to12 weeks					
EQ-5D score – weekly average AUC[b]	0.796 (0.183)	0.783 (0.203)	−0.02	(−0.05 to 0.01)	0.281
Adjusting covariate: baseline EQ-5D score	0.820 (0.221)	0.779 (0.229)			
Stone clearance by 12 weeks	*x* (%): *n* = 59	*x* (%): *n* = 61			
Complete	31 (52.5)	50 (82.0)	0.20	(0.11 to 0.35)	< 0.001
Acceptable	4 (6.8)	8 (13.1)			
Unacceptable	24 (40.7)	3 (4.9)			
Had additional treatment by 12 weeks	*x* (%): *n* = 61	*x* (%): *n* = 64			
Yes	14 (23.0)	2 (3.1)	7.34	(2.37 to 22.75)	0.001
No	47 (77.0)	62 (96.9)			

a Time period (number of weeks) is 12, that is from intervention to 12 weeks post intervention.
b The time period (number of weeks) varies for each participant, because the time period is from randomisation/baseline to 12 weeks post intervention and the number of weeks from randomisation/baseline to intervention varies.

Note
Weekly average AUC is the area under the curve of the EQ-5D scores over time (in number of weeks). Effect estimate: positive favours FURS.

RCT2: multiple imputation analyses

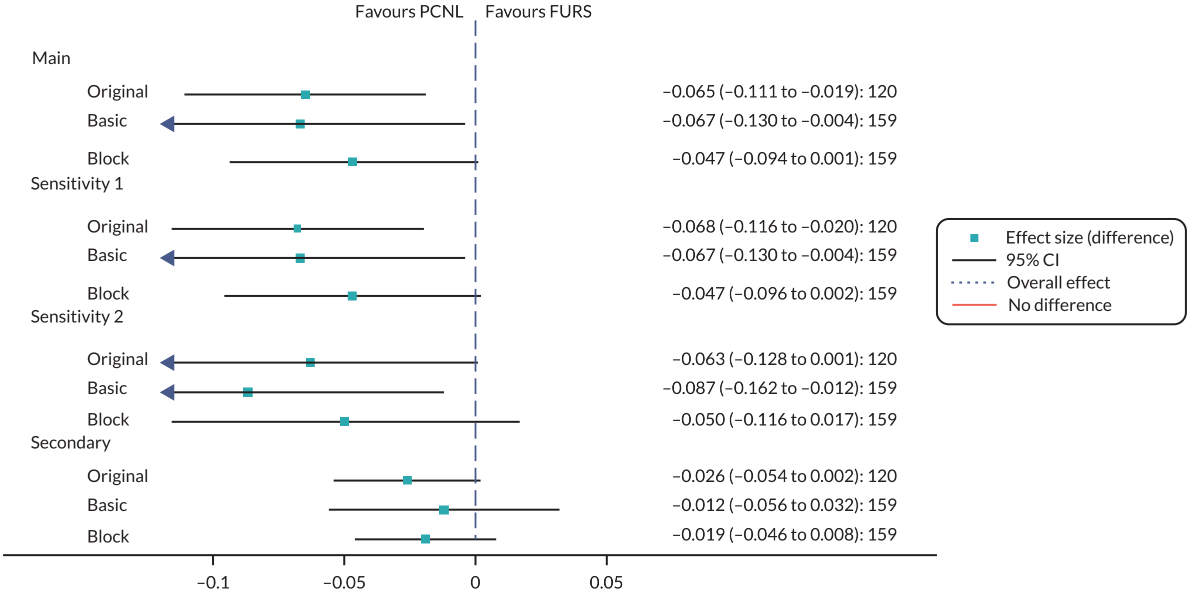


FIGURE 26 RCT2 PO – multiple imputation (MI) difference between FURS and PCNL.

RCT2: pain, health status, quality; pre-intervention to 12-week post-intervention and 12-month post-randomisation summaries

TABLE 57 RCT2: pre-intervention summaries for pain, health status and QoL

Variable	FURS *N* = 73	PCNL *N* = 86
Level of pain today[a,b]	*n* = 60	*n* = 63
Mean (SD)	2.0 (2.7)	2.5 (3.1)
Median (IQI)	1.0 (0.0–4.0)	1.0 (0.0–4.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	60	63
Stone pain in last 7 days[c]	*n* = 59	*n* = 62
Yes	29 (48.3)	34 (54.0)
No	30 (50.0)	28 (44.4)
No. of weekly forms: 1	59	62
Missing	1 (1.7)	1 (1.6)
No. days needed pain relief in last 7 days[a]	*n* = 60	*n* = 61
Mean (SD)	1.8 (2.5)	2.6 (2.9)
Median (IQI)	0.0 (0.0–3.0)	2.0 (0.0–6.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	60	61

continued

Copyright © 2025 Wiseman *et al*. This work was produced by Wiseman *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 57 RCT2: pre-intervention summaries for pain, health status and QoL (*continued*)

Variable	FURS *N* = 73	PCNL *N* = 86
EQ-5D score [a,d]	*n* = 60	*n* = 63
Mean (SD)	0.798 (0.199)	0.760 (0.230)
Median (IQI)	0.850 (0.685–0.986)	0.793 (0.691–0.893)
Minimum, maximum	(–0.110, 0.989)	(–0.104, 0.989)
No. of weekly forms: 1	60	63
EQ-5D-VAS [a,e]	*n* = 60	*n* = 63
Mean (SD)	75.5 (18.9)	74.0 (20.2)
Median (IQI)	78.5 (60.0–90.0)	80.0 (60.0–90.0)
Minimum, maximum	(20, 100)	(20, 100)
No. of weekly forms: 1	60	63
SF-12 scores [a,f]		
Physical component score	*n* = 60	*n* = 63
Mean (SD)	46.1 (10.0)	44.5 (11.3)
Median (IQI)	50.2 (37.0–54.6)	48.0 (37.1–53.4)
Minimum, maximum	(23.8, 61.9)	(15.7, 61.8)
No. of weekly forms: 1	57	59
Mental component score	*n* = 57	*n* = 59
Mean (SD)	49.7 (10.3)	49.7 (10.5)
Median (IQI)	51.8 (42.2–57.8)	50.3 (43.5–57.2)
Minimum, maximum	(20.4, 63.7)	(16.0, 70.9)
No. of weekly forms: 1	57	59

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).
f SF-12 scores range from 0 to 100 (higher score is better).

TABLE 58 RCT2: week 1 post-intervention summaries for pain and health status

Variable	FURS *N* = 73	PCNL *N* = 86
Level of pain today [a,b]	*n* = 52	*n* = 60
Mean (SD)	3.7 (2.6)	3.1 (2.5)
Median (IQI)	3.0 (2.0–5.0)	3.0 (1.0–5.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	49	59
2	3	1

TABLE 58 RCT2: week 1 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 73	PCNL *N* = 86
Worst level of pain since intervention[a]	*n* = 52	*n* = 60
Mean (SD)	7.8 (2.1)	6.0 (3.2)
Median (IQI)	8.0 (7.5–9.0)	7.0 (4.0–8.5)
Minimum, maximum	(2, 10)	(0, 10)
No. of weekly forms: 1	50	59
2	2	1
Stone pain in last 7 days[c]	*n* = 49	*n* = 61
Yes	44 (83.0)	39 (63.9)
No	5 (9.4)	22 (36.1)
No. of weekly forms: 1	46	60
2	3	1
Missing	4 (7.5)	
No. days needed pain relief in last 7 days[a]	*n* = 52	*n* = 60
Mean (SD)	4.8 (2.3)	4.8 (2.5)
Median (IQI)	5.5 (3.0–7.0)	5.5 (3.0–7.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	49	59
2	3	1
EQ-5D score[a,d]	*n* = 52	*n* = 60
Mean (SD)	0.614 (0.289)	0.687 (0.216)
Median (IQI)	0.672 (0.551–0.786)	0.696 (0.584–0.814)
Minimum, maximum	(−0.178, 0.989)	(−0.136, 0.989)
No. of weekly forms: 1	46	58
2	6	2
EQ-5D-VAS[a,e]	*n* = 50	*n* = 54
Mean (SD)	61.8 (22.5)	68.7 (20.9)
Median (IQI)	60.0 (50.0–80.0)	70.0 (50.0–85.0)
Minimum, maximum	(10, 100)	(20, 100)
No. of weekly forms: 1	44	52
2	6	2

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 59 RCT2: week 2 post-intervention summaries for pain and health status

Variable	FURS *N* = 73	PCNL *N* = 86
Level of pain today[a,b]	*n* = 56	*n* = 59
Mean (SD)	2.8 (2.5)	1.8 (2.3)
Median (IQI)	2.0 (0.8–4.5)	1.0 (0.0–2.0)
Minimum, maximum	(0, 9)	(0, 10)
No. of weekly forms: 1	53	55
2	3	4
Worst level of pain since intervention[a]	*n* = 56	*n* = 58
Mean (SD)	7.2 (2.5)	5.0 (3.3)
Median (IQI)	8.0 (7.0–9.0)	6.0 (2.0–8.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	53	55
2	3	3
Stone pain in last 7 days[c]	*n* = 54	*n* = 60
Yes	37 (64.9)	22 (35.5)
No	17 (29.8)	38 (61.3)
No. of weekly forms: 1	51	56
2	3	4
Missing	3 (5.3)	2 (3.2)
No. days needed pain relief in last 7 days[a]	*n* = 56	*n* = 59
Mean (SD)	3.4 (2.7)	2.9 (2.9)
Median (IQI)	2.3 (1.0–6.5)	2.0 (0.0–7.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	53	55
2	3	4
EQ-5D score[a,d]	*n* = 56	*n* = 60
Mean (SD)	0.715 (0.257)	0.783 (0.172)
Median (IQI)	0.753 (0.627–0.872)	0.782 (0.693–0.927)
Minimum, maximum	(–0.179, 0.989)	(0.245, 0.989)
No. of weekly forms: 1	53	57
2	3	3
EQ-5D-VAS[a,e]	*n* = 54	*n* = 57
Mean (SD)	69.9 (20.7)	75.5 (18.4)
Median (IQI)	70.0 (60.0–85.0)	80.0 (60.0–90.0)
Minimum, maximum	(10, 100)	(39, 100)

TABLE 59 RCT2: week 2 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 73	PCNL *N* = 86
No. of weekly forms: 1	51	54
2	3	3

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

TABLE 60 RCT2: week 3 post-intervention summaries for pain and health status

Variable	FURS *N* = 73	PCNL *N* = 86
Level of pain today[a,b]	*n* = 53	*n* = 61
Mean (SD)	1.9 (2.4)	1.4 (1.8)
Median (IQI)	1.0 (0.0–3.0)	1.0 (0.0–2.0)
Minimum, maximum	(0, 8)	(0, 7)
No. of weekly forms: 1	50	60
2	3	1
Worst level of pain since intervention[a]	*n* = 53	*n* = 60
Mean (SD)	6.5 (3.3)	4.7 (3.7)
Median (IQI)	8.0 (4.0–9.0)	6.0 (0.5–8.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	50	59
2	3	1
Stone pain in last 7 days[c]	*n* = 50	*n* = 61
Yes	24 (44.4)	19 (31.1)
No	26 (48.1)	42 (68.9)
No. of weekly forms: 1	47	60
2	3	1
Missing	4 (7.4)	
No. days needed pain relief in last 7 days[a]	*n* = 52	*n* = 61
Mean (SD)	2.0 (2.5)	2.2 (2.6)
Median (IQI)	1.0 (0.0–3.0)	1.0 (0.0–3.5)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	49	60
2	3	1
EQ-5D score[a,d]	*n* = 50	*n* = 61
Mean (SD)	0.760 (0.239)	0.826 (0.179)
Median (IQI)	0.788 (0.669–0.986)	0.864 (0.692–0.987)

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 60 RCT2: week 3 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 73	PCNL *N* = 86
Minimum, maximum	(0.065, 0.989)	(0.159, 0.989)
No. of weekly forms: 1	47	60
2	3	1
EQ-5D-VAS[a,e]	*n* = 48	*n* = 57
Mean (SD)	75.4 (19.9)	80.9 (18.4)
Median (IQI)	80.0 (70.0–90.0)	85.0 (78.0–95.0)
Minimum, maximum	(10, 100)	(25, 100)
No. of weekly forms: 1	45	56
2	3	1

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

TABLE 61 RCT2: week 4 post-intervention summaries for pain and health status

Variable	FURS *N* = 73	PCNL *N* = 86
Level of pain today[a,b]	*n* = 50	*n* = 61
Mean (SD)	1.9 (2.4)	1.3 (2.0)
Median (IQI)	1.0 (0.0–3.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 9)	(0, 9)
No. of weekly forms: 1	48	61
2	2	
Worst level of pain since intervention[a]	*n* = 50	*n* = 60
Mean (SD)	6.8 (3.0)	4.3 (3.6)
Median (IQI)	8.0 (6.0–9.0)	4.5 (0.0–8.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	48	60
2	2	
Stone pain in last 7 days[c]	*n* = 47	*n* = 60
Yes	21 (41.2)	13 (21.0)
No	26 (51.0)	47 (75.8)
No. of weekly forms: 1	45	60
2	2	
Missing	4 (7.8)	2 (3.2)

TABLE 61 RCT2: week 4 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 73	PCNL *N* = 86
No. days needed pain relief in last 7 days[a]	*n* = 50	*n* = 59
Mean (SD)	1.6 (2.4)	1.4 (2.3)
Median (IQI)	0.0 (0.0–2.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	48	59
2	2	
EQ-5D score[a,d]	*n* = 47	*n* = 57
Mean (SD)	0.777 (0.254)	0.848 (0.175)
Median (IQI)	0.868 (0.669–0.987)	0.868 (0.742–0.987)
Minimum, maximum	(0.024, 0.989)	(0.045, 0.989)
No. of weekly forms: 1	45	54
2	2	3
EQ-5D-VAS[a,e]	*n* = 44	*n* = 55
Mean (SD)	73.9 (23.6)	79.5 (22.0)
Median (IQI)	80.0 (62.5–90.0)	85.0 (70.0–99.0)
Minimum, maximum	(15, 100)	(10, 100)
No. of weekly forms: 1	41	52
2	2	3
3	1	

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

TABLE 62 RCT2: week 5 post-intervention summaries for pain and health status

Variable	FURS *N* = 73	PCNL *N* = 86
Level of pain today[a,b]	*n* = 54	*n* = 59
Mean (SD)	1.3 (2.0)	0.9 (1.6)
Median (IQI)	0.0 (0.0–2.0)	0.0 (0.0–1.0)
Minimum, maximum	(0, 8)	(0, 6)
No. of weekly forms: 1	51	59
2	3	
Worst level of pain since intervention[a]	*n* = 54	*n* = 57
Mean (SD)	6.2 (3.4)	4.2 (3.5)
Median (IQI)	8.0 (4.0–8.0)	5.0 (0.0–7.0)

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 62 RCT2: week 5 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 73	PCNL *N* = 86
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	51	57
2	3	
Stone pain in last 7 days[c]	*n* = 52	*n* = 58
Yes	16 (28.6)	13 (22.0)
No	36 (64.3)	45 (76.3)
No. of weekly forms: 1	49	58
2	3	
Missing	4 (7.1)	1 (1.7)
No. days needed pain relief in last 7 days[a]	*n* = 54	*n* = 58
Mean (SD)	1.4 (2.3)	1.1 (2.2)
Median (IQI)	0.0 (0.0–3.0)	0.0 (0.0–1.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	51	58
2	3	
EQ-5D score[a,d]	*n* = 54	*n* = 56
Mean (SD)	0.814 (0.255)	0.886 (0.155)
Median (IQI)	0.937 (0.719–0.987)	0.985 (0.837–0.987)
Minimum, maximum	(–0.087, 0.989)	(0.491, 0.989)
No. of weekly forms: 1	50	56
2	4	
EQ-5D-VAS[a,e]	*n* = 49	*n* = 56
Mean (SD)	74.9 (25.1)	80.8 (22.2)
Median (IQI)	80.0 (65.0–95.0)	90.0 (70.5–98.5)
Minimum, maximum	(10, 100)	(10, 100)
No. of weekly forms: 1	45	56
2	4	

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

TABLE 63 RCT2: week 6 post-intervention summaries for pain and health status

Variable	FURS *N* = 73	PCNL *N* = 86
Level of pain today[a,b]	*n* = 52	*n* = 58
Mean (SD)	1.3 (2.2)	1.1 (1.9)
Median (IQI)	0.0 (0.0–2.0)	0.0 (0.0–1.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	51	57
2	1	1
Worst level of pain since intervention[a]	*n* = 53	*n* = 57
Mean (SD)	6.5 (3.5)	4.0 (3.4)
Median (IQI)	8.0 (5.0–9.0)	3.0 (0.0–7.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	52	56
2	1	1
Stone pain in last 7 days[c]	*n* = 50	*n* = 56
Yes	15 (27.8)	11 (19.0)
No	35 (64.8)	45 (77.6)
No. of weekly forms: 1	49	55
2	1	1
Missing	4 (7.4)	2 (3.4)
No. days needed pain relief in last 7 days[a]	*n* = 53	*n* = 56
Mean (SD)	1.4 (2.5)	1.2 (2.2)
Median (IQI)	0.0 (0.0–2.0)	0.0 (0.0–1.5)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	52	55
2	1	1
EQ-5D score[a,d]	*n* = 52	*n* = 54
Mean (SD)	0.822 (0.265)	0.853 (0.202)
Median (IQI)	0.985 (0.707–0.987)	0.985 (0.757–0.987)
Minimum, maximum	(−0.250, 0.989)	(0.211, 0.989)
No. of weekly forms: 1	50	53
2	2	1
EQ-5D-VAS[a,e]	*n* = 46	*n* = 55
Mean (SD)	78.9 (23.2)	79.9 (22.6)
Median (IQI)	87.3 (70.0–95.0)	87.0 (67.0–98.0)

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 63 RCT2: week 6 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 73	PCNL *N* = 86
Minimum, maximum	(10, 100)	(10, 100)
No. of weekly forms: 1	44	54
2	2	1

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

TABLE 64 RCT2: week 7 post-intervention summaries for pain and health status

Variable	FURS *N* = 73	PCNL *N* = 86
Level of pain today [a,b]	*n* = 51	*n* = 60
Mean (SD)	1.2 (2.4)	1.3 (2.2)
Median (IQI)	0.0 (0.0–2.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 10)	(0, 7)
No. of weekly forms: 1	50	60
2	1	
Worst level of pain since intervention [a]	*n* = 51	*n* = 60
Mean (SD)	6.6 (3.4)	4.5 (3.5)
Median (IQI)	8.0 (5.0–9.0)	5.0 (0.5–8.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	51	60
Stone pain in last 7 days [c]	*n* = 51	*n* = 57
Yes	15 (28.3)	10 (16.7)
No	36 (67.9)	47 (78.3)
No. of weekly forms: 1	50	57
2	1	
Missing	2 (3.8)	3 (5.0)
No. days needed pain relief in last 7 days [a]	*n* = 50	*n* = 58
Mean (SD)	1.5 (2.5)	1.4 (2.5)
Median (IQI)	0.0 (0.0–3.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	49	58
2	1	

TABLE 64 RCT2: week 7 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 73	PCNL *N* = 86
EQ-5D score[a,d]	*n* = 51	*n* = 58
Mean (SD)	0.810 (0.272)	0.824 (0.253)
Median (IQI)	0.986 (0.697–0.987)	0.985 (0.731–0.987)
Minimum, maximum	(−0.052, 0.989)	(−0.136, 0.989)
No. of weekly forms: 1	50	58
2	1	
EQ-5D-VAS[a,e]	*n* = 48	*n* = 57
Mean (SD)	78.8 (22.8)	76.7 (24.1)
Median (IQI)	82.5 (70.0–96.5)	85.0 (60.0–98.0)
Minimum, maximum	(10, 100)	(10, 100)
No. of weekly forms: 1	47	57
2	1	

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

TABLE 65 RCT2: week 8 post-intervention summaries for pain and health status

Variable	FURS *N* = 73	PCNL *N* = 86
Level of pain today[a,b]	*n* = 52	*n* = 60
Mean (SD)	1.5 (2.7)	1.3 (2.3)
Median (IQI)	0.0 (0.0–2.5)	0.0 (0.0–2.0)
Minimum, maximum	(0, 10)	(0, 8)
No. of weekly forms: 1	51	59
2	1	1
Worst level of pain since intervention[a]	*n* = 52	*n* = 59
Mean (SD)	6.8 (3.4)	4.5 (3.6)
Median (IQI)	8.0 (5.0–9.0)	5.0 (0.0–8.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	51	58
2	1	1
Stone pain in last 7 days[c]	*n* = 51	*n* = 55
Yes	14 (25.9)	10 (16.7)
No	37 (68.5)	45 (75.0)

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 65 RCT2: week 8 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 73	PCNL *N* = 86
No. of weekly forms: 1	50	54
2	1	1
Missing	3 (5.6)	5 (8.3)
No. days needed pain relief in last 7 days[a]	*n* = 52	*n* = 60
Mean (SD)	1.1 (2.2)	1.6 (2.7)
Median (IQI)	0.0 (0.0–1.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	51	59
2	1	1
EQ-5D score[a,d]	*n* = 53	*n* = 58
Mean (SD)	0.807 (0.268)	0.833 (0.238)
Median (IQI)	0.985 (0.693–0.987)	0.985 (0.736–0.987)
Minimum, maximum	(–0.052, 0.989)	(–0.136, 0.989)
No. of weekly forms: 1	52	57
2	1	1
EQ-5D-VAS[a,e]	*n* = 50	*n* = 57
Mean (SD)	76.9 (23.4)	76.3 (24.8)
Median (IQI)	80.0 (70.0–95.0)	80.0 (60.0–98.0)
Minimum, maximum	(10, 100)	(10, 100)
No. of weekly forms: 1	48	55
2	2	2

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

TABLE 66 RCT2: week 9 post-intervention summaries for pain and health status

Variable	FURS *N* = 73	PCNL *N* = 86
Level of pain today[a,b]	*n* = 52	*n* = 61
Mean (SD)	1.1 (2.1)	1.0 (1.9)
Median (IQI)	0.0 (0.0–1.3)	0.0 (0.0–1.0)
Minimum, maximum	(0, 9)	(0, 8)
No. of weekly forms: 1	51	59
2	1	2

TABLE 66 RCT2: week 9 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 73	PCNL *N* = 86
Worst level of pain since intervention[a]	*n* = 52	*n* = 60
Mean (SD)	6.8 (3.3)	4.3 (3.6)
Median (IQI)	8.0 (5.5–9.0)	5.0 (0.0–8.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	51	58
2	1	2
Stone pain in last 7 days[c]	*n* = 51	*n* = 58
Yes	14 (25.9)	11 (18.0)
No	37 (68.5)	47 (77.0)
No. of weekly forms: 1	50	56
2	1	2
Missing	3 (5.6)	3 (4.9)
No. days needed pain relief in last 7 days[a]	*n* = 51	*n* = 61
Mean (SD)	1.1 (2.0)	1.4 (2.4)
Median (IQI)	0.0 (0.0–2.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	50	59
2	1	2
EQ-5D score[a,d]	*n* = 51	*n* = 57
Mean (SD)	0.828 (0.270)	0.858 (0.234)
Median (IQI)	0.985 (0.753–0.987)	0.986 (0.791–0.987)
Minimum, maximum	(−0.133, 0.989)	(−0.136, 0.989)
No. of weekly forms: 1	50	55
2	1	1
3		1
EQ-5D-VAS[a,e]	*n* = 49	*n* = 57
Mean (SD)	82.6 (18.0)	77.9 (23.9)
Median (IQI)	90.0 (80.0–96.0)	85.0 (60.0–98.0)
Minimum, maximum	(20, 100)	(10, 100)
No. of weekly forms: 1	48	55
2	1	1
3		1

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 67 RCT2: week 10 post-intervention summaries for pain and health status

Variable	FURS *N* = 73	PCNL *N* = 86
Level of pain today[a,b]	*n* = 53	*n* = 56
Mean (SD)	1.0 (2.0)	0.9 (1.9)
Median (IQI)	0.0 (0.0–1.0)	0.0 (0.0–1.0)
Minimum, maximum	(0, 8)	(0, 7)
No. of weekly forms: 1	51	56
2	2	
Worst level of pain since intervention[a]	*n* = 52	*n* = 56
Mean (SD)	6.6 (3.5)	4.4 (3.4)
Median (IQI)	8.0 (5.5–9.0)	5.0 (0.0–8.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	50	56
2	2	
Stone pain in last 7 days[c]	*n* = 53	*n* = 54
Yes	12 (22.2)	5 (8.8)
No	41 (75.9)	49 (86.0)
No. of weekly forms: 1	51	54
2	2	
Missing	1 (1.9)	3 (5.3)
No. days needed pain relief in last 7 days[a]	*n* = 53	*n* = 55
Mean (SD)	0.9 (1.9)	1.4 (2.5)
Median (IQI)	0.0 (0.0–1.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	51	55
2	2	
EQ-5D score[a,d]	*n* = 52	*n* = 55
Mean (SD)	0.835 (0.271)	0.845 (0.249)
Median (IQI)	0.986 (0.793–0.987)	0.986 (0.745–0.987)
Minimum, maximum	(–0.200, 0.989)	(–0.136, 0.989)
No. of weekly forms: 1	51	55
2	1	
EQ-5D-VAS[a,e]	*n* = 49	*n* = 55
Mean (SD)	84.2 (18.6)	76.5 (23.6)
Median (IQI)	90.0 (80.0–98.0)	80.0 (60.0–97.0)

TABLE 67 RCT2: week 10 post-intervention summaries for pain and health status (*continued*)

Variable	FURS *N* = 73	PCNL *N* = 86
Minimum, maximum	(10, 100)	(10, 100)
No. of weekly forms: 1	48	55
2	1	

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

TABLE 68 RCT2 week 11: Post-intervention summaries for pain and health status

Variable	FURS *N* = 73	PCNL *N* = 86
Level of pain today[a,b]	*n* = 55	*n* = 57
Mean (SD)	0.7 (1.7)	1.1 (1.9)
Median (IQI)	0.0 (0.0–0.0)	0.0 (0.0–2.0)
Minimum, maximum	(0, 7)	(0, 6)
No. of weekly forms: 1	55	57
Mean (SD)	*n* = 55	*n* = 56
Mean (SD)	6.3 (3.6)	4.2 (3.5)
Median (IQI)	8.0 (4.0–9.0)	4.5 (0.0–8.0)
Minimum, maximum	(0, 10)	(0, 10)
No. of weekly forms: 1	55	56
Stone pain in last 7 days[c]	*n* = 54	*n* = 54
Yes	9 (16.4)	11 (19.3)
No	45 (81.8)	43 (75.4)
No. of weekly forms: 1	54	54
Missing	1 (1.8)	3 (5.3)
No. days needed pain relief in last 7 days[a]	*n* = 55	*n* = 56
Mean (SD)	1.0 (1.8)	1.2 (2.3)
Median (IQI)	0.0 (0.0–1.0)	0.0 (0.0–1.5)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	55	56
EQ-5D score[a,d]	*n* = 51	*n* = 55
Mean (SD)	0.828 (0.264)	0.846 (0.247)
Median (IQI)	0.986 (0.753–0.987)	0.985 (0.786–0.987)
Minimum, maximum	(−0.295, 0.989)	(−0.136, 0.989)
No. of weekly forms: 1	51	55

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 68 RCT2 week 11: Post-intervention summaries for pain and health status (*continued*)

Variable	FURS N = 73	PCNL N = 86
EQ-5D-VAS[a,e]	*n* = 49	*n* = 54
Mean (SD)	85.0 (18.4)	78.4 (24.1)
Median (IQI)	90.0 (80.0–99.0)	87.5 (60.0–98.0)
Minimum, maximum	(10, 100)	(10, 100)
No. of weekly forms: 1	49	54

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).

TABLE 69 RCT2 week 12: Post-intervention summaries for pain, health status and QoL

Variable	FURS N = 73	PCNL N = 86
Level of pain today[a,b]	*n* = 54	*n* = 60
Mean (SD)	0.8 (1.8)	0.9 (2.0)
Median (IQI)	0.0 (0.0–0.0)	0.0 (0.0–0.0)
Minimum, maximum	(0, 7)	(0, 9)
No. of weekly forms: 1	53	60
2	1	
Stone pain in last 7 days[c]	*n* = 52	*n* = 61
Yes	12 (21.4)	11 (17.7)
No	40 (71.4)	50 (80.6)
No. of weekly forms: 1	51	61
2	1	
Missing	4 (7.1)	1 (1.6)
No. days needed pain relief in last 7 days[a]	*n* = 51	*n* = 60
Mean (SD)	0.9 (2.2)	0.9 (2.0)
Median (IQI)	0.0 (0.0–0.0)	0.0 (0.0–0.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	50	60
2	1	
EQ-5D score[a,d]	***n* = 55**	***n* = 61**
Mean (SD)	0.833 (0.258)	0.851 (0.236)
Median (IQI)	0.986 (0.753–0.987)	0.985 (0.804–0.987)
Minimum, maximum	(−0.087, 0.989)	(−0.136, 0.989)
No. of weekly forms: 1	54	61
2	1	

TABLE 69 RCT2 week 12: Post-intervention summaries for pain, health status and QoL (*continued*)

Variable	FURS *N* = 73	PCNL *N* = 86
EQ-5D-VAS[a,e]	*n* = 51	*n* = 58
Mean (SD)	83.8 (15.6)	79.1 (22.1)
Median (IQI)	90.0 (75.0–98.0)	85.0 (70.0–98.0)
Minimum, maximum	(40, 100)	(10, 100)
No. of weekly forms: 1	50	58
2	1	
SF-12 scores[a,f]		
Physical component score	*n* = 51	*n* = 58
Mean (SD)	47.6 (11.3)	46.5 (11.5)
Median (IQI)	51.7 (40.3–56.1)	50.5 (40.4–54.8)
Minimum, maximum	(17.4, 64.6)	(15.3, 60.9)
No. of weekly forms: 1	54	58
Mental component score	*n* = 54	*n* = 58
Mean (SD)	49.7 (11.1)	50.2 (12.0)
Median (IQI)	53.8 (42.6–58.5)	55.2 (44.3–59.3)
Minimum, maximum	(20.7, 67.3)	(15.0, 63.4)
No. of weekly forms: 1	54	58

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).
f SF-12 scores range from 0 to 100 (higher score is better).

TABLE 70 RCT2: 12-month post-randomisation summaries for pain, health status and QoL

Variable	FURS *N* = 73	PCNL *N* = 86
Level of pain today[a,b]	*n* = 52	*n* = 57
Mean (SD)	1.5 (2.6)	0.9 (2.0)
Median (IQI)	0.0 (0.0–3.0)	0.0 (0.0–0.0)
Minimum, maximum	(0, 9)	(0, 10)
No. of weekly forms: 1	52	57
Stone pain in last 7 days[c]	*n* = 51	*n* = 54
Yes	13 (24.1)	12 (21.1)
No	38 (70.4)	42 (73.7)

continued

Copyright © 2025 Wiseman *et al*. This work was produced by Wiseman *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 70 RCT2: 12-month post-randomisation summaries for pain, health status and QoL (*continued*)

Variable	FURS *N* = 73	PCNL *N* = 86
No. of weekly forms: 1	51	54
Missing	3 (5.6)	3 (5.3)
No. days needed pain relief in last 7 days[a]	*n* = 53	*n* = 55
Mean (SD)	1.4 (2.5)	0.9 (1.9)
Median (IQI)	0.0 (0.0–2.0)	0.0 (0.0–0.0)
Minimum, maximum	(0, 7)	(0, 7)
No. of weekly forms: 1	53	55
EQ-5D score[a,d]	*n* = 50	*n* = 55
Mean (SD)	0.759 (0.361)	0.776 (0.261)
Median (IQI)	0.986 (0.692–0.987)	0.829 (0.656–0.987)
Minimum, maximum	(–0.253, 0.989)	(–0.174, 0.989)
No. of weekly forms: 1	50	55
EQ-5D-VAS[a,e]	*n* = 46	*n* = 53
Mean (SD)	78.8 (21.5)	75.3 (24.5)
Median (IQI)	85.0 (70.0–95.0)	80.0 (60.0–99.0)
Minimum, maximum	(10, 100)	(20, 100)
No. of weekly forms: 1	46	53
SF-12 scores[a,f]		
Physical component score	*n* = 49	*n* = 56
Mean (SD)	46.2 (11.4)	45.0 (12.1)
Median (IQI)	51.6 (36.8–55.1)	49.0 (34.7–55.4)
Minimum, maximum	(16.2, 63.2)	(16.7, 60.9)
No. of weekly forms: 1	49	56
Mental component score	*n* = 49	*n* = 56
Mean (SD)	50.2 (11.5)	50.4 (10.1)
Median (IQI)	55.4 (43.6–59.6)	53.5 (43.6–58.3)
Minimum, maximum	(28.7, 63.8)	(21.3, 64.2)
No. of weekly forms: 1	49	56

n, number of responses.
a Continuous or pseudo-continuous variable *n*: mean (SD), median (IQI), (minimum, maximum).
b Level of pain today: 0 to 10 (higher score indicates more pain).
c Categorical/ordinal *x* (%).
d NICE: recommend EQ-5D-5L mapped score also accounts for age and sex (higher score is better QoL).
e EQ-5D-VAS: scale of 0 to 100 (higher score indicates better health).
f SF-12 scores range from 0 to 100 (higher score is better).

RCT2: subgroup analyses

Subgroup analyses primary and main secondary outcome treatment comparison estimates: Plots – RCT2.

Graphical representations of the impact of each prespecified subgroup on the treatment effect on the five main outcome measures.

Three PO measures:

- Main PO: The main PO is the AUC of EQ-5D scores from pre intervention to 12 weeks, averaged over 12 weeks and adjusted for baseline EQ-5D scores and where cluster effects are accounted for using a robust variance at centre level.
- Sensitivity 1: The same averaged AUC as above but with covariates of baseline EQ-5D scores and waiting times; clusters adjusted as above.
- Sensitivity 2: The AUC as main PO, adjusted for clusters but without other covariates.

Two secondary outcomes:

- AUC from baseline to 12 weeks – This is an AUC of all EQ-5D scores anchored from baseline right up to 12 weeks post intervention, standardised over the total number of weeks and adjusted for clusters.
- Stone clearance (complete/acceptable/unacceptable) by 12 weeks post intervention also adjusted for clusters.

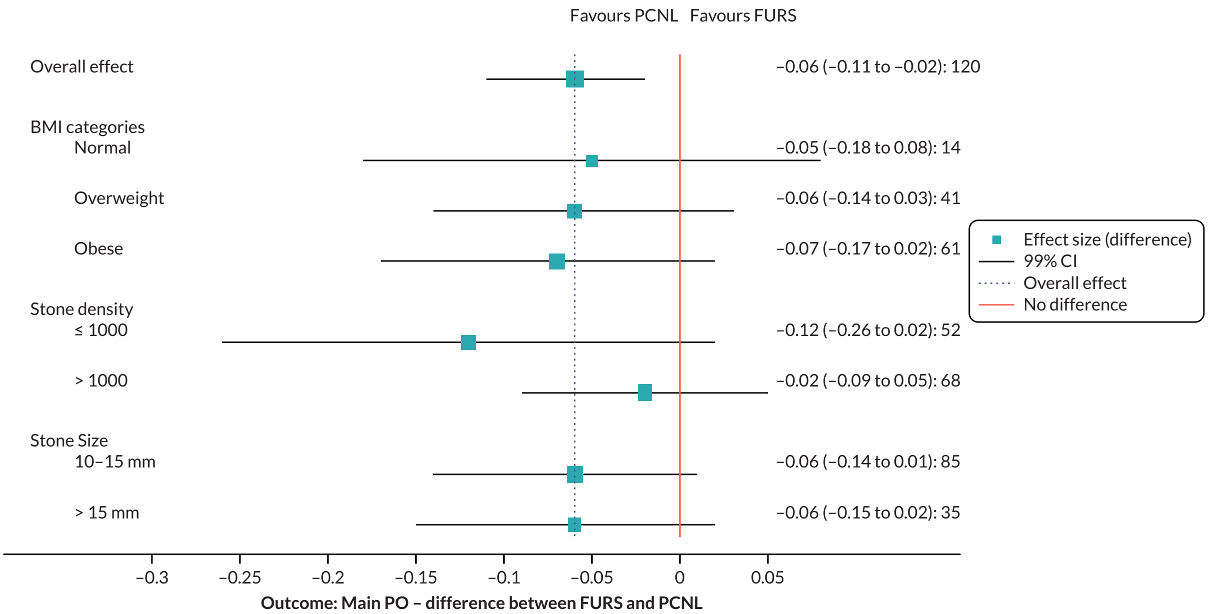


FIGURE 27 RCT2 subgroups – main PO: forest plot showing subgroup model effects.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

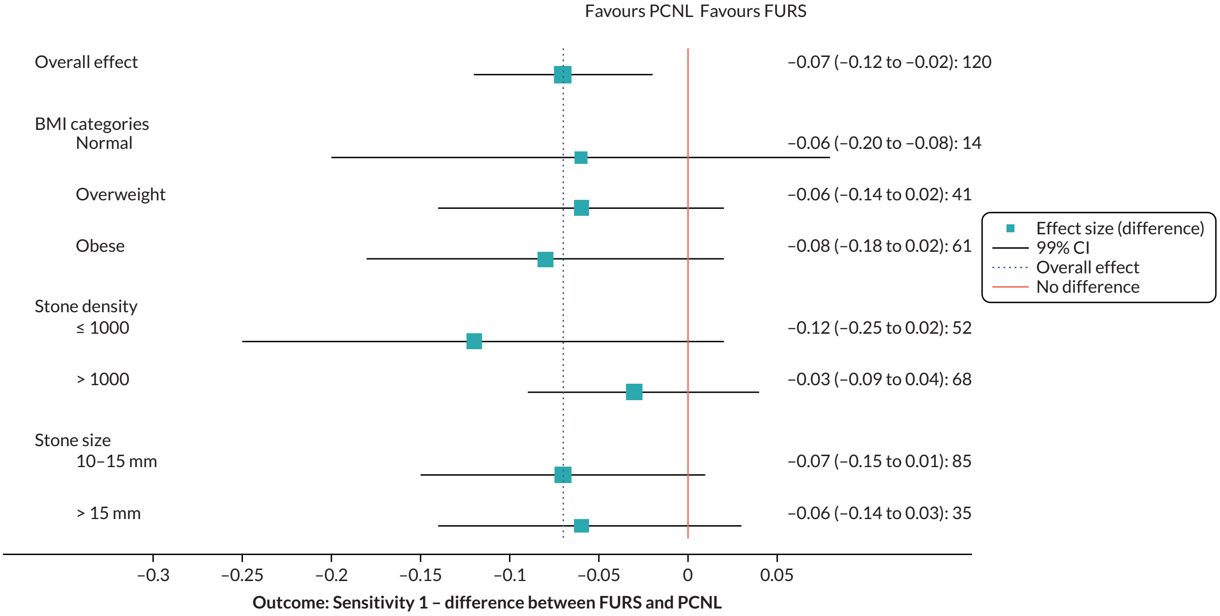


FIGURE 28 RCT2 subgroups – sensitivity 1: forest plot showing subgroup model effects.

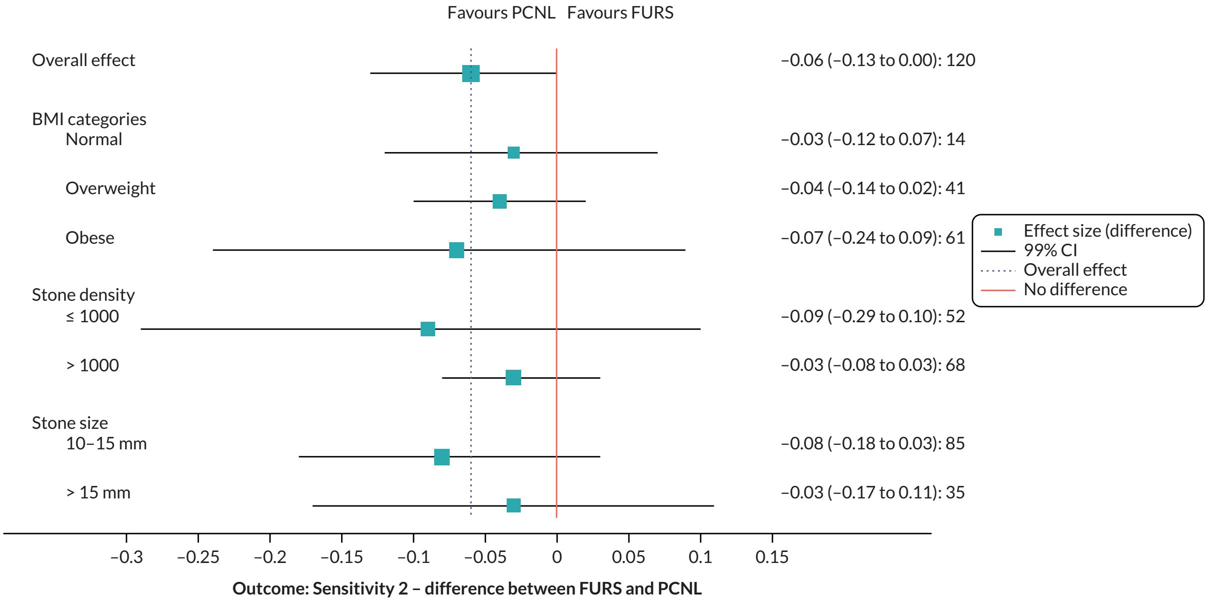


FIGURE 29 RCT2 subgroups – sensitivity 2: forest plot showing subgroup model effects.

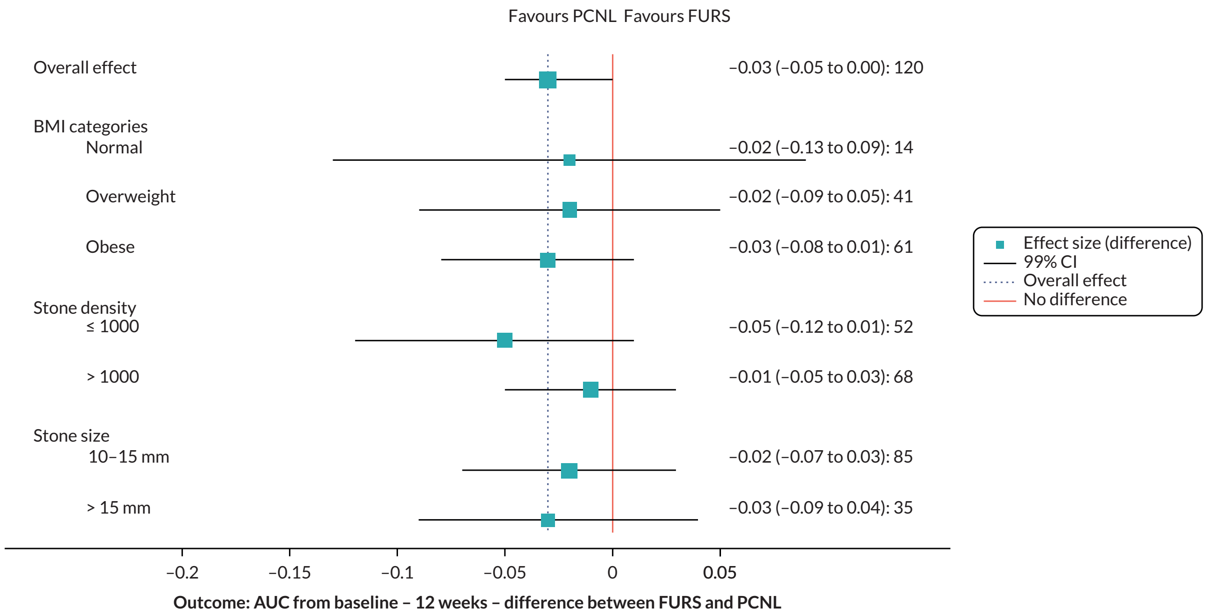


FIGURE 30 RCT2 subgroups – AUC from baseline – 12 weeks: forest plot showing subgroup model effects.

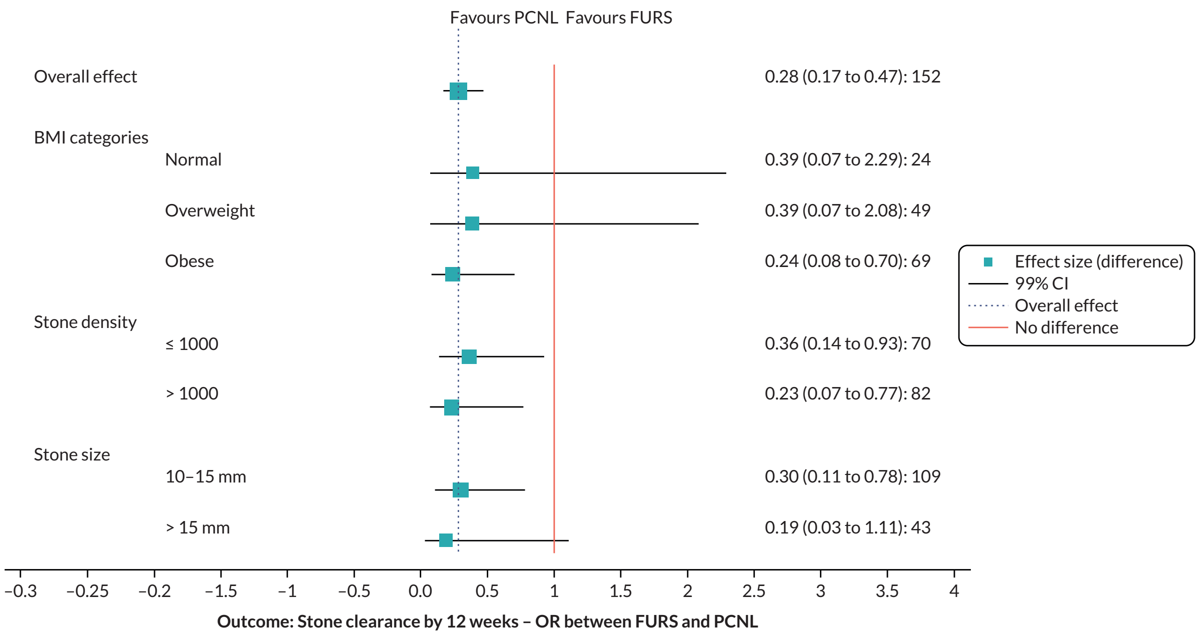


FIGURE 31 RCT2 subgroups – stone clearance at 12 weeks: forest plot showing subgroup model effects.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Appendix 8 Technical factors – primary and main secondary outcome summaries within each technical factor

Three PO measures:

- Main PO: The main PO is the AUC of EQ-5D scores from pre intervention to 12 weeks, averaged over 12 weeks and adjusted for baseline EQ-5D scores and where cluster effects are accounted for using a robust variance at centre level.
- Sensitivity 1: The same averaged AUC as above but with covariates of baseline EQ-5D scores and waiting times; clusters adjusted as above.
- Sensitivity 2: The AUC as main PO, adjusted for clusters but without other covariates.

Two secondary outcomes:

- AUC from baseline – 12 weeks – this is an AUC of all EQ-5D scores anchored from baseline right up to 12 weeks post intervention, standardised over the total number of weeks and adjusted for clusters.
- Stone clearance (complete/acceptable/unacceptable) by 12 weeks post intervention also adjusted for clusters.

Flexible ureterorenoscopy – access sheath versus no sheath

The outcomes and covariates in the primary models (main, sensitivity 1 and 2) and the secondary AUC model in each sheath/no sheath group were summarised. They are virtually identical regardless of whether a sheath was used or not, for the main variable and the adjusting covariates.

Stone clearance at 12 weeks overall showed similar complete stone clearance (70–80%). If the stone was < 10 mm (which was the majority as the overall numbers are dominated by RCT1), the proportions were around 80%. The larger stones 10–15 mm show an apparent difference 45% versus 76% for sheath versus no sheath respectively. There is little difference for largest stones with a low clearance of < 50%. The numbers in the two largest stone size groups are small.

TABLE 71 Flexible ureterorenoscopy intervention – use of access sheath

Outcome	Yes	No
Main PO:[a]	*107*	*97*
EQ-5D score AUC (intervention to 12 weeks)	0.8 (0.2)	0.8 (0.2)
	0.9 (0.7–0.9)	0.9 (0.8–0.9)
	(–0.1, 1.0)	(–0.1, 1.0)
EQ-5D score at baseline	0.8 (0.3)	0.8 (0.2)
	0.9 (0.7–1.0)	0.8 (0.7–1.0)
	(–0.1, 1.0)	(–0.1, 1.0)

TABLE 71 Flexible ureterorenoscopy intervention – use of access sheath (*continued*)

Outcome	Yes	No
Sensitivity 1:[a]	***107***	***97***
EQ-5D score AUC (intervention to 12 weeks)	0.8 (0.2)	0.8 (0.2)
	0.9 (0.7–0.9)	0.9 (0.8–0.9)
	(−0.1, 1.0)	(−0.1, 1.0)
EQ-5D score at baseline	0.8 (0.3)	0.8 (0.2)
	0.9 (0.7–1.0)	0.8 (0.7–1.0)
	(−0.1, 1.0)	(−0.1, 1.0)
Waiting time to intervention in weeks	19.7 (16.4)	19.7 (14.4)
	15.0 (8.7–25.1)	15.1 (10.6–26.1)
	(1.4, 91.7)	(1.9, 75.1)
Sensitivity 2:[a]	***107***	***97***
EQ-5D score AUC (intervention to 12 weeks)	0.8 (0.2)	0.8 (0.2)
	0.9 (0.7–0.9)	0.9 (0.8–0.9)
	(−0.1, 1.0)	(−0.1, 1.0)
EQ-5D score at pre intervention	0.8 (0.2)	0.8 (0.2)
	0.8 (0.7–0.9)	0.8 (0.7–1.0)
	(−0.1, 1.0)	(0.1, 1.0)
AUC from baseline – 12 weeks[a]	***107***	***97***
EQ-5D score AUC (randomisation to 12 weeks)	0.8 (0.2)	0.8 (0.2)
	0.8 (0.7–0.9)	0.8 (0.7–0.9)
	(−0.1, 1.0)	(−0.0, 1.0)
EQ-5D score at baseline	0.8 (0.3)	0.8 (0.2)
	0.9 (0.7–1.0)	0.8 (0.7–1.0)
	(−0.1, 1.0)	(−0.1, 1.0)
Stone clearance at 12 weeks[b]	***136***	***119***
Complete	92 (67.6)	95 (79.8)
Acceptable	6 (4.4)	7 (5.9)
Unacceptable	38 (27.9)	17 (14.3)
Stone clearance at 12 weeks – (< 10 mm stones)	91	93
Complete	71 (78.0)	77 (82.8)
Acceptable	4 (4.4)	4 (4.3)
Unacceptable	16 (17.6)	12 (12.9)
Stone clearance at 12 weeks – (10–15 mm stones)	32	21
Complete	15 (46.9)	16 (76.2)
Acceptable	2 (6.3)	2 (9.5)
Unacceptable	15 (46.9)	3 (14.3)

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 71 Flexible ureterorenoscopy intervention – use of access sheath (*continued*)

Outcome	Yes	No
Stone clearance at 12 weeks – (> 15 mm stones)	13	5
Complete	6 (46.2)	2 (40.0)
Acceptable	-	1 (20.0)
Unacceptable	7 (53.8)	2 (40.0)

a Mean (SD), median (IQI) and (minimum, maximum).
b *N n* (%).

Flexible ureterorenoscopy – digital versus non-digital instruments

All the PO summaries (main, sensitivity 1 and 2) and both secondary measures AUC and stone clearance at 12 weeks are virtually comparable for digital and non-digital instruments (see *Table 72*). The only difference was waiting times were longer for treatment with non-digital instruments.

TABLE 72 Flexible ureterorenoscopy intervention – digital vs. non-digital instruments

Outcome	Summary	Digital	Non-digital
Main PO:	*N*	**49**	**153**
EQ-5D score AUC (intervention to 12 weeks)	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.9 (0.8–1.0)	0.9 (0.7–0.9)
	Minimum, maximum	(−0.1, 1.0)	(−0.1, 1.0)
EQ-5D score at baseline	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.8 (0.6–1.0)	0.9 (0.7–1.0)
	Minimum, maximum	(0.0, 1.0)	(−0.1, 1.0)
Sensitivity 1:	*N*	**49**	**153**
EQ-5D score AUC (intervention to 12 weeks)	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.9 (0.8–1.0)	0.9 (0.7–0.9)
	Minimum, maximum	(−0.1, 1.0)	(−0.1, 1.0)
EQ-5D score at baseline	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.8 (0.6–1.0)	0.9 (0.7–1.0)
	Minimum, maximum	(0.0, 1.0)	(−0.1, 1.0)
Waiting time to intervention	Mean (SD)	14.5 (12.2)	21.4 (16.0)
	Median (IQI)	11.0 (5.7–17.6)	17.7 (11.0–26.6)
	Min, max	(1.9, 51.1)	(1.4, 91.7)

TABLE 72 Flexible ureterorenoscopy intervention – digital vs. non-digital instruments (*continued*)

Outcome	Summary	Digital	Non-digital
Sensitivity 2:	*N*	**49**	**153**
EQ-5D score AUC (intervention to 12 weeks)	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.9 (0.8–1.0)	0.9 (0.7–0.9)
	Minimum, maximum	(−0.1, 1.0)	(−0.1, 1.0)
EQ-5D score at pre intervention	Mean (SD)	0.7 (0.3)	0.8 (0.2)
	Median (IQI)	0.8 (0.6–1.0)	0.8 (0.7–1.0)
	Minimum, maximum	(−0.1, 1.0)	(0.2, 1.0)
AUC from baseline – 12 weeks	*N*	**49**	**153**
EQ-5D score AUC (randomisation to 12 weeks)	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.8 (0.7–0.9)	0.8 (0.7–0.9)
	Minimum, maximum	(−0.1, 1.0)	(−0.0, 1.0)
EQ-5D score at baseline	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.8 (0.6–1.0)	0.9 (0.7–1.0)
	Minimum, maximum	(0.0, 1.0)	(−0.1, 1.0)
Stone clearance at 12 weeks	*N*	70	182
Complete	*n* (%)	55 (78.6)	131 (72.0)
Acceptable	*n* (%)	2 (2.9)	11 (6.0)
Unacceptable	*n* (%)	13 (18.6)	40 (22.0)

Extracorporeal shockwave lithotripsy – fixed versus mobile device

The PO summaries and the secondary AUC measure were comparable for fixed and mobile devices. Waiting times were slightly shorter for fixed devices. Complete stone clearance at 12 weeks did differ; mobile: (*n* = 131, 30%) compared to fixed (*n* = 89, 47%) (see *Table 73*).

TABLE 73 Extracorporeal shockwave lithotripsy intervention – lithotripter type

Outcome	Summary	Mobile	Fixed
Main PO:	*N*	**114**	**73**
EQ-5D score AUC (intervention to 12 weeks)	Mean (SD)	0.8 (0.2)	0.9 (0.1)
	Median (IQI)	0.9 (0.7–1.0)	0.9 (0.8–1.0)
	Minimum, maximum	(−0.0, 1.0)	(0.5, 1.0)
EQ-5D score at baseline	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.9 (0.7–1.0)	0.9 (0.7–1.0)
	Minimum, maximum	(−0.2, 1.0)	(0.2, 1.0)

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 73 Extracorporeal shockwave lithotripsy intervention – lithotripter type (*continued*)

Outcome	Summary	Mobile	Fixed
Sensitivity 1:	*N*	**114**	**73**
EQ-5D score AUC (intervention to 12 weeks)	Mean (SD)	0.8 (0.2)	0.9 (0.1)
	Median (IQI)	0.9 (0.7–1.0)	0.9 (0.8–1.0)
	Min, max	(−0.0, 1.0)	(0.5, 1.0)
EQ-5D score at baseline	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.9 (0.7, 1.0)	0.9 (0.7, 1.0)
	Minimum, maximum	(−0.2, 1.0)	(0.2, 1.0)
Waiting time to intervention	Mean (SD)	11.7 (11.3)	7.5 (6.8)
	Median (IQI)	7.9 (4.9–14.0)	5.3 (3.3–9.3)
	Minimum, maximum	(0.0, 55.9)	(0.1, 42.0)
Sensitivity 2:	*N*	**114**	**73**
EQ-5D score AUC (intervention to 12 weeks)	Mean (SD)	0.8 (0.2)	0.9 (0.1)
	Median (IQI)	0.9 (0.7–1.0)	0.9 (0.8–1.0)
	Minimum, maximum	(−0.0, 1.0)	(0.5, 1.0)
EQ-5D score at pre intervention	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.9 (0.7–1.0)	0.9 (0.7–1.0)
	Minimum, maximum	(−0.0, 1.0)	(0.3, 1.0)
AUC from baseline – 12 weeks	*N*	**114**	**73**
EQ-5D score AUC (randomisation to 12 weeks)	Mean (SD)	0.8 (0.2)	0.8 (0.1)
	Median (IQI)	0.9 (0.8–1.0)	0.9 (0.8–0.9)
	Minimum, maximum	(0.0, 1.0)	(0.5, 1.0)
EQ-5D score at baseline	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.9 (0.7–1.0)	0.9 (0.7–1.0)
	Minimum, maximum	(−0.2, 1.0)	(0.2, 1.0)
Stone clearance at 12 weeks	*N*	**131**	**89**
Complete	*n* (%)	39 (29.8)	42 (47.2)
Acceptable	*n* (%)	18 (13.7)	9 (10.1)
Unacceptable	*n* (%)	74 (56.5)	38 (42.7)

Extracorporeal shockwave lithotripsy – skin-to-stone distance: ≤ 120 mm versus > 120 mm

The PO summaries (main, sensitivity 1 and 2) and both secondary measures were comparable regardless of skin-to-stone distance (see *Table 74*). This was also the case for any adjusting variables too, with comparable median waiting times.

TABLE 74 Extracorporeal shockwave lithotripsy intervention – skin-to-stone distance

Outcome	Summary	≤ 120 mm	> 120 mm
Main PO:	*N*	**157**	**34**
EQ-5D score AUC (intervention to 12 weeks)	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.9 (0.7–1.0)	0.9 (0.7–0.9)
	Minimum, maximum	(−0.2, 1.0)	(0.2, 1.0)
EQ-5D score at baseline	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.9 (0.7–1.0)	0.9 (0.7–1.0)
	Minimum, maximum	(−0.2, 1.0)	(0.4, 1.0)
Sensitivity 1:	*N*	**157**	**34**
EQ-5D score AUC (intervention to 12 weeks)	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.9 (0.7–1.0)	0.9 (0.7–0.9)
	Minimum, maximum	(−0.2, 1.0)	(0.2, 1.0)
EQ-5D score at baseline	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.9 (0.7–1.0)	0.9 (0.7–1.0)
	Minimum, maximum	(−0.2, 1.0)	(0.4, 1.0)
Waiting time to intervention	Mean (SD)	9.7 (9.1)	9.5 (10.8)
	Median (IQI)	6.9 (4.0–12.0)	4.9 (3.3–10.0)
	Minimum, maximum	(0.0, 46.0)	(0.1, 46.0)
Sensitivity 2:	*N*	**157**	**34**
EQ-5D score AUC (intervention to 12 weeks)	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.9 (0.7–1.0)	0.9 (0.7–0.9)
	Minimum, maximum	(−0.2, 1.0)	(0.2, 1.0)
EQ-5D score at pre intervention	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.9 (0.7–1.0)	0.9 (0.7–1.0)
	Minimum, maximum	(−0.1, 1.0)	(0.4, 1.0)
AUC from baseline – 12 weeks	*N*	**157**	**34**
EQ-5D score AUC (randomisation to 12 weeks)	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.9 (0.8–1.0)	0.8 (0.7–0.9)
	Minimum, maximum	(−0.2, 1.0)	(0.4, 1.0)
EQ-5D score at baseline	Mean (SD)	0.8 (0.2)	0.8 (0.2)
	Median (IQI)	0.9 (0.7–1.0)	0.9 (0.7–1.0)
	Minimum, maximum	(−0.2, 1.0)	(0.4, 1.0)
Stone clearance at 12 weeks	*N*	**184**	**43**
Complete	*n* (%)	66 (35.9)	17 (39.5)
Acceptable	*n* (%)	23 (12.5)	4 (9.3)
Unacceptable	*n* (%)	95 (51.6)	22 (51.2)

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Percutaneous nephrolithotomy – calibre of access tract: < 22 versus ≥ 22

In the main, these came from RCT2 only and so the numbers are small. All the primary measures and their adjusting variables are similar with waiting means looking different but medians similar (see *Table 75*). Stone clearance at 12 weeks has again been split according to the size of the original, but mainly has 10–15 mm stones so any differences in the other stone size groups are unreliable.

TABLE 75 Percutaneous nephrolithotomy intervention – access tract size

Outcome	< 22	≥ 22
Main PO:[a]	**21**	**38**
EQ-5D score AUC (intervention to 12 weeks)	0.9 (0.2)	0.8 (0.2)
	0.9 (0.7–1.0)	0.9 (0.7–1.0)
	(0.5, 1.0)	(0.1, 1.0)
EQ-5D score at baseline	0.8 (0.2)	0.8 (0.3)
	0.9 (0.7–1.0)	0.8 (0.7–1.0)
	(0.5, 1.0)	(−0.1, 1.0)
Sensitivity 1:[a]	**21**	**38**
EQ-5D score AUC (intervention to 12 weeks)	0.9 (0.2)	0.8 (0.2)
	0.9 (0.7–1.0)	0.9 (0.7–1.0)
	(0.5, 1.0)	(0.1, 1.0)
EQ-5D score at baseline	0.8 (0.2)	0.8 (0.3)
	0.9 (0.7–1.0)	0.8 (0.7–1.0)
	(0.5, 1.0)	(−0.1, 1.0)
Waiting time to intervention	30.2 (24.1)	21.9 (13.9)
	18.1 (11.0–46.0)	16.6 (11.7–32.9)
	(5.0, 73.7)	(1.9, 59.9)
Sensitivity 2:[a]	**21**	**38**
EQ-5D score AUC (intervention to 12 weeks)	0.9 (0.2)	0.8 (0.2)
	0.9 (0.7–1.0)	0.9 (0.7–1.0)
	(0.5, 1.0)	(0.1, 1.0)
EQ-5D score at pre intervention	0.8 (0.1)	0.7 (0.2)
	0.8 (0.7–0.9)	0.8 (0.7–0.9)
	(0.5, 1.0)	(−0.1, 1.0)
AUC from baseline – 12 weeks[a]	**21**	**38**
EQ-5D score AUC (randomisation to 12 weeks)	0.8 (0.2)	0.8 (0.2)
	0.9 (0.7–0.9)	0.8 (0.7–0.9)
	(0.5, 1.0)	(0.1, 1.0)
EQ-5D score at baseline	0.8 (0.2)	0.8 (0.3)
	0.9 (0.7–1.0)	0.8 (0.7–1.0)
	(0.5, 1.0)	(−0.1, 1.0)

TABLE 75 Percutaneous nephrolithotomy intervention – access tract size (*continued*)

Outcome	< 22	≥ 22
Stone clearance at 12 weeks[b]	**30**	**38**
Complete	26 (86.7)	28 (73.7)
Acceptable	1 (3.3)	8 (21.1)
Unacceptable	3 (10.0)	2 (5.3)
Stone clearance at 12 weeks (< 10 mm stones)	3	.
Complete	3 (100.0)	-
Stone clearance at 12 weeks (10–15 mm stones)	20	26
Complete	17 (85.0)	22 (84.6)
Acceptable	1 (5.0)	3 (11.5)
Unacceptable	2 (10.0)	1 (3.8)
Stone clearance at 12 weeks (> 15 mm stones)	7	12
Complete	6 (85.7)	6 (50.0)
Acceptable	-	5 (41.7)

a Mean (SD), median (IQI) and (minimum, maximum).
b *N n* (%).

Copyright © 2025 Wiseman *et al*. This work was produced by Wiseman *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Appendix 9 Health economics

TABLE 76 Disposables and equipment assumed

Item	Unit cost (£)
Disposables	
FURS	
Sensor™ Wire	24.40
LithoVue™ ureteroscope	800.00
200 micron Lumenis® Moses™ laser fibre or Flexiva™ TracTip laser fibre	269.00
1.9Fr Zero Tip™ basket	151.74
24/26 cm 6 Fr Percuflex™ stent	56.55
24/26 cm 6Fr ImaJin® silicone stent	48.50
access sheath Navigator™ (half of cases)	80.47
FURS total cost of disposables	1391.93
PCNL	
Mini-access kit (e.g. Cook) with 18G needle (includes micro-wire and transitional dilator)	48.00
Two Amplatz Super Stiff™ wires	82.08
12Fr dilator	19.50
Ultrasound probe cover	2.90
Cover US machine	3.53
Baloon dilator (Nephromax™)	196.62
Fragmentation device (Swiss LithoClast® Trilogy probe 3.4 mm)	384.54
24/26 cm 6Fr Percuflex™ stent/drainage tube	56.55
PCNL Total cost of disposables	793.72
Equipment (assumes 5-year depreciation)	
ESWL	
Full lithotripsy system (including x-ray and ultrasound machine)	230,000.00
Service cost per year	30,000.00
Annual patient throughput	200.00
Total equipment per intervention	230.00
Total maintenance per intervention	150.00
FURS	
35W Laser	40,000.00
Service cost per year	7500.00
Annual patient throughput	100.00
Total equipment per intervention	80.00
Total maintenance per intervention	75.00

TABLE 76 Disposables and equipment assumed (*continued*)

Item	Unit cost (£)
PCNL	
Trilogy	42,000.00
Service cost per year	3374.00
Annual patient throughput	40.00
Total equipment per intervention	210.00
Total maintenance per intervention	84.35

TABLE 77 Regression model selection test results

Total costs RCT1		
Family	χ^2	***p-value***
Poisson:	0.8605	0.3536
Gaussian NLLS:	1.3260	0.2495
Gamma:	9.0405	0.0026
Inverse Gaussian	25.8661	0.0000
Results of tests for link; p-values		
GLM, Gaussian family		
	Identify link	
Pearson correlation	1.0000	
Pregibon link test:	0.0293	
Modified Hosmer and Lemeshow	0.0559	
Total QALY RCT1		
Family	χ^2	***p-value***
Gaussian NLLS:	2.05	0.1526
Poisson:	6.33	0.0119
Gamma:	12.96	0.0003
Inverse Gaussian	21.94	0.0000
Results of tests for link; p-values		
GLM identity link		
	Identify link	
Pearson correlation test:	1.0000	
Pregibon link test:	0.3266	
Modified Hosmer and Lemeshow:	0.1146	
Total costs RCT2		
Family	χ^2	***p-value***
Inverse Gaussian or Wald:	0.48	0.0000

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 77 Regression model selection test results (*continued*)

Gamma:	1.61	0.0000
Poisson:	3.39	0.0000
Gaussian NLLS:	5.83	0.0000
Results of tests for link; p-values		
GLM identity link		
	Identify link	
Pearson correlation test:	1.0000	
Pregibon link test:	0.0143	
Modified Hosmer and Lemeshow:	0.8444	
Total QALYs RCT2		
Family	χ^2	***p-value***
Gaussian NLLS:	1.16	0.2820
Poisson:	6.29	0.0122
Gamma:	15.52	0.0001
Inverse Gaussian or Wald:	28.85	0.0000
Results of tests for link; p-values		
GLM identity link		
	Identify link	
Pearson correlation test:	1.0000	
Pregibon link test:	0.0888	
Modified Hosmer and Lemeshow:	0.9921	

NLLS, non-linear least squares.

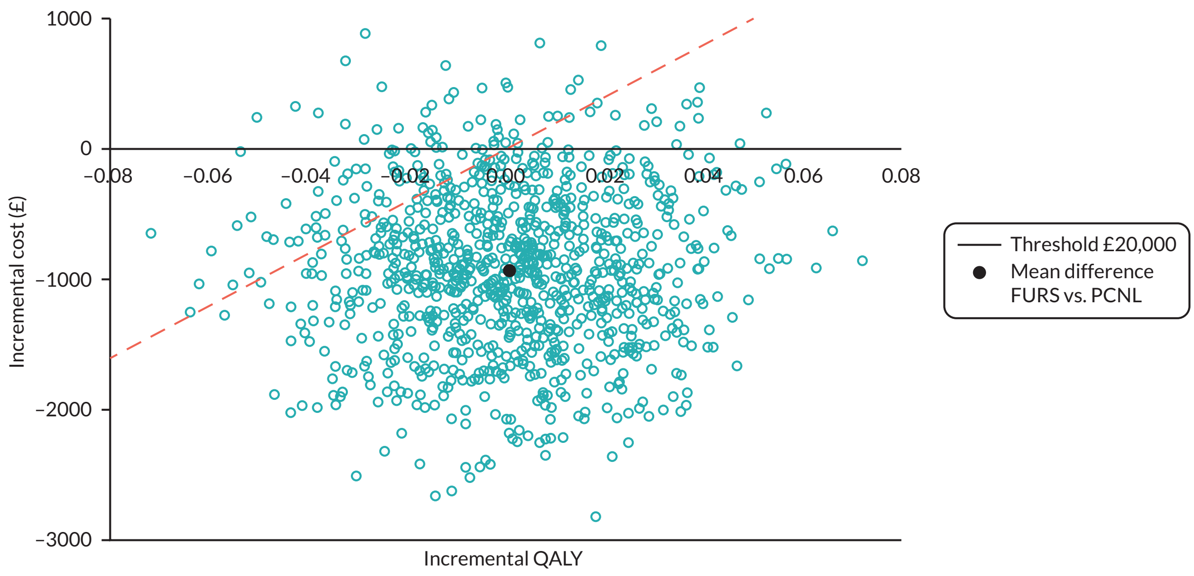


FIGURE 32 Trial-based incremental cost-effectiveness scatterplot for FURS vs. PCNL. (base case; imputed data – 1000 bootstrap iterations) HRG-based costing for index and additional interventions. Iterations below the red dashed £20,000 threshold line mean FURS is cost-effective; and above the line, PCNL is cost-effective.

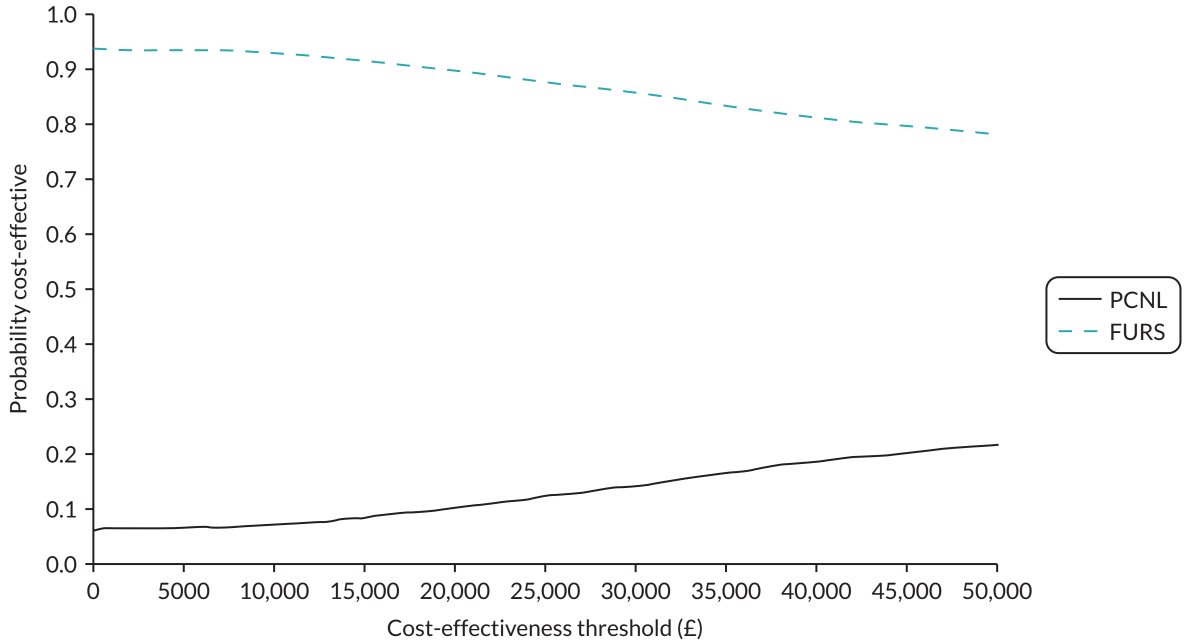


FIGURE 33 Cost-effectiveness acceptability curves for FURS and PCNL groups. (trial-based analysis; base case; imputed data – 1000 bootstrap iterations) HRG-based costing for index and additional interventions.

TABLE 78 Time to return to paid work, productivity costs, and out-of-pocket expenditure, RCT1

Variable	No. of observations	ESWL	FURS
Time off from paid work			
Twelve weeks post surgery (days); *n* (%); mean (SD)	316	166 (61); 2.3 (12.2)	150 (54); 6.6 (31)
Additional days up to 12 months post randomisation (days); *n* (%); mean (SD)	278	148 (55); 0.6 (3.3)	130 (52); 0.5 (2.6)
Productivity costs (£)			
Cost of time from paid work			
Twelve weeks post surgery; mean (SD)	311	377 (2047)	1022 (4835)
Additional cost up to 12 months post randomisation; mean (SD)	268	94 (515)	85 (422)
Total cost up to 12 months post randomisation; mean (SD)	340	416 (2028)	1012 (4686)
Out-of-pocket-expenditure (£)			
Medication costs; *n*, *n*', (%); mean (SD)	356	183; 45 (25); 4.6 (13.5)	173; 52 (30); 7.6 (33.9)
Private healthcare costs; *n*, *n*', (%); mean (SD)	357	183; 2 (1); 1 (14.5)	174; 5 (3); 0.7 (9.9)
Travel cost to:			
Hospital admission; *n* (%); mean (SD)	164	72 (31); 3.9 (6.9)	92 (39); 6.9 (13.4)
Outpatient consultation; *n* (%); mean (SD)	236	119 (52); 6.5 (11.6)	117 (50); 8.7 (13.6)
GP visit; *n* (%); mean (SD)	200	91 (39); 1 (2.6)	109 (47); 1.2 (2.8)

Note
n' = number of individuals declaring they bought medicines/private health care.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 79 Time to return to paid work, productivity costs, and out-of-pocket expenditure, RCT2

Variable	No. of observations	PCNL	FURS
Time off from paid work			
Twelve weeks post surgery (days); *n* (%); mean (SD)	109	58 (36); 2.4 (6.1)	51 (49); 7.6 (14.7)
Additional days up to 12 months post randomisation (days); *n* (%); mean (SD)	99	50 (34); 7 (42.8)	49 (43); 4 (21.8)
Productivity costs (£)			
Cost of time from paid work			
Twelve weeks post surgery; mean (SD)	108	364 (909)	1672 (5085)
Additional cost up to 12 months post randomisation; mean (SD)	97	1112 (6871)	673 (3692)
Total cost up to 12 months post randomisation; mean (SD)	117	1200 (6593)	2146 (6499)
Out-of-pocket expenditure (£)			
Medication costs; *n*, *n*', (%); mean (SD)	129	69; 9 (13); 0.6 (3.2)	60; 15 (25); 5.3 (18.7)
Private healthcare costs; *n*, *n*', (%); mean (SD)	129	70; 0 (0); 0 (0)	59; 2 (3); 1.8 (14.6)
Travel cost to:			
Hospital admission; *n* (%); mean (SD)	79	40 (47); 9.5 (16.2)	39 (53); 7.8 (12.4)
Outpatient consultation; *n* (%); mean (SD)	87	44 (51); 9 (13)	43 (59); 7.3 (9.8)
GP visit; *n* (%); mean (SD)	78	42 (49); 1.6 (3.4)	36 (49); 1.1 (2.4)

Note
n' = number of individuals declaring they bought medicines/private health care.

TABLE 80 Decision model parameter values and probability distributions

		RCT1					RCT2					
				Probability distribution					Probability distribution			
Variable name	Description	Mean value	SE (mean)	Type	Param 1	Param 2	Mean value	SE (mean)	Type	Param 1	Param 2	Source
Starting age	Cohort starting age	54					59					PUrE
Proportion of male in the cohort		0.62					0.63					PUrE
Time horizon	Number of months	60					60					Assumption
Costs												
Pre-surgical events	Cost for those waiting for index surgery	424					167					PUrE
Twelve weeks after treatment	Costs after index or further treatment	589					684					PUrE

TABLE 80 Decision model parameter values and probability distributions (*continued*)

		RCT1					RCT2					
				Probability distribution					Probability distribution			
Variable name	**Description**	**Mean value**	**SE (mean)**	**Type**	**Param 1**	**Param 2**	**Mean value**	**SE (mean)**	**Type**	**Param 1**	**Param 2**	**Source**
ESWL	LB36Z; extracorporeal lithotripsy – day case	894	89	Gamma	100.90	0.11	894	89	Gamma	100.90	0.11	NHS national cost collection
FURS	LB65E; major endoscopic, kidney or ureter procedures Nineteen years and over, with CC score 0–2 – day case	3350	335	Gamma	100.00	0.03	3350	335	Gamma	100.00	0.03	NHS national cost collection
PCNL	LB75B; PCNL with CC score 0–1 – Elective inpatient	8285	828	Gamma	100.00	0.01	8285	828	Gamma	100.00	0.01	NHS national cost collection
Probability no residual fragments 12 weeks after treatment												
ESWL		0.368		Beta	84	164						PUrE
PCNL							0.753		Beta	61	20	PUrE
RR	FURS vs. ESWL for RCT1 and FURS vs. PCNL for RCT2 no residual fragments FURS	1.99	0.21	Normal	1.99	0.21	0.655	0.072	Normal	0.655	0.072	PUrE
Probability residual fragments ≤ 4 mm												
ESWL	Probability having residual fragments ≤ 4 mm given that fragments were left	0.188		Beta	27	117						PUrE
FURS		0.154		Beta	9	50	0.139		Beta	5	31	PUrE
PCNL		0.55					0.55		Beta	11	9	PUrE
Rate of pre-surgical events	Applied to wait for index surgery Markov state	0.065					0.087					PUrE
Utilities												
Waiting for index surgery	Baseline EQ5D score pooled for both trial arms	0.793	0.01	Beta	1300.92	339.59	0.767	0.022	Beta	282.44	85.80	PUrE
Stone clear		0.965	0.015	Beta	143.89	5.22	0.965	0.015	Beta	143.89	5.22	PUrE
Residual fragments ≤ 4 mm		0.965	0.015	Beta	143.89	5.22	0.965	0.015	Beta	143.89	5.22	PUrE
Residual fragments > 4 mm		0.916	0.021	Beta	158.90	14.57	0.916	0.021	Beta	158.90	14.57	PUrE

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 80 Decision model parameter values and probability distributions (*continued*)

		RCT1					RCT2					
				Probability distribution					Probability distribution			
Variable name	Description	Mean value	SE (mean)	Type	Param 1	Param 2	Mean value	SE (mean)	Type	Param 1	Param 2	Source
Utility post treatment												
ESWL												*PUrE*
Pre treatment		0.808	0.016	Beta	488.8	116.16						PUrE
Week 1	After treatment	0.789	0.016	Beta	512.3	137.00						PUrE
Week 2		0.814	0.016	Beta	480.6	109.82						PUrE
Week 3		0.816	0.018	Beta	377.3	85.08						PUrE
Week 4		0.819	0.018	Beta	373.9	82.63						PUrE
Week 5		0.84	0.016	Beta	440.2	83.84						PUrE
Week 6		0.839	0.016	Beta	441.9	84.79						PUrE
Week 7		0.831	0.018	Beta	359.4	73.08						PUrE
Week 8		0.837	0.018	Beta	351.6	68.47						PUrE
Week 9		0.826	0.019	Beta	328.0	69.10						PUrE
Week 10		0.83	0.018	Beta	360.6	73.86						PUrE
Week 11		0.84	0.019	Beta	311.9	59.41						PUrE
Week 12		0.829	0.019	Beta	324.7	66.98						PUrE
FURS												*PUrE*
Pre treatment		0.808	0.016	Beta	488.84	116.16	0.763	0.028	Beta	175.22	54.43	PUrE
Week 1	After treatment	−0.106	0.023	Normal	−0.106	0.023	−0.114	0.036	Normal	−0.114	0.036	PUrE
Week 2		−0.008	0.02	Normal	−0.008	0.02	−0.092	0.04	Normal	−0.092	0.04	PUrE
Week 3		0.028	0.022	Normal	0.028	0.022	−0.092	0.032	Normal	−0.092	0.032	PUrE
Week 4		0.034	0.021	Normal	0.034	0.021	−0.099	0.04	Normal	−0.099	0.04	PUrE
Week 5		0.026	0.02	Normal	0.026	0.02	−0.085	0.039	Normal	−0.085	0.039	PUrE
Week 6		0.045	0.019	Normal	0.045	0.019	−0.029	0.032	Normal	−0.029	0.032	PUrE
Week 7		0.05	0.015	Normal	0.05	0.015	−0.026	0.038	Normal	−0.026	0.038	PUrE
Week 8		0.057	0.02	Normal	0.057	0.02	−0.041	0.044	Normal	−0.041	0.044	PUrE
Week 9		0.06	0.019	Normal	0.06	0.019	−0.038	0.039	Normal	−0.038	0.039	PUrE
Week 10		0.052	0.02	Normal	0.052	0.02	−0.029	0.037	Normal	−0.029	0.037	PUrE
Week 11		0.047	0.021	Normal	0.047	0.021	−0.027	0.038	Normal	−0.027	0.038	PUrE
Week 12		0.05	0.016	Normal	0.05	0.016	−0.008	0.04	Normal	−0.008	0.04	PUrE

TABLE 80 Decision model parameter values and probability distributions (*continued*)

		RCT1					RCT2					
				Probability distribution					Probability distribution			
Variable name	Description	Mean value	SE (mean)	Type	Param 1	Param 2	Mean value	SE (mean)	Type	Param 1	Param 2	Source
PCNL												
Pre treatment							0.763	0.028	Beta	175.22	54.43	PUrE
Week 1	After treatment						0.678	0.029	Beta	175.32	83.27	PUrE
Week 2							0.763	0.026	Beta	203.34	63.16	PUrE
Week 3							0.818	0.024	Beta	210.61	46.86	PUrE
Week 4							0.839	0.025	Beta	180.49	34.64	PUrE
Week 5							0.878	0.022	Beta	193.44	26.88	PUrE
Week 6							0.836	0.031	Beta	118.43	23.23	PUrE
Week 7							0.807	0.038	Beta	86.24	20.62	PUrE
Week 8							0.843	0.03	Beta	123.13	22.93	PUrE
Week 9							0.872	0.028	Beta	123.27	18.10	PUrE
Week 10							0.864	0.032	Beta	98.28	15.47	PUrE
Week 11							0.853	0.034	Beta	91.67	15.80	PUrE
Week 12							0.851	0.03	Beta	119.04	20.84	PUrE

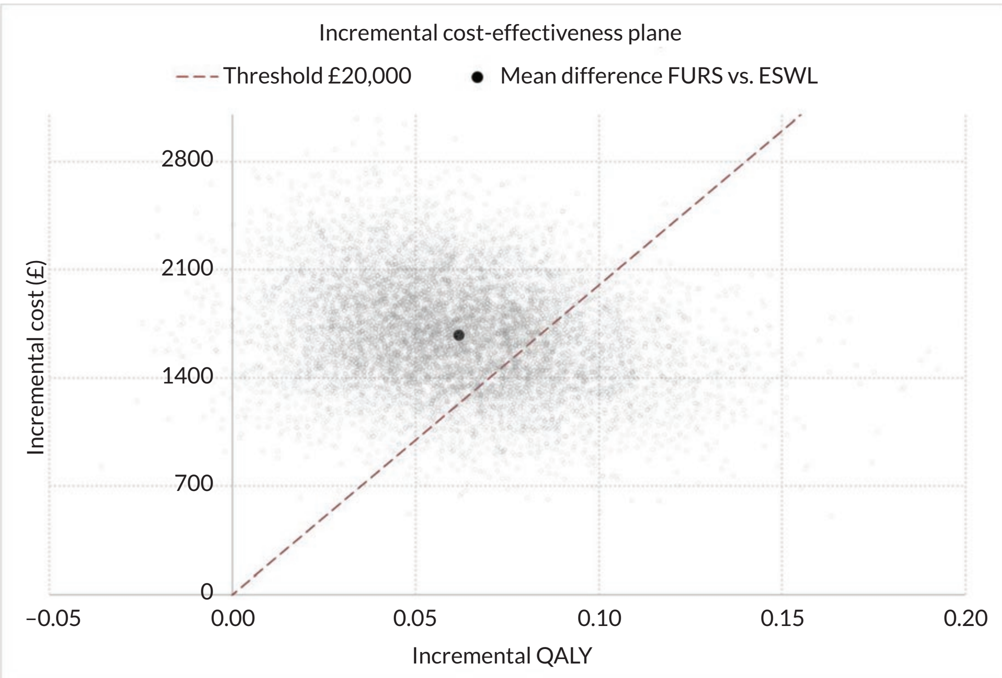


FIGURE 34 Cost-effectiveness scatterplot. Markov model – base-case analysis RCT1: FURS vs. ESWL.

Copyright © 2025 Wiseman *et al*. This work was produced by Wiseman *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

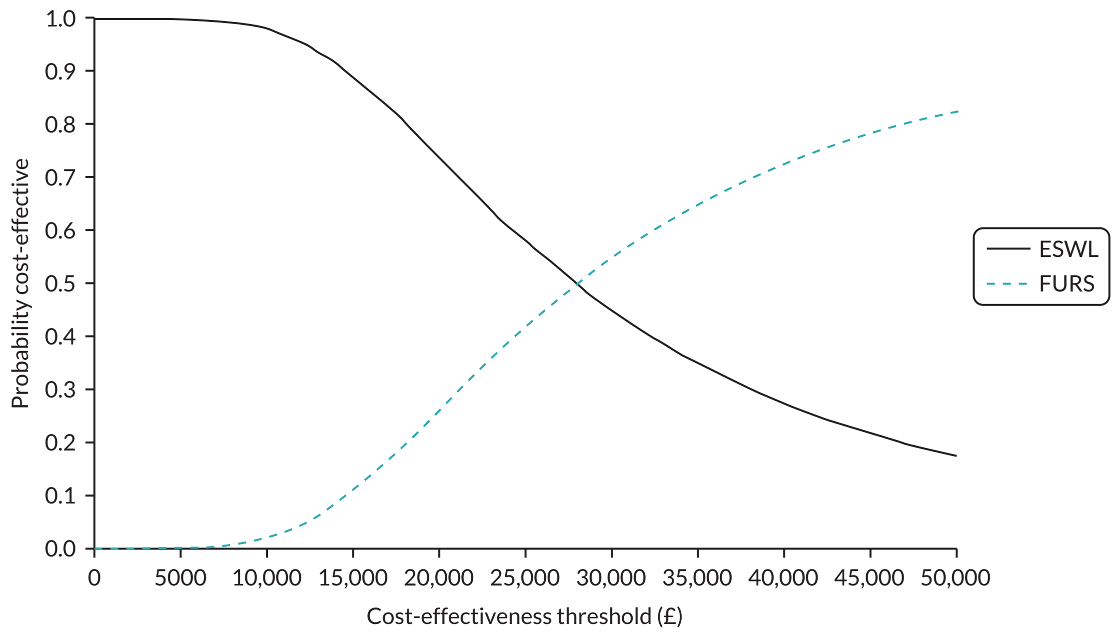


FIGURE 35 Cost-effectiveness acceptability curves. Markov model – base-case analysis RCT1: FURS vs. ESWL.

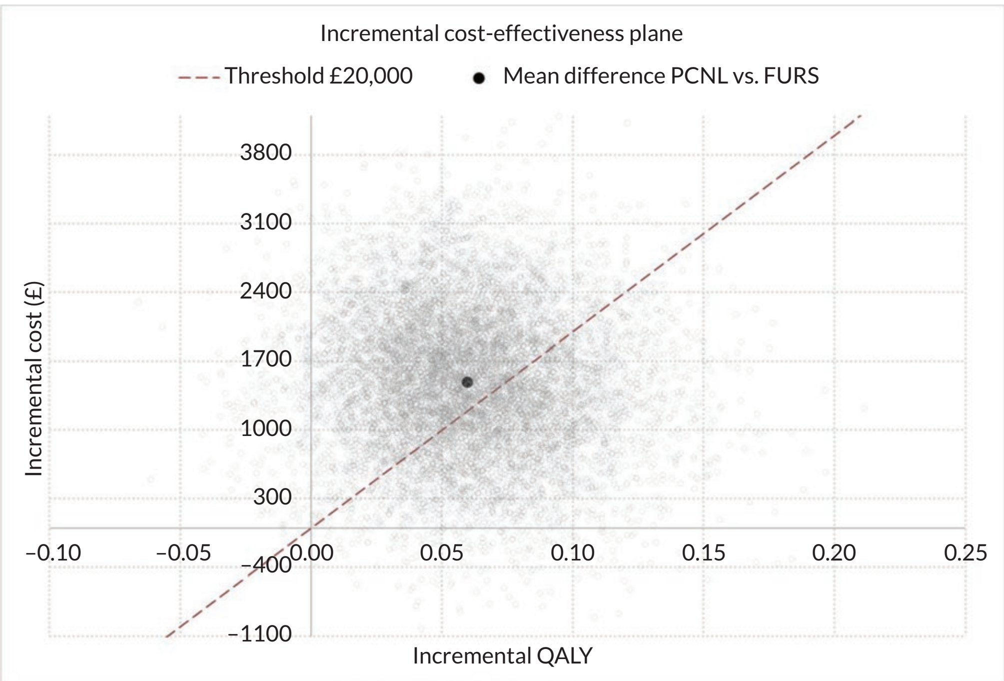


FIGURE 36 Cost-effectiveness scatterplot. Markov model – base-case analysis RCT2: FURS vs. PCNL.

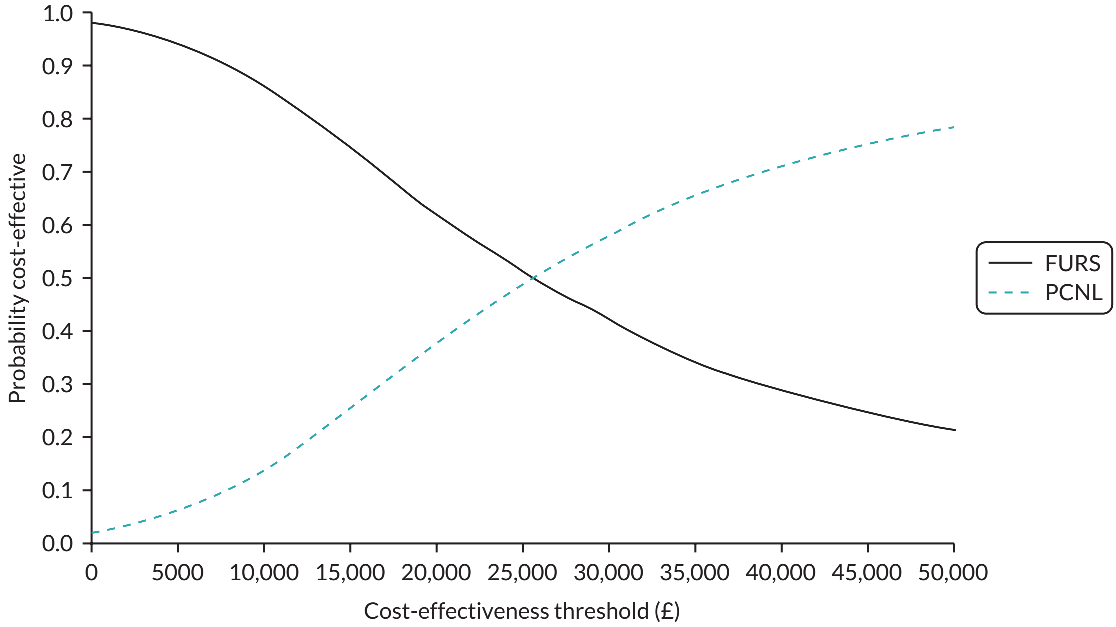


FIGURE 37 Cost-effectiveness acceptability curves. Markov model – base-case analysis RCT2: FURS vs. PCNL.

RCT2

Flexible ureterorenoscopy vs. percutaneous nephrolithotomy

Table 81 reports the results for the base-case analysis for RCT2 when micro-costing was used to cost the initial trial intervention. ITT with PCNL is less costly than ITT with FURS for this analysis; therefore, PCNL is reported first in *Table 81*. For every time horizon, FURS is more costly and produces fewer QALYs than PCNL. This means that FURS is dominated by PCNL and that PCNL is the cost-effective option. The probabilistic analysis results show the probability of PCNL being cost-effective above 80% for all the model runs and thresholds considered.

TABLE 81 One-way sensitivity analysis on Markov model time horizon. Micro-costing. RCT2 FURS vs. PCNL

Intervention	Mean total cost (£)	Δ Cost (£)	Mean total QALYs	Δ QALY	ICER	Probability cost-effective		
						£13,000	£20,000	£30,000
Time horizon 12 months								
PCNL	4565		0.775			0.90	0.87	0.81
FURS	5298	733	0.773	−0.001	−692,724	0.10	0.13	0.19
Time horizon 24 months								
PCNL	5768		1.632			0.89	0.89	0.89
FURS	6246	478	1.611	−0.021	−23,033	0.11	0.11	0.11
Time horizon 36 months								
PCNL	6475		2.502			0.88	0.91	0.93
FURS	6782	308	2.466	−0.036	−8451	0.12	0.09	0.07
Time horizon 48 months								
PCNL	7033		3.342			0.89	0.93	0.95
FURS	7295	262	3.292	−0.050	−5256	0.11	0.07	0.05

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 81 One-way sensitivity analysis on Markov model time horizon. Micro-costing. RCT2 FURS vs. PCNL (*continued*)

Intervention	Mean total cost (£)	Δ Cost (£)	Mean total QALYs	Δ QALY	ICER	Probability cost-effective £13,000	£20,000	£30,000
Time horizon 60 months								
PCNL	7534		4.150			0.89	0.93	0.95
FURS	7713	179	4.089	−0.061	−2933	0.11	0.07	0.05
Same utility weight for RF size below or above 4 mm								
PCNL	7534		4.162			0.74	0.77	0.80
FURS	7713	179	4.140	−0.021	−8464	0.26	0.23	0.20

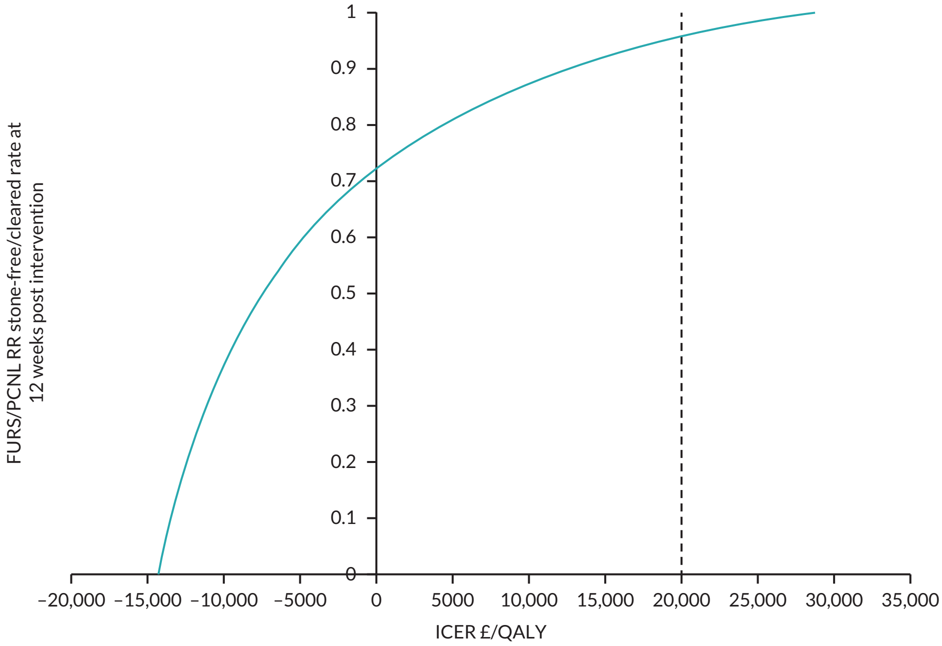


FIGURE 38 One-way sensitivity analysis varying SFR RR between FURS and PCNL (< 1 PCNL better). Markov model. Micro-costing used for index intervention for trial follow-up (12-monthly cycles).

Appendix 10 Summary of update of systematic review for PUrE Health Technology Assessment monograph

Context

The prevalence of urolithiasis is increasing. LPS are the commonest renal calculi.

Objective

To update our original review comparing SFRs, unplanned procedure rate, retreatment rate, complications, convalescence, QoL and economic outcomes for SWL, RIRS and PNL and to incorporate the results from the PUrE RCT.

Methods

Search strategy

Databases searched were MEDLINE and MEDLINE-in-process (1 April 2014 – 30 June 2022), EMBASE (1 April 2014 – 30 June 2022), Conference Proceedings Citation Index – Science (1 January 2014 – 30 June 2022), Latin American and Caribbean Health Sciences (LILACS) (June 2022), the International HTA database (1 January 2014 – 30 June 2022), Database of Abstracts of Review of Effects (1 January 2014 – 31 December 2015, last update), and the Cochrane Central Register of Controlled Trials and Cochrane Database of Systematic Reviews (1 January 2014 – 30 June 2022). No language restrictions were applied.

Types of studies included

We included RCTs and quasi-RCTs. Non-randomised studies were excluded.

Types of participants

Adults (≥ 18 years) with renal stones ≤ 20 mm in the lower pole calyx were included. Accepted diagnostic imaging modalities included ultrasound, radiograph/nephrotomogram, intravenous urography or computerised tomography. Studies including renal stones from any location but not reporting results for LPS separately were excluded. Other exclusions were anatomical abnormalities such as horseshoe kidney, transplanted kidney, polycystic kidney disease, calyceal diverticular calculi and pregnant women.

Types of interventions and comparators

The following interventions were compared in any combination:

- SWL
- FURS
- PNL.

All procedure protocols including all types/generations of SWL machines and any size/type of PCNL were included.

In this summary we focus on the comparisons and stone sizes which were included in PUrE RCT1 (ESWL vs. FURS, ≤ 10 mm) and RCT2 (FURS vs. PCNL 10–≤ 25 mm).

Types of outcome measures

The outcome measures of effectiveness are outlined below.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Primary outcome
Stone-free rate (up to 3 months).

Secondary outcomes
Operative duration
Duration of hospital stay
Complications (as defined by trialists)
Time to return to usual activities (e.g. driving, work)
Presence of clinically insignificant residual fragments
Quality of life
Need for unplanned procedures (e.g. ureteric stenting, nephrostomy insertion)
Retreatment rate (in excess of the protocol's number of sessions; same or alternative modality)

Economic analysis
Analgesic use
Cost of intervention
Resources implication
Cost–utility
Renal scarring

Assessment of risk of bias
Risk of bias was assessed using the Cochrane RoB assessment tool for RCTs[95] by two reviewers independently (SM and/or SMcC, OW, ZG, DS). Any disagreements were resolved by discussion or by arbitration with a third reviewer.

Assessment of the certainty of evidence
The GRADE system was used to assess the certainty of evidence[96] for four outcomes judged to be of critical importance based on consultations with patients and expert clinicians. Two reviewers undertook assessment (SM and SMcC). Any disagreements were resolved by discussion or arbitration.

Data collection and analysis
Two reviewers independently screened titles, abstracts, and full texts against inclusion criteria. Disagreements were resolved by arbitration. Data were extracted on study design, characteristics of participants and interventions, and outcome measures. Outcomes were abstracted from included studies in the form of total numbers and proportions for dichotomous outcomes, and total numbers, means and SDs, or median and interquartile range, for continuous outcomes. Dichotomous outcomes were estimated using risk ratios with 95% CIs. Continuous outcomes were estimated using (weighted) mean difference and 95% CI. We used random effects models throughout because clinical heterogeneity was suspected.

Heterogeneity between studies was assessed using the χ^2 test for heterogeneity and the I^2 statistic. Analysis was performed using the Cochrane RevMan software [The Cochrane Collaboration. Review Manager (RevMan) Version 5.4 2020] and GRADEpro [McMaster University and Evidence Prime. GRADEpro Guideline Development Tool. 2022].

Results

We screened 1348 titles and abstracts and 44 full texts, ultimately including 15 RCTs more in this update, as outlined in the PRISMA diagram shown in *Figure 39*.

Our original systematic review in 2015 before the PURE trial began included 5 RCTs comparing ESWL versus FURS, and 1 comparing FURS versus PCNL, and in our update we have included a further 15 studies, 7 of which compared ESWL versus FURS and 8 compared FURS versus PCNL (inclusive of PURE 1 and 2, and note that some studies include all 3 interventions in 1 report).

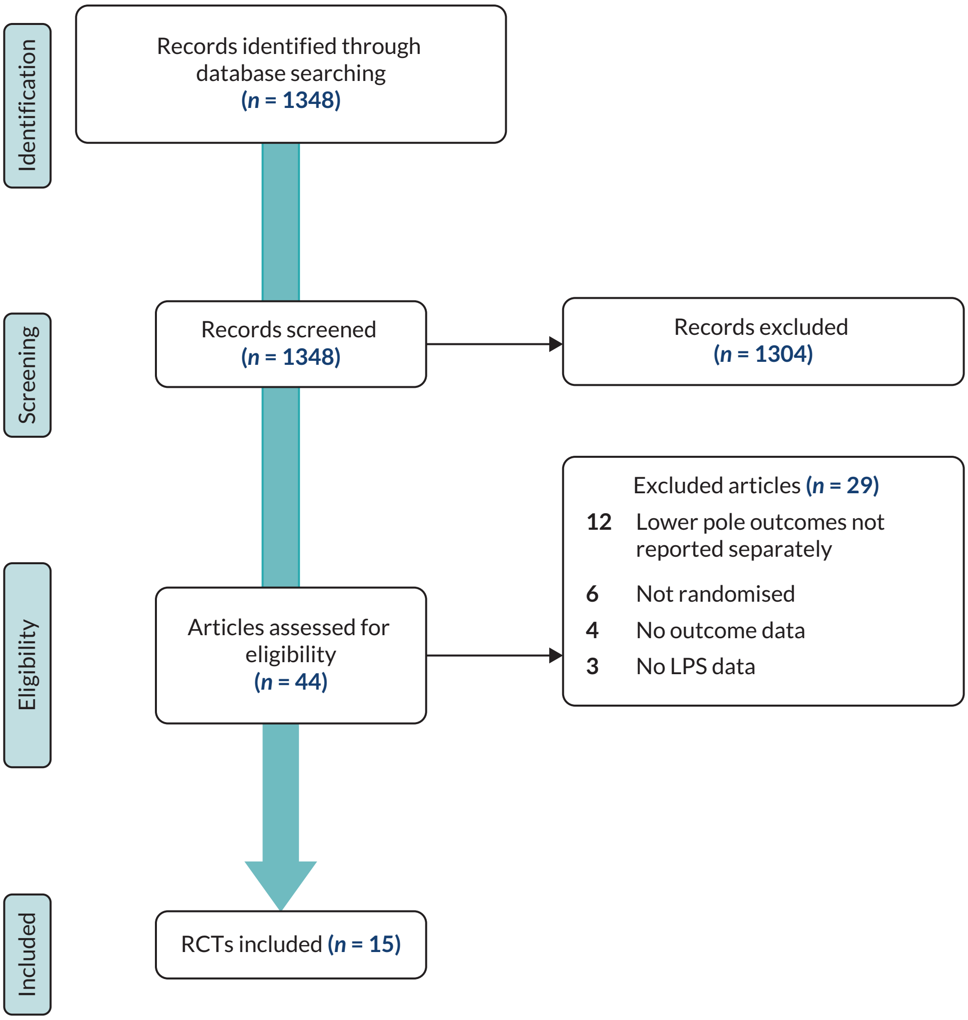


FIGURE 39 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.

The baseline characteristics are summarised in *Table 82* and treatment protocols in *Table 83*.

Risk of bias

The RoB is summarised in *Figure 40*. Generally, across the included studies, selection biases were judged as unclear or high in around half because they were not well reported or used methods of sequence generation or allocation concealment potentially open to bias. Blinding of trial personnel and participants was not possible given the intervention types so there was a high RoB throughout. Blinding of outcome assessors was possible but was judged as mostly as unclear for all outcomes indicating an uncertain RoB. Attrition was minimal and where it occurred appeared balanced between study arms indicating low RoB. Selective outcome reported was judged to be a low RoB overall. Other biases were frequently unknown.

Results

Extracorporeal shockwave lithotripsy versus flexible ureterorenoscopy

The results for all outcomes for each study are shown in *Table 84*. Forest plots are shown for outcomes where meta-analyses were done and for some outcomes to show estimates across studies but with no pooled estimate.

Stone-free rate at 3 months

The SFR at 3 months was reported in 11 studies,[30,32–34,36,37,80,85,86,98,99] and 5 of these were included in our original review.[30,32–34,36]

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 82 Baseline characteristics of included RCTs

Study ID, country, recruitment period	Intervention	*N* patients randomised (men)	Drop-outs (SFR)	Age/years: mean, range (SD)	BMI/kg/m^2: mean (SD)	FU/m	Modality	Stone sizes included[a]/mm, multiplicity	Stone size (mm): mean (SD) range
Pearle 2008, USA, 2000–3[30]	FURS	35 (17)	3	49.3, NR, (14.2)	28.1, 6.6	3	CT	≤ 10 mm Single	35.9 mm^2, 18.4
	ESWL	32 (19)	6	52.5, N/R, (12.3)	26.9, 8.1				42.2 mm^2, 21.2
Salem 2013, Egypt, 2010–2[33]	FURS	30 (N/R)	N/R	44.2	N/R	3	XR	≤ 20 mm Single or Multiple	11.5
	ESWL	30 (N/R)		35.5					11.3
Kumar 2013, N/R, 2011–2[32]	FURS	90 (N/R)	0	N/R	N/R	3	N/R	< 20 mm Single	12.3
	ESWL	90 (N/R)	0						12.1
Sener 2014, Turkey, 2012[34]	FURS	70 (41)	0	45.4 (6.4)	N/R	3[c]	XR + US[d]	< 10 mm Single	7.8 (1.3)
	ESWL	70 (31)	0	42.9 (5.6)					8.2 (1.2)
Singh 2014, India, 2011–3[36]	FURS	35 (22)	0	37.7 (11.8)	23.5 (3.3)	1	XR + US	10–20 mm Single	15.05[e], 3.56 mm
	ESWL	35 (20)	0	34.5 (13.1)	22.7 (4.35)				16.45[e], 2.28 mm
Kuo 2003, USA, NR[35]	FURS	13 (N/R)	0	N/R	N/R	3	CT	11–25 mm Single	N/R
	PCNL	15 (N/R)	0						
Kumar 2014, India, 2012–3[97]	Mini PNL	53	12	33.7 (1.6)	23.5 (1.2)	3	CT	10–20 mm Single	13.3 (1.3)
	SWL	52	10	33.1 (1.3)	23.4 (1.2)		US		13.2 (1.2)
	RIRS	53	10	33.1 (1.4)	23.6 (1.1)		US + CT		13.1 (1.1)
Sener 2015, Turkey, 2011–4[98]	FURS	50	N/R	36.8 (11.7)	N/R	3	CT	< 10 mm Single	8.2 (1.2)
	ESWL	50	N/R	34.5 (11.0)	N/R				7.9 (1.1)
Vilches, 2015 Chile, 2009–0[99]	FURS	31	7	43.7 (9.2)	26.6 (1.5)	2	CT	< 15 mm Single	9.7 (0.5)
	ESWL	32	1	45.6 (13.7)	27.7 (3.1)				9.6 (0.6)
Bozzini 2017, Italy, 2010–4[86]	PCNL	206	25	54.8 (7.2)	N/R	3	CT	10–20 mm Single	15.23 (3.3)
	ESWL	217	23	53.3 (14.8)					13.78 (3.1)
	FURS	226	19	55.8 (16.1)					14.82 (2.7)
Fayad 2017, Egypt, 2012–5[89]	Min PCNL	60	5	37.2 (9.2)	N/R	3	CT	< 20 mm Single	14.7 (3), 8–20
	FURS	60	9	37.7 (9.8)					14.1 (3), 8–20
Kandemir 2017, Turkey, 2013–5[100]	Mic PCNL	30	0	49.7, 1–78	N/R	3	CT	≤ 15 mm Single	10.6, 5–15
	FURS	30	0	51.8, 21–81					11.5. 7–15
Zeng 2018, China, India, Turkey, 2015–7[87]	SMin PCNL	80	3	49.4 (12.8)	24.6 (4.1)	3	KUB and US on day 1; CT at 3 m	10–20 mm	15 (2.9)
	FURS	80	4	47.1 (13.9)	24.1 (3.0)				14.3 (3.4)

TABLE 82 Baseline characteristics of included RCT (*continued*)

Study ID, country, recruitment period	Intervention	*N* patients randomised (men)	Drop-outs (SFR)	Age/years: mean, range (SD)	BMI/kg/m^2: mean (SD)	FU/m	Modality	Stone sizes included[a]/mm, multiplicity	Stone size (mm): mean (SD) range
Jiang 2019, China, 2013–7[88]	Mic PCNL	58	1	43.4 (11.6)	23.9 (4.5)	3	CT/IVU	< 20 mm Single	16.12 (3.0)
	FURS	58	2	45.4 (11.2)	24.1 (4.3)				15.2 (3.2)
Jin 2019, China, 2017–9[92]	Min PCNL	110	4	53.2 (13.7)	24.8 (3.8)	3	CT	10–20 mm Single	14.9 (3.9)
	FURS	110	0	51.4 (11.9)	25.3 (4.2)				16.4 (3.3)
Zhang 2019, China, 2015–7[85]	Min PNL	60	N/R	48.9 (11.1)	24.3 (3.0)	3	CT	10–20 mm Single	15.5 (2.5)
	SWL	60	N/R	50.5 (12.6)	24.3 (3.0)				14.9 (2.9)
	RIRS	60	N/R	50.1 (11.9)	24.3 (3.1)				14.6 (2.7)
Yavuz 2020, Turkey, 2017[90]	FURS	35	2	48.1 42.8–49.2 (13.1)	25.4 (2.8)	3	CT	10–20 mm Single	Surface area 401 (85) mm^2 424 – 435
	PCNL	140(4x35)	5						
Bosio 2022, Italy, 2015–8[80]	FURS	74 (29 LPS)	4	53 (12.8)	25.97 (3.4)	3	US and KUB	6–20 mm Single	11.5 (3.4)
	ESWL	74 (25 LPS)	6	51 (14.6)	26.2 (3.6)				10.8 (3.3)
PURE RCT1 2023, UK, 2016–21	FURS	230	16	54.1 (14.9)	29.1 (5.4)	3	CT/no further treat	≤ 10 mm	7.1 (1.9)
	ESWL	231	10	52.8 (14.2)	29.3 (5.5)				7.0 (2.0)
PURE RCT2 2023, UK, 2016–21	FURS	73	3	57.1 (13.3)	30.4 (6.9)	3	CT/no further treat	(> 10 mm, ≤ 25 mm)	13.9 (2.9)
	PCNL	86	5	60.3 (12.0)	31.4 (6.3)	3			14.5 (3.2)

CT, computerised tomography; Mic, micro; Min, mini; N/R, not reported; SMin, supermini; XR, X-Ray; Umin, ultramini.
a Greatest stone diameter.
b Personal communication.
c One-week SFRs were also reported; but are not analysed.
d Two patients underwent CT (radio-opaque stone and equivocal ultrasound).
e Approached significant difference (p = 0.0542).

Note
Blue shading: 2015 systematic review. Pink shading: 2023 update.

TABLE 83 Summary of intervention protocols used in included RCTs

	Intervention	*N* of sessions/stages	Energy used[a]	Procedural aspects (per session)
Pearle 2008[30]	FURS	N/R	Ureteral access sheath, intact stone retrieval and stent placement were at investigator's discretion	
	ESWL	N/R	EHL, EML, Piezo-electric[a]	Discretion of investigator; goal < 3 mm
Salem 2013[33]	FURS	N/R	365 um laser fibre	N/A
	ESWL	N/R	N/R	N/R
Kumar 2013[32]	FURS	N/R	Holmium laser	6.8/8.8Fr dual channel FURS
	ESWL	N/R	EML[a]	N/R

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 83 Summary of intervention protocols used in included RCTs (*continued*)

	Intervention	*N* of sessions/ stages	Energy used[a]	Procedural aspects (per session)
Sener 2014[34]	FURS	1	Holmium laser, 11–13F access sheath, displaced (upper pole/pelvis) for fragmentation, goal < 3 mm	
	ESWL	≤ 3 (mean 2.7)	EHL[a]	2500–3000 shocks at 14–17 kV; goal < 3 mm
Singh 2014[36]	FURS	1	Holmium laser, 7.5 French FLEX-X2 (Karl Storz), access sheath, displaced (upper pole/pelvis) for fragmentation, stent placed for 3 weeks routinely	
	ESWL	≤ 3	EML[a]	3500–4500 shocks (initially 60 Hz/level 1): increased each 200 shocks (≤ 120 Hz/level4). IV sedation.
Bosio 2022[80]	FURS	1	Holmium	Flex-X2, access sheath, stone basket
	ESWL	1	EHL	4000 shocks max
Vilches 2015[99]	FURS	1	Holmium 200 um	12–14 Ch Access sheath, Laser at 1.9J 4 Hz. Olympus URF P3, Stone basket. All cases routinely stented
	ESWL	1	ESWL	Modulith Storz SLX, Stone localised using fluoroscopy, 3000 shocks max
Sener 2015[98]	FURS	1	Holmium laser	Flex-X2, access sheath
	ESWL	1–3	2500–3000 shocks	Multimedia Classic EHL
Bozzini 2017[86]	PCNL	1	Holmium laser	Stone basket for fragments
	ESWL	1	Centre dependent	2500 shocks max
	FURS	1	Holmium laser	Access sheath, flexi URS, stone basket
Kumar 2014[97]	M PCNL	1	Pneumatic	MiniPCNL, 18Fr track
	ESWL	1 (1–4)	EML	2500 shocks max
	FURS	1	Holmium 200 um	8/9.8Fr ureteroscope, Stone basket
Zhang 2019[85]	UminiPCNL	1	365 um Holmium	13F UMP sheath
	FURS	1	220 um Holmium	Access sheath, URFP5 Olympus scope basket used to retrieve fragments. Stent placed
	ESWL	2	3000 shocks[a]	Dornier Compact S
PURE 2023	PCNL	1	Ultrasound, ballistic, combination or laser	Centre dependent
	FURS	2 (second to take place within 6 weeks)	Laser	Centre dependent
	ESWL	2 (second to take place within 8 weeks)	Centre dependent	Centre dependent
Kuo 2003[35]	FURS	N/R	N/R	N/A
	PCNL	N/R	N/R	N/R
Yavuz 2020[90]	FURS	1	Holmium 272 um	7.5F Flex X2 STORZ
	PCNL	1	Pneumatic, ultrasound or laser	Four different techniques used – standard, mini, ultra-mini and micro PCNL

TABLE 83 Summary of intervention protocols used in included RCTs (*continued*)

	Intervention	*N* of sessions/ stages	Energy used[a]	Procedural aspects (per session)
Bozzini 2017[86]	FURS	1	Holmium laser	Access sheath, flexi URS, stone basket
	PCNL	1	30 W Holmium YAG Laser	rigid nephroscope 20.8–24 Fr, kidney puncture was mainly performed under US guidance with a 'flash' X-ray check
Jiang 2019[88]	MicroPCNL	1	Holmium 200 um	Micro PCNL 16-gauge all-seeing needle
	FURS	1	Holmium 200 um	Access sheath for all 7.5Fr flexi ureteroscope. Stent as required
Kandemir 2017[100]	MicroPCNL	1	Holmium 200 um	Micro PCNL
	FURS	1	Holmium	Access sheath, Olympus URF-V, stone basket
Fayad 2017[89]	Mini PCNL	1	Holmium 365 um	16Fr sheath 10Fr Ureteroscope
	FURS	1	Holmium 365 um	Access sheath, 7.5Fr flexi ureteroscope
Jin 2019[92]	Mini PCNL	1	Holmium	16Fr sheath/URS
	FURS	1	Holmium 200 um	8/9.8Fr Wolf semirigid Then 12/14Fr Flexi scope Nitinol basket for fragments
Zeng 2018[87]	PCNL	1	Holmium 200 um	14Fr tract
	FURS	1	Holmium 200 um	12–14 access sheath, Olympus P5 or Flex X-2, stone repositioned to upper pole if possible, stone basket

EHL, electro-hydraulic lithotripsy; EML, electro-magnetic lithotripsy; IV, intra-venous; N/R, not reported; US, ultrasonic.

a Lithotripters used (where stated) were:

- Pearle 2008; Dornier T (Doli-S, HM3, MFL-5000), Healthtronics (Lithotron®), Siemens (Lithostar®, Lithostar Plus®), Storz Modulith®, Medstone STS, Medispec Econolith™.
- Kumar 2013: Dornier Alpha Compact.
- Sener 2014: Multimed Classic; Elmed, Ankara, Turkey.
- Singh 2014: Dornier (compact alpha).
- Zhang 2019: Dornier Compact S.
- Sener 2015: MULTIMED Classic – as for 2014 paper.
- Kumar 2015: Dornier Alpha Compact.
- Bosio 2022: HMT LithoTron® LSH171, HMT, Milan, Italy.
- Vilches 2015: STORZ Modulith SLX.

Note

Blue shading: 2015 systematic review. Pink shading: 2023 update.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcomes assessment? (detection bias) – SFR	Blinding of outcomes assessment? (detection bias) – Need for further interventions	Blinding of outcomes assessment? (detection bias) – Complications	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Bosio 2022	+	+	–	?	?	?	+	+	?
Bozzini 2017	+	?	–	?	?	?	+	?	?
Fayad 2017	?	–	?	?	?	?	+	?	?
Jiang 2019	?	?	–	?	?	?	+	?	?
Jin 2019	+	?	–	?	?	?	+	+	?
Kandemir 2017	+	?	–	–	–	–	+	+	?
Kumar 2013	?	?	–	?	?	?	+	+	?
Kumar 2014	+	?	–	?	?	?	+	+	?
Kumar 2015	+	?	–	?	?	?	+	+	?
Kuo 2003	?	?	–	?	?	?	+	+	–
McClinton 2023 (PUrE 1)	+	+	–	+	+	+	+	+	+
McClinton 2023 (PUrE 2)	+	+	–	+	+	+	+	+	+
Pearle 2008	+	?	–	?	?	?	+	+	?
Salem 2013	?	?	–	?	+	?	+	+	?
Sener 2014	–	?	–	?	?	?	+	?	?
Sener 2015	–	?	–	?	?	?	+	?	?
Singh 2014	+	?	–	–	?	?	?	?	?
Vilches 2015	+	?	–	?	?	?	+	?	?
Yavuz 2020	?	?	–	?	?	?	+	?	?
Zeng 2018	+	+	–	+	+	+	+	?	?
Zhang 2019	–	?	–	–	–	–	+	?	?

FIGURE 40 Risk of bias.

TABLE 84 Summary of results: FURS vs. ESWL

Study ID, Imaging, SFR definition	Outcome	Number of patients		%/no pts or mean (SD)		RR or MD (95% CI)[a]	*p*-value[a]
		FURS	ESWL	FURS	ESWL		
Pearle 2008[30] CT ≤ 4 mm	SFR (3 m) ≤ 10 mm	32	26	72/23	65/17	1.10 (0.77 to 1.57)	0.60
	Unplanned prcdrs	32	26	3.1/1	7.7/2	0.41 (0.04 to 4.23)	0.45
	Retreatment	32	26	3.1/1	11.5/3	3.69 (0.41 to 33.43)	0.25
	Complications: post op	33	30	21/7	23/7	0.91 (0.36 to 2.29)	0.84
	Procedure time/minute	N/R	N/R	90.4 (43.8)	66.5 (27.9)	Unavailable[b]	0.01[b]
	Hospital stay/day	35	32	0.06	0	Unavailable[b]	0.68[b]
Salem 2013[33] XR ≤ 3 mm	SFR (3 m) ≤ 20 mm	30	30	96.7/29	56.7/17	1.71 (1.24 to 2.35)	0.001
	Complications	30	30	16.7/5	23.3/7	0.71 (0.25 to 2.00)	0.52
Kumar 2013[32] NR Not defined	SFR ≤ 20 mm (3 m)	90	90	86.6/78	66.6/60	1.30 (1.10 to 1.54)	0.002
	SFR < 10 mm (3 m)	49	53	87.7/43	71.7/38	1.22 (1.00 to 1.49)	0.05
	SFR 10–20 mm (3 m)	41	37	85.4/35	59.5/22	1.44 (1.07 to 1.93)	0.02
	Unplanned prcdrs	90	90	17.7/16	21.1/19	0.84 (0.46 to 1.53)	0.57
	Retreatment	90	90	1.1/1	67.1/60	60.00 (8.50 to 423.64)	< 0.0001
	Complications	90	90	11.1/10	6.6/6	1.67 (0.63 to 4.39)	0.30
Sener 2014[34] XR+US ≤ 3 mm	SFR < 10 mm (3 m)	70	70	100/70	91.5/64	1.09 (1.01 to 1.18)	0.02
	Unplanned prcdrs	70	70	0/0	1.4/1	0.33 (0.01 to 8.04)	0.50
	Retreatment	70	70	0/0	8.6/6	13.00 (0.75 to 226.45)	0.08
	Complications	70	70	2.8/3	5.7/4	0.75 (0.17 to 3.23)	0.70
Singh 2014[36] XR+US ≤ 3 mm	SFR 10–20 mm (1 m)	35	35	85.7/30	54.3/19	1.58 (1.13 to 2.20)	0.007
	Unplanned prcdrs	35	35	0/0	5.7/2	0.20 (0.01 to 4.02)	0.29
	Retreatment	35	35	14.3/5	45/16	3.20 (1.32 to 7.78)	0.01
	Complications	35	35	31.4/11	48.6/17	0.65 (0.36 to 1.17)	0.15
Kumar 2014[97] CT ≤ 4 mm	SFR (3 m) 10–20 mm	43	42	86.1/37	73.8/31	1.17 (0.94 to 1.45)	0.1652
	Unplanned prcdrs	43	42	8.8/4	20.2/8	0.49 (016 to 1.50)	0.2108
	Retreatment	43	42	2.1/1	63.4/27	27.64 (3.93 to 194.31)	0.0008
	Complications	43	42	9.3/4	7.1/3	1.30 (0.31 to 5.47)	0.7183
	Hospital stay/days	43	42	1.3	0.13	Not estimable	N/A
	Procedure time/minutes	43	42	47.5 (1.1)	43.6 (1.4)	3.90 (3.36 to 4.44)	< 0.0001
Sener 2015[98] CT ≤ 3 mm	SFR (3 m) (< 10 mm)	50	50	92/46	92/46	1.00 (0.89 to 1.12)	1
	Retreatment	50	50	4	6 20 had more than 1 session	1.50 (0.45 to 4.99)	0.6963
	Complications	50	50	8/4	6/3	1.33 (0.31 to 5.65)	0.6963

continued

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 84 Summary of results: FURS vs. ESWL (*continued*)

Study ID, Imaging, SFR definition	Outcome	Number of patients		%/no pts or mean (SD)		RR or MD (95% CI)[a]	*p*-value[a]
Vilches 2015[99] CT ≤ 3 mm	SFR (2 m) < 15 mm	24	31	70.8/17	48.3/15	1.46 (0.94 to 2.28)	0.093
	Complications	24	31	37.25/9	16.1/5	2.33 (0.90 to 6.04)	0.0832
	Hospital stay/days	24	31	1	0	Not estimable	N/A
	Procedure time/minutes	24	31	59.6 (16.5)	42.7 (12.2)	16.90 (9.02 to 24.78)	< 0.0001
Bozzini 2017[86] CT ≤ 3 mm	SFR (3 m) 10–20 mm	207	194	82.1/170	61.8/120	1.33 (1.17 to 1.51)	< 0.0001
	Unplanned prcdrs	207	194	8.7/18	22.1/43	0.07 (0.04 to 0.14)	< 0.0001
	Retreatment	207	194	4.3/9	61.3 119	14.11 (7.38 to 26.99)	< 0.0001
	Complications	207	194	14.5/30	6.7/13	2.16 (1.16 to 4.02)	0.0148
	Hospital stay/days	207	194	1.3 (0.4)	0.12 (0.1)	1.18 (1.12 to 1.24)	< 0.0001
	Procedure time/minutes	207	194	55.8 (11.4)	40.9 (7.7)	14.90 (13.01 to 16.79)	< 0.0001
Zhang 2019[85] CT ≤ 3 mm	SFR (3 m) (10–20 mm)	60	60	92/55	73/44	1.25 (1.05 to 1.48)	0.0104
	Complications	60	60	10/6	6.67/4	1.50 (0.45 to 5.05)	0.5125
	Hospital stay/days	60	60	3.2 (0.5)	1.1 (0.3)	2.10 (1.95 to 2.25)	< 0.0001
	Procedure time/minutes	60	60	93.4 (21.6)	46.3 (5.8)	47.10 (40.59 to 53.61)	< 0.0001
Bosio 2022[80] CT No fragment	SFR (1 m) 6–20 mm	29	25	48.3/14	24/6	2.01 (0.91 to 4.44)	0.084
	SFR (6 m) 6–20 mm	29	25	37.5/21	48.3/14	1.29 (0.85 to 1.96)	0.2234
PURE 1 No imaging/ further Rx	SFR (3 m) ≤ 10 mm	230	231	71.7/165	36.4/84	1.97 (1.63 to 2.28)	< 0.001
	Additional intervention (unplanned procedures)	230	231	9.1/21	27.3/63	2.99 (1.89 to 4.73)	< 0.001
	Complications	230	231	20/48	13/30	0.62 (0.41 to 0.95)	0.03

m, month(s); NS, not significant; Prcdrs, procedures; RR, risk ratio; Rx, treatment.
a Risk ratios and 95% CIs and associated *p*-values were calculated from available numerators and denominators reported in primary studies where possible.
b Insufficient data were reported to enable calculation of these values. Where possible the reported *p*-value is stated. No studies reported risk ratios.
c Numerators were not reported in the primary study. These are estimates using reported percentages and denominators.

Note
Blue shading: 2015 systematic review. Pink shading: 2023 update.

One study in the original (Singh, 2014)[36] and two in the update (Bosio, 2017;[80] Vilches, 2015)[99] were excluded from the meta-analysis because they reported SFR at time points other than three months, but these estimates can still be seen in *Table 84*. We decided also to exclude PUrE1 from the meta-analysis because the focus of the study was on QoL and the definition of stone free was not imaging based (while it was in the other studies) and was instead inferred from the lack of further hospital visit/treatment. We also excluded Kumar, 2013[32] from the meta-analysis because no definition of SFR was provided. The trend across these studies finds in the same vein as the meta-analysis below, that FURS generally achieves a better SFR (see *Table 28* for estimates).

The meta-analysis of SFR is shown in *Figure 41*. Considering any stone size 1–20 mm, among seven RCTs that defined SFR as stone free plus fragments of either ≤ 3 mm or ≤ 4 mm, FURS is more effective than ESWL for SFR (RR 1.19, 95% CI 1.05 to 1.35). When restricting to ≤ 10 mm, that is, the inclusion criteria for PURE1, three RCTs providing data in this subgroup and the effect size is attenuated and not statistically significant (RR 1.06, 95% CI 1.00 to 1.13). Stone size likely accounts for most of the observed heterogeneity in the overall estimate (I^2 77% in main analysis but 0% in the subgroups). The certainty of evidence was moderate and was downgraded for RoB, this judgement is unchanged from the original review. The RRs in our update are similar to those seen our original review RR overall: 1.31, 95% CI 1.08 to 1.59; ≤ 10 mm subgroup (RR 1.11, 95% CI 1.03 to 1.19) but do not find as strongly in favour of FURS.

PUrE 1 found significantly better SFR at 3 months for FURS (72.1%) versus ESWL (36.4%), which is a pragmatic estimate in terms of what happens in a health system when not all patients are followed up with imaging.

Unplanned procedures

Our original review found that FURS was favoured in terms of unplanned procedures, across four studies but only statistically significantly in one, and the results were not pooled. The certainty of evidence was judged to be very low.

In our update, unplanned procedures were reported in a further three RCTs and we opted still not to meta-analyse these data due to uncertainty in the definition of unplanned procedures across studies and further heterogeneity in the size of stones included, and instead offer a narrative synthesis. No subgroups stratified by stone size were available. Considering these caveats, FURS appears more effective in terms of proportionally fewer unplanned procedures. The certainty of evidence remains very low due to heterogeneity in population and definitions and RoB. The estimates for each study are shown in *Table 84* and plotted in *Figure 42*.

Retreatment rate

Retreatment rate was reported in four trials in our original review. They were narratively synthesised and generally favoured FURS. There was heterogeneity in the estimates, likely due to differing thresholds for retreatment which factored into our decision not to meta-analyse these data. The certainty of evidence was very low.

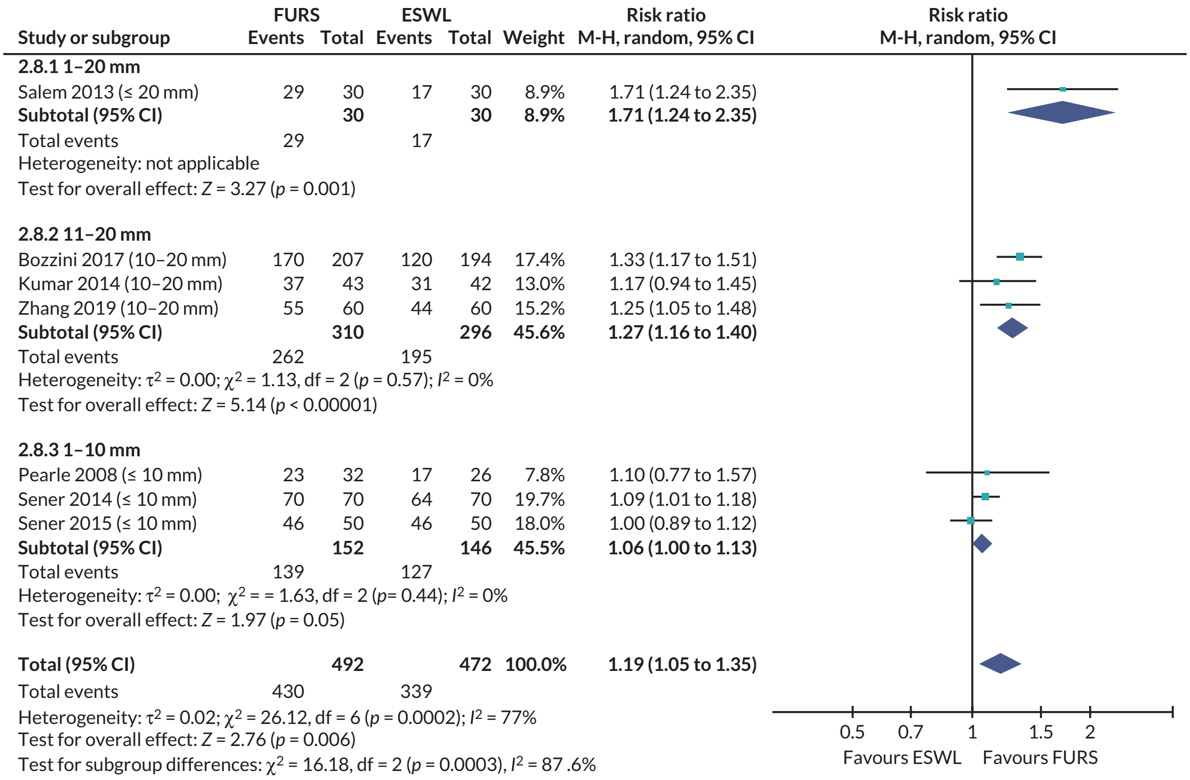


FIGURE 41 Extracorporeal shockwave lithotripsy vs. FURS SFR at 3 months.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Study or subgroup	ESWL Events	ESWL Total	FURS Events	FURS Total	Risk ratio M-H, random, 95% CI	Year
Pearle 2008	2	26	1	32	2.46 (0.24 to 25.66)	2008
Kumar 2013	19	90	16	90	1.19 (0.65 to 2.16)	2013
Singh 2014	16	35	5	35	3.20 (1.32 to 7.78)	2014
Kumar 2014	8	42	4	43	2.05 (0.67 to 6.29)	2014
Sener 2014	1	70	0	70	3.00 (0.12 to 72.40)	2014
Bozzini 2017	43	194	18	207	2.55 (1.52 to 4.26)	2017
PURE 1 2023 (≤ 10 mm)	63	231	21	230	2.99 (1.89 to 4.73)	2023

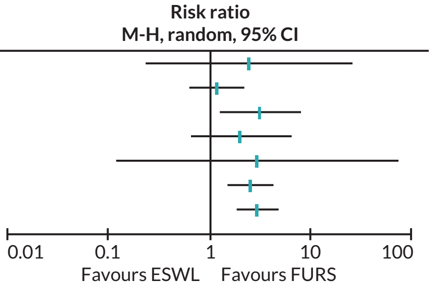


FIGURE 42 Extracorporeal shockwave lithotripsy vs. FURS unplanned procedures.

There are now an additional three RCTs available, but the same problems regarding outcome definition remain and meta-analysis was inappropriate. The estimates for each study are shown in *Table 84* and they range from RR 1.50 (95% CI 0.45 to 4.99) (Sener 2015)[34] to RR 60 (95% CI 8.50 to 423.64) (Kumar, 2013)[32] indicating between study heterogeneity and within study imprecision, as seen in *Figure 43*. These factors influence our judgement that the evidence for this outcome remains of very low certainty.

Complications

Complications were reported in five studies in our original review. There were conflicting findings likely due to heterogenous classification of complications and the certainty of evidence was judged as very low. We narratively synthesised these data instead of meta-analysing. That analysis decision remains unchanged.

In our update an additional 6 studies were included, giving 11 in total (see *Figure 44*). The proportion of complications were statistically significantly lower for ESWL compared to FURS in 2 of the newer studies (Bozzini, 2017[86] and PUrE1,[37] both with notably larger sample sizes than the other studies) but the problems with classification heterogeneity remain so it remains uncertain which intervention is more effective in terms of reducing the risk of complications. The certainty of evidence remains very low.

Study or subgroup	ESWL Events	ESWL Total	FURS Events	FURS Total	Risk ratio M-H, random, 95% CI	Year
Pearle 2008	3	26	1	32	3.69 (0.41 to 33.43)	2008
Kumar 2013	60	90	1	90	60.00 (8.50 to 423.63)	2013
Singh 2014	16	35	5	35	3.20 (1.32 to 7.78)	2014
Kumar 2014	27	42	1	43	27.64 (3.93 to 194.31)	2014
Sener 2014	6	70	0	70	13.00 (0.75 to 226.45)	2014
Sener 2015	6	50	4	50	1.50 (0.45 to 4.99)	2015
Bozzini 2017	119	194	9	207	14.11 (7.38 to 26.99)	2017

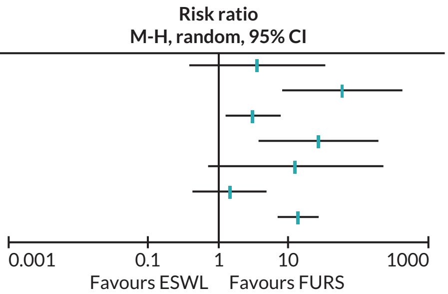


FIGURE 43 Extracorporeal shockwave lithotripsy vs. FURS retreatment rate.

Study or subgroup	ESWL Events	ESWL Total	FURS Events	FURS Total	Risk ratio M-H, random, 95% CI	Year
Pearle 2008	7	30	7	33	1.10 (0.44 to 2.77)	2008
Kumar 2013	6	90	10	90	0.60 (0.23 to 1.58)	2013
Salem 2013	7	30	5	30	1.40 (0.50 to 3.92)	2013
Sener 2014	4	70	3	70	1.33 (0.31 to 5.74)	2014
Singh 2014	17	35	11	35	1.55 (0.85 to 2.81)	2014
Kumar 2014	3	42	4	43	0.77 (0.18 to 3.23)	2014
Vilches 2015	5	31	9	24	0.43 (0.17 to 1.12)	2015
Sener 2015	3	50	4	50	0.75 (0.18 to 3.18)	2015
Bozzini 2017	13	194	30	207	0.46 (0.25 to 0.86)	2017
Zhang 2019	4	60	6	60	0.67 (0.20 to 2.24)	2019
PURE1 2023	30	231	48	230	0.62 (0.41 to 0.95)	2023

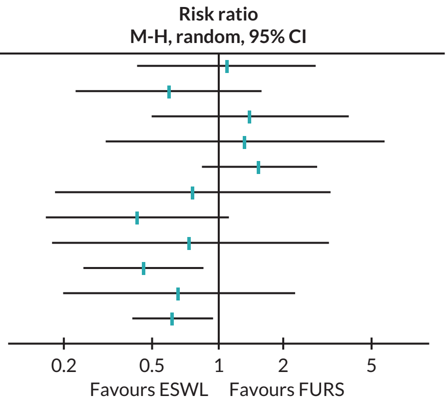


FIGURE 44 Extracorporeal shockwave lithotripsy vs. FURS complications.

Quality of life, satisfaction and economic outcomes

Quality of life was seldom reported in either the original review or the update. This was concluded to be an area for improvement in the original review and was part of the impetus to make QoL a PO in PUrE.

In our original review ESWL was favoured to FURS for QoL, convalescence time, less analgesic requirement, and willingness to undergo the procedure again in one study (Pearle, 2008)[30] while conversely another found higher satisfaction and willingness to undergo the procedure again with FURS (Singh, 2014).[36] However, these results could be influenced by differing treatment protocols regarding the number of sessions for ESWL and it was difficult to say anything definitive based on this heterogeneous low certainty evidence.

In our update, two RCTs provide data on QoL or satisfaction outcomes. One reports on patient satisfaction with the procedure (Bosio, 2022,[80] no clinically or statistically significant differences), and PUrE 1[37] showed a marginal and not clinically or statistically significant improvement with FURS in terms of QoL.

The evidence base on QoL remains inconclusive and of very low certainty due to RoB and heterogeneity of measurement.

Zhang (2018)[85] found FURS to be more than eight times the cost of ESWL in the Chinese healthcare system, and similarly PUrE 1[37] showed that ESWL was likely to be more cost-effective for the UK NHS. However, these estimates may not be generalisable outside of the country included in the study.

Hospital stay

Hospital stay was not clearly reported in the studies included in our original review. In our update two RCTS reported hospital stay and both clearly favoured ESWL (see *Figure 45*).

Procedure time

No conclusive results were found in our original review regarding operative duration likely due to heterogeneity in the treatment protocols. In our update the procedure time was shorter in all studies for ESWL compared to FURS but again heterogeneity in protocols hamper our ability to say anything meaningful on this outcome other than procedure time is probably shorter with ESWL (see *Figure 46*).

Flexible ureterorenoscopy versus percutaneous nephrolithotomy

The estimates per study are reported in *Table 85*. Forest plots are shown for outcomes where meta-analysis was done and for some outcomes to show estimates across studies but with no pooled estimate.

	ESWL			FURS			Mean difference
Study or subgroup	Mean (days)	SD (days)	Total	Mean (days)	SD (days)	Total	IV, random, 95% CI
Bozzini 2017	0.12	0.1	194	1.3	0.4	207	−1.18 (−1.24 to −1.12)
Zhang 2019	1.1	0.3	60	3.2	0.5	60	−2.10 (−2.25 to −1.95)

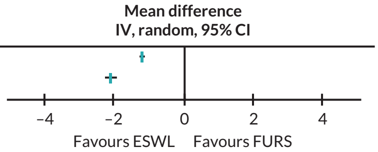


FIGURE 45 Extracorporeal shockwave lithotripsy vs. FURS hospital stay.

	ESWL			FURS			Mean difference	
Study or subgroup	Mean (minutes)	SD (minutes)	Total	Mean (minutes)	SD (minutes)	Total	IV, random, 95% CI	Year
Kumar 2014	43.6	1.4	42	47.5	1.1	43	−3.90 (−4.44 to −3.36)	2014
Vilches 2015	42.7	12.2	31	59.6	16.5	24	−16.90 (−24.78 to −9.02)	2015
Bozzini 2017	40.9	7.7	194	55.8	11.4	207	−14.90 (−16.79 to −13.01)	2017
Zhang 2019	46.3	14	60	93.4	21.6	60	−47.10 (−53.61 to −40.59)	2019

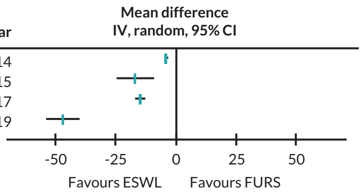


FIGURE 46 Extracorporeal shockwave lithotripsy vs. FURS procedure time.

Copyright © 2025 Wiseman *et al.* This work was produced by Wiseman *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 85 Summary of results: FURS vs. PCNL

Study ID Type of PCNL Imaging Definition	Outcome	Number of patients		%/no pts or mean (SD)		RR or MD (95% CI or SD*)	*p*-value[a]
		FURS	PCNL	FURS	PCNL		
Kuo 2003[35] CT Not defined	SFR (3 m) 11–25 mm	13	15	45.6/6[b]	66.7/10[b]	1.44 (0.73 to 2.87)	0.29
	Secondary treatment	N/R	N/R	25.0	9.1	Unavailable[c]	0.59[c]
	Complications	N/R	N/R	0.0	6.7	Unavailable[c]	0.999[c]
	Hospital stay (days)	N/R	N/R	0	2.8 (2.2)	Unavailable[c]	< 0.001[c]
	Procedure time (minutes)	N/R	N/R	125 (49)	111(38)	Unavailable[c]	NS[c]
	Recovery (days)	N/R	N/R	10.0 (7.7)	23.5 (20.5)	Unavailable[c]	$p < 0.05$[c]
Bozzini 2017[87] Standard PCNL CT ≤ 3 mm	SFR (3 m) 10–20 mm	207	181	82.1/170	87.3/158	1.06 (0.98 to 1.16)	0.1566
	Unplanned prcdrs	207	181	8.7/18	6.6/12	0.76 (0.38 to 1.54)	0.4493
	Retreatment	207	181	4.3/9	2.2/4	0.51 (0.16 to 1.62)	0.2532
	Complications	207	181	14.5/30	19.3/35	1.33 (0.86 to 2.08)	0.2040
	Hospital stay/days	207	181	1.3 (0.4)	3.7 (1.5)	2.40 (2.17 to 2.63)	0.0000
	Procedure time/minutes	207	181	55.8 (11.4)	40.9 (7.7)	−14.90 (−16.82 to −12.98)	0.0000
Fayad 2017[89] Mini PCNL CT ≤ 2 mm	SFR (3 m) ≤ 20 mm	51	55	84.3/43	92.7/51	1.10 (0.96 to 1.26)	0.1734
	Retreatment	51	55	5.9/3	3.6/2	1.62 (0.28 to 9.29)	0.5897
	Complications	51	55	9.8/5	9.1/5	0.93 (0.29 to 3.02)	0.9003
	Procedure time/minutes	51	55	109.7 (20.8)	71.7 (10.4)	−38.00 (−44.34 to −31.66)	0.0000
Kandemir 2017[100] Micro PCNL CT ≤ 4 mm	SFR (3 m) ≤ 15 mm	30	30	86.7/26	83.3/25	0.96 (0.78 to 1.19)	0.7180
	Complications	30	30	20/6	16.6/5	0.83 (0.28 to 2.44)	0.7392
	Procedure time/minutes	30	30	51.05	59.04	Not estimable	N/R
Zeng 2018[87] SMini PCNL	SFR (1d) 10–20 mm	80	80	71.2/57	91.2/73	1.28 (1.10 to 1.50)	0.0017
	SFR (3 m) 10–20 mm	80	80	82.5/66	93.8/75	1.16 (0.97 to 1.38)	0.0947
	Unplanned prcdrs	80	80	23.8/19	5/4	0.21 (0.07 to 0.59)	0.0031
	Complications (< 30 days)	80	80	1.25/1	1.25/1	1.00 (0.06 to 15.71)	1
	Complications (> 30 days)	80	80	8.75/7	8.75/7	1.00 (0.37 to 2.72)	1
	Procedure time/minutes	80	80	52.3 (22.4)	58.6 (21.6)	6.30 (−0.52 to 13.12)	0.0721
Jiang 2019[88] Micro PCNL CT ≤ 4 mm	SFR (3 m) ≤ 20 mm	58	58	79.3/46	84.5/49	1.07 (0.90 to 1.26)	0.4705
	SFR (6 m) ≤ 20 mm	56	57	92.9/52	94.7/54	1.02 (0.93 to 1.12)	0.6792
	Unplanned prcdrs	58	58	20.7/12	15.5/9	0.75 (0.34 to 1.64)	0.4720
	Retreatment	58	58	20.7/12	15.5/9	0.75 (0.34 to 1.64)	0.4720
	Complications	58	58	5.2/3	6.9/4	1.33 (0.31 to 5.70)	0.6978
	Hospital stay/days	58	58	3.2 (0.6)	3.2 (0.5)	0.00 (−0.20 to 0.20)	0.856
	Procedure time/minutes	58	58	60.3 (8.5)	54.0 (8.2)	−6.30 (−9.34 to −3.26)	0.0001

TABLE 85 Summary of results: FURS vs. PCNL (*continued*)

Study ID Type of PCNL Imaging Definition	Outcome	Number of patients		%/no pts or mean (SD)		RR or MD (95% CI or SD*)	*p*-value[a]
Jin 2019[92] Mini PCNL CT ≤ 3 mm	SFR (3 m) 10–20 mm	110	106	97.3/107	99.1/105	1.02 (0.98 to 1.06)	0.3277
	Complications	110	110	8.18/9	14.54/16	1.78 (0.82 to 3.85)	0.1445
	Hospital stay/days	110	110	3.15 (0.72)	5.59 (0.82)	2.44 (2.24 to 2.64)	0.007
	Procedure time/minutes	110	110	87.2 ± 13.34	79.6 (14.86)	−7.60 (−11.33 to −3.87)	0.124
Zhang 2019[85] UMini PCNL	SFR (3 m) 10–20 mm	60	60	92/55	98/59	1.07 (0.99 to 1.17)	0.0978
	Complications	60	60	10/6	17/10	1.67 (0.65 to 4.30)	0.2903
	Hospital stay/days	60	60	3.2 (0.5)	5.3 (1.2)	2.10 (1.77 to 2.43)	< 0.0001
	Procedure time/minutes	60	60	93.4 (21.6)	68.6 (15.8)	−24.80 (−31.57 to −18.03)	0.0000
Yavuz 2020[90] Various PCNL CT ≤ 3 mm	SFR (3 m) 10–20 mm	33	35 33 34 33	76/25	Mic 77/27 U 90.1/30 M 94.1/32 S 94/31	1.02 (0.78 to 1.33) 1.20 (0.96 to 1.50) 1.24 (1.01 to 1.53) 1.24 (1.00 to 1.53)	0.8930 0.1061 0.0433 0.0463
	Hospital stay/ days (median ± minimum–maximum)	33	35 33 34 33	1.0 (0–5.3)	Mic 1.5 (1–4) UM 2.0 (1–6) M 2.0 (1–14) S 3.0 (2–16)	Not estimable	< 0.001
	Complications (Clavien > 2)	33	35 33 34 33	6/2	Mic 8.5/3 UM 6/2 M 8.8/3 S 6/2	1.41 (0.25 to 7.94) 1.00 (0.15 to 6.68) 1.46 (0.26 to 8.16) 1.00 (0.15 to 6.68)	0.6937 1 0.6693 1
	Procedure time/minutes (Mean ± SD)	33	35 33 34 33	60.7 (13)	109 (31) 63.2 (21) 61.6 (18.5) 66.0 (14.4)	48.30 (37.11 to 59.49) 2.50 (−5.93 to 10.93) 0.90 (−6.74 to 8.54) 5.30 (−1.32 to 11.92)	< 0.0001 0.5630 0.8190 0.1215
PURE 2 2023 No imaging/ further Rx	SFR (3 m) 11–25 mm	73	86	35 (47.9)	61 (70.9)	0.25 (0.10 to 0.64)	0.004
	Additional intervention (unplanned procedures)	73	86	17 (23.3)	5 (5.8)	4.01 (1.55 to 10.33)	0.004
	Complications	73	86	17 (23.3)	17 (19.7)	1.18 (0.65 to 2.14)	0.54

m, month(s); Mic, micro; M, mini; NS, not significant; Prcdrs, procedures; RR, risk ratio; Rx, treatment; U, ultra Mini; S, Supermini.
a Risk ratios and 95% confidence intervals and associated *p*-values were calculated from available numerators and denominators reported in primary studies where possible.
b Insufficient data were reported to enable calculation of these values. Where possible the reported *p*-value is stated. No studies reported risk ratios.
c Numerators were not reported in the primary study. These are estimates using reported percentages and denominators.

Note
Blue shading: 2015 systematic review. Pink shading: 2023 update.

Stone-free rate at 3 months

In our original review, one RCT[35] was included but sample size was very small (13 FURS versus 15 PCNL), SFR was not defined, and the results were not statistically significant.

In our update, we found eight RCTs (including PUrE 2). There was variation in the types of PCNL included, some were 'standard', where others were 'mini', 'ultra-mini', 'super-mini', or 'micro' PCNL. There was also heterogeneity in the definition of SFR. We omitted PUrE 2 from the meta-analysis because our definition of SFR differed from the other studies in that it was pragmatic and related to additional hospital visits/treatment and not always imaging-based. Also,

Copyright © 2025 Wiseman *et al*. This work was produced by Wiseman *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

although Yavuz[90] included various PCNL types, we included only the estimate for standard PCNL in this meta-analysis of eight RCTs.

The results show that regardless of stone size PCNL is marginally more effective than FURS for SFR (RR 1.07, 95% CI 1.01 to 1.12) (see *Figure 47*). The certainty of evidence was judged as 'low' because there are concerns about RoB, heterogeneity in imaging types (or lack thereof) and defining the outcome (variously defining stone free plus fragments ≤ 2, ≤ 3, or ≤ 4 mm).

Study or subgroup	PCNL Events	PCNL Total	FURS Events	FURS Total	Weight	Risk ratio M-H, random, 95% CI	Year
3.1.1 1–25 mm							
Kuo 2003	10	15	6	13	0.6%	1.44 (0.73 to 2.87)	2003
Subtotal (95% CI)		**15**		**13**	**0.6%**	**1.44 (0.73 to 2.87)**	
Total events	10		6				
Heterogeneity: not applicable							
Test for overall effect: $Z = 1.05$ ($p = 0.29$)							
3.1.2 1–20 mm							
Fayad 2017	51	55	43	51	9.8%	1.10 (0.96 to 1.26)	2017
Jiang 2019	49	58	46	58	7.2%	1.07 (0.90 to 1.26)	2019
Yavuz 2020	31	33	25	33	5.1%	1.24 (1.00 to 1.53)	2020
Subtotal (95% CI)		**146**		**142**	**22.0%**	**1.12 (1.01 to 1.23)**	
Total events	131		114				
Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 1.28$, df = 2 ($p = 0.53$); $I^2 = 0\%$							
Test for overall effect: $Z = 2.24$ ($p = 0.03$)							
3.1.3 10–20 mm							
Zeng 2018	66	80	57	80	7.1%	1.16 (0.97 to 1.38)	2018
Jin 2019	105	106	107	110	29.2%	1.02 (0.98 to 1.06)	2019
Zhang 2019	59	60	55	60	18.1%	1.07 (0.99 to 1.17)	2019
Subtotal (95% CI)		**246**		**250**	**54.5%**	**1.06 (0.97 to 1.17)**	
Total events	230		219				
Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 7.62$, df = 2 ($p = 0.02$); $I^2 = 74\%$							
Test for overall effect: $Z = 1.25$ ($p = 0.21$)							
3.1.4 1–15 mm							
Kandemir 2017	25	30	26	30	5.0%	0.96 (0.78 to 1.19)	2017
Subtotal (95% CI)		**30**		**30**	**5.0%**	**0.96 (0.78 to 1.19)**	
Total events	25		26				
Heterogeneity: not applicable							
Test for overall effect: $Z = 0.36$ ($p = 0.72$)							
3.1.5 1–10 mm							
Bozzini 2017	158	181	170	207	17.9%	1.06 (0.98 to 1.16)	2017
Subtotal (95% CI)		**181**		**207**	**17.9%**	**1.06 (0.98 to 1.16)**	
Total events	158		170				
Heterogeneity: not applicable							
Test for overall effect: $Z = 1.42$ ($p = 0.16$)							
Total (95% CI)		**618**		**642**	**100.0%**	**1.07 (1.01 to 1.12)**	
Total events	554		535				
Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 13.42$, df = 8 ($p = 0.10$); $I^2 = 40\%$							
Test for overall effect: $Z = 2.44$ ($p = 0.01$)							
Test for subgroup differences: $\chi^2 = 2.47$, df = 4 ($p = 0.65$), $I^2 = 0\%$							

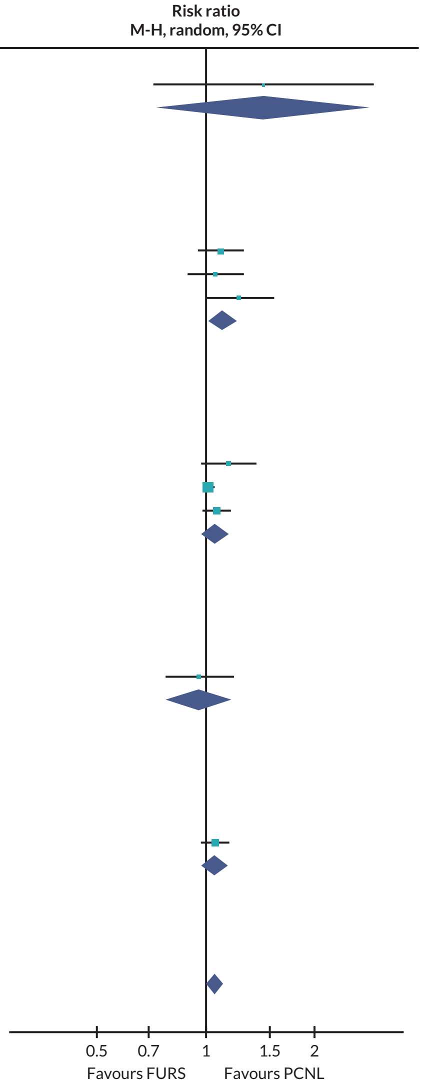


FIGURE 47 Flexible ureterorenoscopy vs. PCNL SFR at 3 months.

Unplanned procedures

Data on unplanned procedures were not available in our original review. In our update four RCTs reported on it but we did not meta-analyse due to heterogenous definitions (see *Figure 48*). The study level estimates statistically significantly favour PCNL in two studies and there is no evidence of a difference in two others. The certainty of evidence is very low because of concerns about RoB imprecise estimates.

Study or subgroup	FURS Events	FURS Total	PCNL Events	PCNL Total	Risk ratio M-H, random, 95% CI	Year
Bozzini 2017	18	207	12	181	1.31 (0.65 to 2.65)	2017
Zeng 2018	19	80	4	80	4.75 (1.69 to 13.34)	2018
Jiang 2019	12	58	9	58	1.33 (0.61 to 2.92)	2019
PURE 2 2023	17	73	5	86	4.01 (1.55 to 10.33)	2023

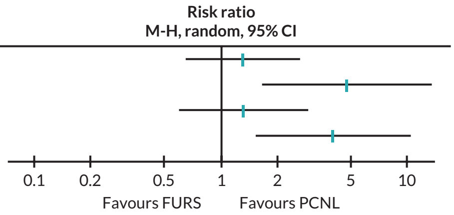


FIGURE 48 Flexible ureterorenoscopy vs. PCNL unplanned procedures.

Retreatment rates

No studies in our original review reported retreatment rates. Three studies reported in retreatment rates in our update. There were proportionately more retreatments in the FURS arm across studies, but event rates were low and there is nothing conclusive that can be said about this outcome and the certainty is very low (see *Figure 49*).

Study or subgroup	FURS Events	FURS Total	PCNL Events	PCNL Total	Risk ratio M-H, random, 95% CI
Bozzini 2017	9	207	4	181	1.97 (0.62 to 6.28)
Fayad 2017	3	51	2	55	1.62 (0.28 to 9.29)
Jiang 2019	12	58	9	58	1.33 (0.61 to 2.92)

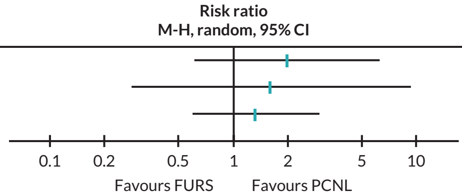


FIGURE 49 Flexible ureterorenoscopy vs. PCNL retreatment rates.

Complications

No studies reported complications in our original review and eight RCTs reported complications in our update. We did not meta-analyse these data because there was lack of standardised reporting (see *Figure 50*). Event rates were generally low and there is no discernible difference between FURS and PCNL.

Study or subgroup	FURS Events	FURS Total	PCNL Events	PCNL Total	Risk ratio M-H, random, 95% CI	Year
Kandemir 2017	6	30	5	30	1.20 (0.41 to 3.51)	2017
Bozzini 2017	30	207	35	181	0.75 (0.48 to 1.17)	2017
Fayad 2017	5	51	5	55	1.08 (0.33 to 3.51)	2017
Zeng 2018	7	80	7	80	1.00 (0.37 to 2.72)	2018
Jin 2019	9	110	16	110	0.56 (0.26 to 1.22)	2019
Zhang 2019	6	60	10	60	0.60 (0.23 to 1.55)	2019
Jiang 2019	3	58	4	58	0.75 (0.18 to 3.20)	2019
PURE 2 2023	17	73	17	86	1.18 (0.65 to 2.14)	2023

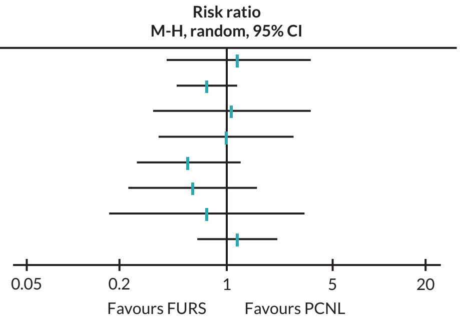


FIGURE 50 Flexible ureterorenoscopy vs. PCNL complications.

Quality of life, satisfaction, and economic outcomes

No studies reported QoL for FURS versus PCNL in our original review. In our update two RCTs reported on various aspects. Yavuz, 2020[90] found that return to daily activities was 3.9, 4.5, 6.5, 9.3, and 13.5 days for FURS, micro, ultra-mini, mini, and standard PNL, respectively and the difference was statistically significant. PUrE 2[37] reported on the AUC from treatment to 12 weeks and found that QoL for PCNL was marginally superior.

Hospital stay

Mean days hospital stay was statistically significantly shorter with FURS by over 2 days in three studies and not different in one (see *Figure 51*).

Copyright © 2025 Wiseman *et al*. This work was produced by Wiseman *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Study or subgroup	FURS Mean (days)	FURS SD (days)	FURS Total	PCNL Mean (days)	PCNL SD (days)	PCNL Total	Mean difference IV, random, 95% CI	Year
Bozzini 2017	1.3	0.4	207	3.7	1.5	181	−2.40 (−2.63 to −2.17)	2017
Jiang 2019	3.2	0.6	58	3.2	0.5	58	0.00 (−0.20 to 0.20)	2019
Jin 2019	3.15	0.72	110	5.59	0.82	110	−2.44 (−2.64 to −2.24)	2019
Zhang 2019	3.2	0.5	60	5.3	1.2	60	−2.10 (−2.43 to −1.77)	2019

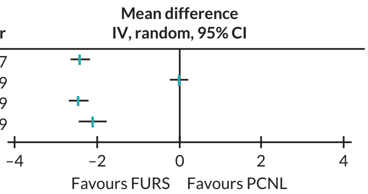


FIGURE 51 Flexible ureterorenoscopy vs. PCNL hospital stay.

Procedure time

Six studies in our update reported procedure time. PCNL was shorter in all studies but one (see *Figure 52*). The mean differences across studies ranged from PCNL taking over 6 minutes longer to 38 minutes shorter. The certainty of evidence is 'very low' given the RoB concerns and inconsistent estimates, which potentially reflect local practice differences rather than anything meaningful.

Study or subgroup	FURS Mean (minutes)	FURS SD (minutes)	FURS Total	PCNL Mean (minutes)	PCNL SD (minutes)	PCNL Total	Mean difference IV, random, 95% CI
Bozzini 2017	55.8	11.4	207	40.9	7.7	181	14.90 (12.98 to 16.82)
Fayad 2017	109.7	20.8	51	71.7	10.4	55	38.00 (31.66 to 44.34)
Jiang 2019	60.3	8.5	58	54	8.2	58	6.30 (3.26 to 9.34)
Jin 2019	87.2	13.34	110	79.6	14.86	110	7.60 (3.87 to 11.33)
Zeng 2018	52.3	22.4	80	58.6	21.6	80	−6.30 (−13.12 to 0.52)
Zhang 2019	93.4	21.6	60	68.6	15.8	60	24.80 (18.03 to 31.57)

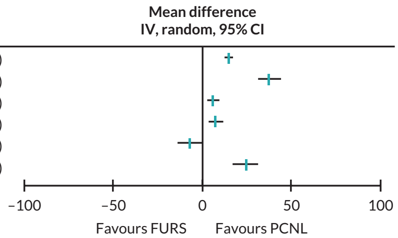


FIGURE 52 Flexible ureterorenoscopy vs. PCNL procedure time.

Other costs

Yavuz (2020)[90] reported on the mean total cost of the procedures and found that it was higher in the RIRS group compared to PCNL in general but that between various sizes of PCNL there were no significant differences. Zhang (2019)[85] found that FURS was more expensive than PCNL (USD 4657.28 ± 679.28 versus 4085.51 ± 416.69).

EME
HSDR
HTA
PGfAR
PHR

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

This report presents independent research funded by the National Institute for Health and Care Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care

Published by the NIHR Journals Library