

Central Lancashire Online Knowledge (CLoK)

Title	Surgical treatments compared with early structured physiotherapy in secondary care for adults with primary frozen shoulder: the UK FROST three-arm RCT
Type	Article
URL	https://clock.uclan.ac.uk/36082/
DOI	https://doi.org/10.3310/hta24710
Date	2020
Citation	Brealey, Stephen, Northgraves, Matthew, Kottam, Lucksy, Keding, Ada, Corbacho, Belen, Goodchild, Lorna, Srikesavan, Cynthia, Rex, Saleema, Charalambous, Charalambos P. et al (2020) Surgical treatments compared with early structured physiotherapy in secondary care for adults with primary frozen shoulder: the UK FROST three-arm RCT. Health Technology Assessment, 24 (71). pp. 1-162. ISSN 1366-5278
Creators	Brealey, Stephen, Northgraves, Matthew, Kottam, Lucksy, Keding, Ada, Corbacho, Belen, Goodchild, Lorna, Srikesavan, Cynthia, Rex, Saleema, Charalambous, Charalambos P., Hanchard, Nigel, Armstrong, Alison, Brooksbank, Andrew, Carr, Andrew, Cooper, Cushla, Dias, Joseph, Donnelly, Iona, Hewitt, Catherine, Lamb, Sarah E, McDaid, Catriona, Richardson, Gerry, Rodgers, Sara, Sharp, Emma, Spencer, Sally, Torgerson, David, Toye, Francine and Rangan, Amar

It is advisable to refer to the publisher's version if you intend to cite from the work.
<https://doi.org/10.3310/hta24710>

For information about Research at UCLan please go to <http://www.uclan.ac.uk/research/>

All outputs in CLoK are protected by Intellectual Property Rights law, including Copyright law. Copyright, IPR and Moral Rights for the works on this site are retained by the individual authors and/or other copyright owners. Terms and conditions for use of this material are defined in the <http://clock.uclan.ac.uk/policies/>

Health Technology Assessment

Volume 24 • Issue 71 • December 2020

ISSN 1366-5278

Surgical treatments compared with early structured physiotherapy in secondary care for adults with primary frozen shoulder: the UK FROST three-arm RCT

Stephen Brealey, Matthew Northgraves, Lucksy Kottam, Ada Keding, Belen Corbacho, Lorna Goodchild, Cynthia Srikesavan, Saleema Rex, Charalambos P Charalambous, Nigel Hanchard, Alison Armstrong, Andrew Brooksbank, Andrew Carr, Cushla Cooper, Joseph Dias, Iona Donnelly, Catherine Hewitt, Sarah E Lamb, Catriona McDaid, Gerry Richardson, Sara Rodgers, Emma Sharp, Sally Spencer, David Torgerson, Francine Toye and Amar Rangan



Surgical treatments compared with early structured physiotherapy in secondary care for adults with primary frozen shoulder: the UK FROST three-arm RCT

Stephen Brealey¹, Matthew Northgraves¹,
Lucksy Kottam², Ada Keding¹, Belen Corbacho¹,
Lorna Goodchild³, Cynthia Srikesavan⁴, Saleema Rex¹,
Charalambos P Charalambous^{5,6}, Nigel Hanchard⁷,
Alison Armstrong⁸, Andrew Brooksbank⁹,
Andrew Carr⁴, Cushla Cooper⁴, Joseph Dias⁸,
Iona Donnelly⁹, Catherine Hewitt¹, Sarah E Lamb⁴,
Catriona McDaid¹, Gerry Richardson¹⁰, Sara Rodgers¹,
Emma Sharp⁹, Sally Spencer¹¹, David Torgerson¹,
Francine Toye¹² and Amar Rangan^{1,2,4*}

¹York Trials Unit, Department of Health Sciences, University of York, York, UK

²The James Cook University Hospital, South Tees Hospitals NHS Foundation Trust, Middlesbrough, UK

³The Physiotherapy Practice, South Shields, UK

⁴Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Botnar Research Centre, University of Oxford, Oxford, UK

⁵Department of Orthopaedics, Blackpool Victoria Hospital, Blackpool, UK

⁶School of Medicine, University of Central Lancashire, Preston, UK

⁷School of Health & Life Sciences, Teesside University, Middlesbrough, UK

⁸University Hospitals of Leicester NHS Trust, Leicester, UK

⁹Glasgow Royal Infirmary, Glasgow, UK

¹⁰Centre for Health Economics, University of York, York, UK

¹¹Postgraduate Medical Institute, Edge Hill University, Ormskirk, UK

¹²Physiotherapy Research Unit, Oxford University Hospitals NHS Foundation Trust, Oxford, UK

*Corresponding author

Declared competing interests of authors: Lucksy Kottam reports grants from the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) programme for other work during the conduct of this study. South Tees Hospitals NHS Foundation Trust receives an educational grant to the Department of Trauma and Orthopaedic Surgery from DePuy Synthes (Warsaw, IN, USA; part of the Johnson & Johnson Medical Devices group). It also receives payment from DePuy Synthes for Lucksy Kottam as a study co-ordinator for the GLOBAL ICON Stemless Shoulder System Post Market

Clinical Follow Up Study: CT 1401. These payments are outside and unrelated to the submitted work. Catherine Hewitt is a member of the NIHR HTA Commissioning Board. Catriona McDaid receives funding from the British Orthopaedic Association (2014 to present). She is a member of the NIHR HTA and Efficacy and Mechanism Evaluation Editorial Board (2017 to present). Sarah E Lamb reports membership of the following boards: HTA Additional Capacity Funding Board 2012–15, HTA Clinical Trials Board 2010–15, HTA End of Life Care and Add on Studies Board 2015, HTA Funding Boards Policy Group (formerly Clinical Studies Group) 2010–15, HTA Maternal, Neonatal and Child Health Methods Group 2013–15, HTA Post-board funding teleconference 2010–15, HTA Primary Care Themed Call Board 2013–14, HTA Prioritisation Group 2010–15 and the NIHR Clinical Trials Unit Standing Advisory Committee 2012–16. Amar Rangan reports other grants from the NIHR HTA programme, Orthopaedic Research UK (London, UK) and Horizon 2020 during the conduct of the study. South Tees Hospitals NHS Foundation Trust receives an educational grant to the department from DePuy Synthes. The institution also receives payment from DePuy Synthes for Amar Rangan as the co-ordinating investigator for the GLOBAL ICON Stemless Shoulder System Post Market Clinical Follow Up Study: CT 1401. These are outside and unrelated to the submitted work. Joseph Dias reports grants from NIHR during the conduct of the study, outside the submitted work.

Published December 2020

DOI: 10.3310/hta24710

This report should be referenced as follows:

Brealey S, Northgraves M, Kottam L, Keding A, Corbacho B, Goodchild L, *et al.* Surgical treatments compared with early structured physiotherapy in secondary care for adults with primary frozen shoulder: the UK FROST three-arm RCT. *Health Technol Assess* 2020;**24**(71).

Health Technology Assessment is indexed and abstracted in *Index Medicus/MEDLINE*, *Excerpta Medica/EMBASE*, *Science Citation Index Expanded (SciSearch®)* and *Current Contents®/Clinical Medicine*.

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 3.370

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, the Cochrane Library and Clarivate Analytics Science Citation Index.

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. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports 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 report

The research reported in this issue of the journal was funded by the HTA programme as project number 13/26/01. The contractual start date was in October 2014. The draft report began editorial review in October 2019 and was accepted for publication in June 2020. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. This report has been published following a shortened production process and, therefore, did not undergo the usual number of proof stages and opportunities for correction. The HTA editors and publisher have tried to ensure the accuracy of the authors' report 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 report.

This report presents independent research funded by the National Institute for Health 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, NETSCC, 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, NETSCC, the HTA programme or the Department of Health and Social Care.

© Queen's Printer and Controller of HMSO 2020. This work was produced by Brealey *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

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

Editor-in-Chief of *Health Technology Assessment* and NIHR Journals Library

Professor Ken Stein Professor of Public Health, University of Exeter Medical School, UK

NIHR Journals Library Editors

Professor John Powell Chair of HTA and EME Editorial Board and Editor-in-Chief of HTA and EME journals. Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK, and Professor of Digital Health Care, Nuffield Department of Primary Care Health Sciences, University of Oxford, UK

Professor Andrée Le May Chair of NIHR Journals Library Editorial Group (HS&DR, PGfAR, PHR journals) and Editor-in-Chief of HS&DR, PGfAR, PHR journals

Professor Matthias Beck Professor of Management, Cork University Business School, Department of Management and Marketing, University College Cork, Ireland

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Eugenia Cronin Senior Scientific Advisor, Wessex Institute, UK

Dr Peter Davidson Consultant Advisor, Wessex Institute, University of Southampton, UK

Ms Tara Lamont Senior Scientific Adviser (Evidence Use), Wessex Institute, University of Southampton, UK

Dr Catriona McDaid Senior Research Fellow, York Trials Unit, Department of Health Sciences, University of York, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Emeritus Professor of Wellbeing Research, University of Winchester, UK

Professor John Norrie Chair in Medical Statistics, University of Edinburgh, UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Great Ormond Street Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Professor Ken Stein Professor of Public Health, University of Exeter Medical School, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Professor Martin Underwood Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, UK

Please visit the website for a list of editors: www.journalslibrary.nihr.ac.uk/about/editors

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

Abstract

Surgical treatments compared with early structured physiotherapy in secondary care for adults with primary frozen shoulder: the UK FROST three-arm RCT

Stephen Brealey¹, Matthew Northgraves¹, Lucksy Kottam²,
Ada Keding¹, Belen Corbacho¹, Lorna Goodchild³,
Cynthia Srikesavan⁴, Saleema Rex¹, Charalambos P Charalambous^{5,6},
Nigel Hanchard⁷, Alison Armstrong⁸, Andrew Brooksbank⁹,
Andrew Carr⁴, Cushla Cooper⁴, Joseph Dias⁸, Iona Donnelly⁹,
Catherine Hewitt¹, Sarah E Lamb⁴, Catriona McDaid¹,
Gerry Richardson¹⁰, Sara Rodgers¹, Emma Sharp⁹, Sally Spencer¹¹,
David Torgerson¹, Francine Toye¹² and Amar Rangan^{1,2,4*}

¹York Trials Unit, Department of Health Sciences, University of York, York, UK

²The James Cook University Hospital, South Tees Hospitals NHS Foundation Trust, Middlesbrough, UK

³The Physiotherapy Practice, South Shields, UK

⁴Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Botnar Research Centre, University of Oxford, Oxford, UK

⁵Department of Orthopaedics, Blackpool Victoria Hospital, Blackpool, UK

⁶School of Medicine, University of Central Lancashire, Preston, UK

⁷School of Health & Life Sciences, Teesside University, Middlesbrough, UK

⁸University Hospitals of Leicester NHS Trust, Leicester, UK

⁹Glasgow Royal Infirmary, Glasgow, UK

¹⁰Centre for Health Economics, University of York, York, UK

¹¹Postgraduate Medical Institute, Edge Hill University, Ormskirk, UK

¹²Physiotherapy Research Unit, Oxford University Hospitals NHS Foundation Trust, Oxford, UK

*Corresponding author amar.rangan@york.ac.uk

Background: Frozen shoulder causes pain and stiffness. It affects around 10% of people in their fifties and is slightly more common in women. Costly and invasive surgical interventions are used, without high-quality evidence that these are effective.

Objectives: To compare the clinical effectiveness and cost-effectiveness of three treatments in secondary care for adults with frozen shoulder; to qualitatively explore the acceptability of these treatments to patients and health-care professionals; and to update a systematic review to explore the trial findings in the context of existing evidence for the three treatments.

Design: This was a pragmatic, parallel-group, multicentre, open-label, three-arm, randomised superiority trial with unequal allocation (2 : 2 : 1). An economic evaluation and a nested qualitative study were also carried out.

Setting: The orthopaedic departments of 35 hospitals across the UK were recruited from April 2015, with final follow-up in December 2018.

Participants: Participants were adults (aged ≥ 18 years) with unilateral frozen shoulder, characterised by restriction of passive external rotation in the affected shoulder to $< 50\%$ of the opposite shoulder, and with plain radiographs excluding other pathology.

Interventions: The interventions were early structured physiotherapy with a steroid injection, manipulation under anaesthesia with a steroid injection and arthroscopic capsular release followed by manipulation. Both of the surgical interventions were followed with post-procedural physiotherapy.

Main outcome measures: The primary outcome and end point was the Oxford Shoulder Score at 12 months post randomisation. A difference of 5 points between early structured physiotherapy and manipulation under anaesthesia or arthroscopic capsular release or of 4 points between manipulation under anaesthesia and arthroscopic capsular release was judged clinically important.

Results: The mean age of the 503 participants was 54 years; 319 were female (63%) and 150 had diabetes (30%). The primary analyses comprised 473 participants (94%). At the primary end point of 12 months, participants randomised to arthroscopic capsular release had, on average, a statistically significantly higher (better) Oxford Shoulder Score than those randomised to manipulation under anaesthesia (2.01 points, 95% confidence interval 0.10 to 3.91 points; $p = 0.04$) or early structured physiotherapy (3.06 points, 95% confidence interval 0.71 to 5.41 points; $p = 0.01$). Manipulation under anaesthesia did not result in statistically significantly better Oxford Shoulder Score than early structured physiotherapy (1.05 points, 95% confidence interval -1.28 to 3.39 points; $p = 0.38$). No differences were deemed of clinical importance. Serious adverse events were rare but occurred in participants randomised to surgery (arthroscopic capsular release, $n = 8$; manipulation under anaesthesia, $n = 2$). There was, however, one serious adverse event in a participant who received non-trial physiotherapy. The base-case economic analysis showed that manipulation under anaesthesia was more expensive than early structured physiotherapy, with slightly better utilities. The incremental cost-effectiveness ratio for manipulation under anaesthesia was £6984 per additional quality-adjusted life-year, and this intervention was probably 86% cost-effective at the threshold of £20,000 per quality-adjusted life-year. Arthroscopic capsular release was more costly than early structured physiotherapy and manipulation under anaesthesia, with no statistically significant benefit in utilities. Participants in the qualitative study wanted early medical help and a quicker pathway to resolve their shoulder problem. Nine studies were identified from the updated systematic review, including UK FROST, of which only two could be pooled, and found that arthroscopic capsular release was more effective than physiotherapy in the long-term shoulder functioning of patients, but not to the clinically important magnitude used in UK FROST.

Limitations: Implementing physiotherapy to the trial standard in clinical practice might prove challenging but could avoid theatre use and post-procedural physiotherapy. There are potential confounding effects of waiting times in the trial.

Conclusions: None of the three interventions was clearly superior. Early structured physiotherapy with a steroid injection is an accessible and low-cost option. Manipulation under anaesthesia is the most cost-effective option. Arthroscopic capsular release carries higher risks and higher costs.

Future work: Evaluation in a randomised controlled trial is recommended to address the increasing popularity of hydrodilatation despite the paucity of high-quality evidence.

Trial registration: Current Controlled Trials ISRCTN48804508.

Funding: This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 24, No. 71. See the NIHR Journals Library website for further project information.

Contents

List of tables	xiii
List of figures	xvii
List of supplementary material	xix
List of abbreviations	xxi
Plain English summary	xxiii
Scientific summary	xxv
Chapter 1 Introduction	1
Frozen shoulder	1
Diagnosis of frozen shoulder	2
Treatments for frozen shoulder	2
Rationale for the UK FROzen Shoulder Trial	3
Aim and objectives	3
Chapter 2 Trial design and methods	5
Trial design	5
Participants	5
<i>Inclusion criteria</i>	5
<i>Exclusion criteria</i>	5
Setting	5
Interventions	6
<i>Manipulation under anaesthesia with an intra-articular steroid</i>	6
<i>Arthroscopic capsular release with manipulation under anaesthesia</i>	6
<i>Early structured physiotherapy</i>	7
<i>Post-procedural physiotherapy</i>	7
<i>Steroid injections</i>	8
<i>Modifications to interventions</i>	8
<i>Adherence to interventions</i>	8
<i>Concomitant care</i>	8
Outcomes	8
<i>Primary outcome</i>	8
<i>Secondary outcomes</i>	9
Sample size	10
Internal pilot study	10
<i>Phase 1 (months 4–9)</i>	10
<i>Phase 2 (months 10–27)</i>	10
Recruitment	11
Randomisation	11
Blinding	12
Statistical methods	12
<i>Trial progression</i>	12
<i>Baseline characteristics</i>	12
<i>Intervention delivery/fidelity</i>	12

CONTENTS

<i>Missing data</i>	12
<i>Primary outcome (Oxford Shoulder Score) analysis</i>	13
<i>Secondary analyses</i>	13
<i>Subgroup analyses</i>	14
<i>Analysis of secondary outcomes</i>	14
<i>Other analyses</i>	15
Data management	15
Adverse event management	16
Ethics approval and monitoring	17
<i>Ethics committee approval and any changes to the project protocol</i>	17
<i>Trial Management Group</i>	17
<i>Trial Steering Committee</i>	17
<i>Data Monitoring and Ethics Committee</i>	17
Patient and public involvement	17
Chapter 3 Trial results	19
Summary findings of the internal pilot	19
Summary findings of the shoulder capsule tissue and blood samples study	20
Recruitment into UK FROST	21
Flow of participants	21
Baseline characteristics	21
Intervention delivery	21
Primary outcome	31
<i>Descriptives</i>	31
<i>Primary analysis</i>	32
<i>Secondary analyses</i>	33
<i>Subgroup analyses</i>	37
Secondary outcomes	42
<i>QuickDASH</i>	42
<i>Shoulder pain</i>	43
<i>Extent of recovery</i>	44
<i>Stiffness</i>	45
<i>Complications/adverse events</i>	45
Other analyses	48
<i>Treatment preferences</i>	48
<i>Oxford Shoulder Score change scores</i>	48
<i>Outcomes for patients receiving no treatment</i>	48
Systematic review: integrating the new evidence	48
Chapter 4 Economic evaluation	49
Objective	49
Overview	49
Methods	49
<i>Data sources</i>	49
<i>Measurement of resource use and costs</i>	50
<i>Estimation of costs</i>	51
<i>Health-related quality of life and quality-adjusted life-years</i>	52
<i>Missing data</i>	53
Cost-effectiveness analyses	54
<i>Sensitivity analyses and uncertainty</i>	55
Results	55
<i>Study population</i>	55
<i>Health-care resource use and costs</i>	55

<i>Health-related quality of life and quality-adjusted life-years</i>	61
<i>Cost-effectiveness analysis</i>	63
<i>Sensitivity analyses and uncertainty</i>	64
Conclusion	67
Discussion	67
Chapter 5 Qualitative study	69
Introduction	69
Methods	69
Results	70
<i>Interviews with trial participants</i>	70
<i>Interviews with surgeons and physiotherapists</i>	78
Discussion	83
<i>Principal findings</i>	83
<i>Study limitations</i>	86
<i>Strengths</i>	86
<i>Implications for clinical practice</i>	86
Conclusion	86
Chapter 6 Discussion and conclusion	87
Principal findings of clinical effectiveness	87
<i>Primary outcome</i>	87
<i>Secondary outcomes</i>	87
<i>Fidelity of treatment</i>	88
<i>Further treatment</i>	88
<i>Safety</i>	88
<i>Systematic review update of the currently available evidence</i>	88
Cost-effectiveness	89
Qualitative study findings	89
Trial validity and minimising bias	89
Applicability of results	91
<i>Characteristics of the trial population</i>	91
Application of the trial results to clinical practice	93
Conclusion	94
<i>Recommendations for research</i>	94
Acknowledgements	95
References	101
Appendix 1 Participating trusts	111
Appendix 2 Table of amendments	113
Appendix 3 Recruitment	115
Appendix 4 Practitioner characteristics	117
Appendix 5 Elements of physiotherapy	119
Appendix 6 Oxford Shoulder Score subdomains	125
Appendix 7 Illustration of treatment effects	127

CONTENTS

Appendix 8 Oxford Shoulder Score by treatment completion at start and end of trial	129
Appendix 9 Other secondary analyses	131
Appendix 10 Subgroup descriptive statistics (Oxford Shoulder Score)	133
Appendix 11 Secondary outcomes descriptives	137
Appendix 12 Adverse events	141
Appendix 13 Treatment preferences	143
Appendix 14 Oxford Shoulder Score change scores	147
Appendix 15 Outcomes for patients receiving no treatment	149
Appendix 16 Health-related quality of life and quality-adjusted life-years	151
Appendix 17 Missing data in health economics analysis	155

List of tables

TABLE 1 Reasons for exclusion	23
TABLE 2 Baseline characteristics of different populations	24
TABLE 3 Baseline characteristics of randomised participants	24
TABLE 4 Completed treatment	28
TABLE 5 Time to start/end of treatment	29
TABLE 6 Further treatment (any treatment following completion of trial treatment)	29
TABLE 7 Fidelity (surgery and injection)	30
TABLE 8 Fidelity (physiotherapy)	30
TABLE 9 Unadjusted OSS by treatment arm: follow-up since randomisation	30
TABLE 10 Estimated mean OSS differences by treatment arm (estimates from primary analysis model)	32
TABLE 11 Unadjusted OSS by treatment arm: pre treatment and 6 months post treatment	33
TABLE 12 Estimated mean OSS differences by treatment arm (estimates from model incorporating follow-ups before and after treatment in addition to post-randomisation outcomes)	34
TABLE 13 Comparison of baseline characteristics by treatment compliance	35
TABLE 14 Comparison of patient characteristics by missingness of OSS over time	36
TABLE 15 Subgroup analyses summary	37
TABLE 16 Estimated mean QuickDASH differences by treatment arm	43
TABLE 17 Estimated shoulder Numeric Rating Scale for Pain differences by treatment arm	44
TABLE 18 Estimated mean extent of recovery visual analogue scale differences by treatment arm	45
TABLE 19 Serious adverse events (itemised)	46
TABLE 20 Non-serious AEs (itemised)	46
TABLE 21 Resource use and costs related to initial surgery: MUA and ACR	52
TABLE 22 Costs related to ESP	56

TABLE 23 Cost related to complications and additional treatments by trial arm	56
TABLE 24 Average primary care and community care resource use (shoulder related) and days missed off work per treatment group	57
TABLE 25 Unit costs used for the analysis (£, 2017–18 prices)	59
TABLE 26 Costs for patients with complete data by treatment arm and cost category (£, 2017–18 prices)	60
TABLE 27 Health-related quality of life: EQ-5D-5L summary scores (available cases) at each time point adjusted for baseline utility	62
TABLE 28 Adjusted mean differences in QALYs and costs between interventions (base case)	64
TABLE 29 Sensitivity analysis (scenarios 1 and 2): summary for incremental analysis (ITT), cost-effectiveness results and uncertainty under different cost scenarios	65
TABLE 30 Sensitivity analyses (scenarios 3, 4 and 5): summary for incremental analysis (ITT), cost-effectiveness results and uncertainty under different missing data assumptions	66
TABLE 31 Characteristics of the trial participants interviewed	71
TABLE 32 Participant-reported problems mapped to ICF categories: body functions and structures	72
TABLE 33 Participant-reported problems mapped to ICF categories: activities and participation	72
TABLE 34 Participant-reported problems mapped to ICF categories: environmental factors	73
TABLE 35 Characteristics of the surgeons and physiotherapists interviewed	79
TABLE 36 Recruitment by site	115
TABLE 37 Practitioner characteristics	117
TABLE 38 Physiotherapy elements received (common treatments)	120
TABLE 39 Physiotherapy elements received (pain-specific treatments)	122
TABLE 40 Physiotherapy elements received (stiffness-specific treatments)	123
TABLE 41 Unadjusted OSS pain subdomain by treatment arm	125
TABLE 42 Unadjusted OSS function subdomain by treatment arm	126
TABLE 43 Unadjusted OSS scores at 12 months by treatment completion	129

TABLE 44 Estimated mean OSS differences by treatment arm (estimates from analysis, excluding data received after 6 weeks)	131
TABLE 45 Estimated mean OSS differences by treatment arm (estimates from analysis, adjusted for employment status)	132
TABLE 46 Unadjusted OSS by treatment arm and diabetes status	133
TABLE 47 Unadjusted OSS by treatment arm and previous physiotherapy	134
TABLE 48 Unadjusted OSS by treatment arm and patient preference	135
TABLE 49 Unadjusted OSS by treatment arm and duration of symptoms at baseline	136
TABLE 50 Unadjusted OSS for patients who completed treatment by receipt of steroid injection	136
TABLE 51 Unadjusted QuickDASH by treatment arm	137
TABLE 52 Unadjusted Numeric Rating Scale for Pain by treatment arm	138
TABLE 53 Unadjusted extent of recovery by treatment arm	139
TABLE 54 Predominant shoulder problem (among patients who received their allocated treatment)	139
TABLE 55 Serious adverse events (summary by treatment arm)	141
TABLE 56 Non-serious AEs (summary by treatment arm)	142
TABLE 57 Treatment preferences of non-consenting patients	143
TABLE 58 Strength of treatment preference of non-randomised patients (scale 1–10)	143
TABLE 59 Treatment expectations	144
TABLE 60 Treatment preferences for randomised patients at baseline	144
TABLE 61 Treatment preferences for randomised patients at 12 months	145
TABLE 62 Anchoring of OSS change scores	147
TABLE 63 Unadjusted trial outcomes by receipt of treatment	149
TABLE 64 Health-related quality of life: number of questionnaires returned and completed EQ-5D-5L scores	151
TABLE 65 Health-related quality of life: proportion reporting EQ-5D-5L levels 1–5 by dimension, treatment arm and time point (complete cases)	152
TABLE 66 Description of economic variables in UK FROST	156
TABLE 67 Number and proportion of patients with complete data by treatment arm	158

LIST OF TABLES

TABLE 68 Logistic regression for (1) missingness of costs and QALYs on baseline variables and (2) missingness between missing costs and QALYs and observed outcomes	160
TABLE 69 Sensitivity analysis (scenario 6): summary for incremental analysis (ITT), cost-effectiveness results and uncertainty of different methods for handling missing data (MUA vs. ESP)	161

List of figures

FIGURE 1 Diagram showing site of pathology of frozen shoulder	1
FIGURE 2 The CONSORT flow diagram	22
FIGURE 3 Unadjusted mean OSS and 95% CIs by treatment arm	31
FIGURE 4 Unadjusted mean OSS and 95% CIs by treatment arm (using scores at baseline, follow-up before treatment and 6 months after treatment; and at 3, 6 and 12 months post randomisation)	33
FIGURE 5 Unadjusted mean OSS function items and 95% CIs by treatment arm and diabetes: (a) diabetic; and (b) not diabetic	38
FIGURE 6 Unadjusted mean OSS function items and 95% CIs by treatment arm and previous physiotherapy: (a) had previous physiotherapy; and (b) did not have previous physiotherapy	39
FIGURE 7 Unadjusted mean OSS function items and 95% CIs by treatment arm and baseline preference: (a) preference for physiotherapy; (b) preference for surgery; and (c) no preference	40
FIGURE 8 Unadjusted mean OSS function items and 95% CIs by treatment arm and length of time with symptoms: (a) symptoms for < 9 months; and (b) symptoms for ≥ 9 months	41
FIGURE 9 Unadjusted mean QuickDASH scores and 95% CIs by treatment arm	42
FIGURE 10 Unadjusted mean shoulder Numeric Rating Scale for Pain and 95% CIs by treatment arm	43
FIGURE 11 Unadjusted mean extent of recovery visual analogue scale and 95% CIs by treatment arm	44
FIGURE 12 The EQ-5D-5L scores distribution at the different time points over the 12 months	62
FIGURE 13 Base-case CEACs	64
FIGURE 14 Flow diagram of the trial participant interviews	71
FIGURE 15 Flow diagram of surgeons' and physiotherapists' interviews	79
FIGURE 16 Unadjusted mean OSS pain items and 95% CIs by treatment arm	125
FIGURE 17 Unadjusted mean OSS function items and 95% CIs by treatment arm	126
FIGURE 18 Estimated mean OSS differences from primary analysis model by treatment arm (ESP vs. surgery)	127

LIST OF FIGURES

FIGURE 19 Estimated mean OSS differences from primary analysis model by treatment arm (MUA vs. ACR)	127
FIGURE 20 Pattern of missing data in UK FROST data set	159
FIGURE 21 Comparison of the distribution of imputed values (imputations 1–10) with the observed data (imputation number 0) for (a) costs; and (b) QALYs	160

List of supplementary material

- Report Supplementary Material 1 MUA surgery form
- Report Supplementary Material 2 ACR surgery form
- Report Supplementary Material 3 Frozen shoulder information sheet
- Report Supplementary Material 4 Frozen shoulder home exercises
- Report Supplementary Material 5 Structured physiotherapy log-book
- Report Supplementary Material 6 Post-procedural physiotherapy log-book
- Report Supplementary Material 7 52-week complication form
- Report Supplementary Material 8 Study eligibility form
- Report Supplementary Material 9 Trial participant information sheet
- Report Supplementary Material 10 Trial consent form
- Report Supplementary Material 11 Baseline form
- Report Supplementary Material 12 Consent status form
- Report Supplementary Material 13 Patient preference form
- Report Supplementary Material 14 Pre-treatment form
- Report Supplementary Material 15 Change in status form
- Report Supplementary Material 16 Three-month patient questionnaire
- Report Supplementary Material 17 Adverse event form
- Report Supplementary Material 18 Serious adverse event form
- Report Supplementary Material 19 Adverse event review form
- Report Supplementary Material 20 Systematic review report
- Report Supplementary Material 21 Interview schedule: trial participants
- Report Supplementary Material 22 Interview schedule: health professionals

Supplementary material can be found on the NIHR Journals Library report page (<https://doi.org/10.3310/hta24710>).

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

List of abbreviations

ACR	arthroscopic capsular release	ITT	intention to treat
AE	adverse event	MAR	missing at random
CACE	complier-average causal effect	MNAR	missing not at random
CEAC	cost-effectiveness acceptability curve	MUA	manipulation under anaesthesia
CI	confidence interval	NICE	National Institute for Health and Care Excellence
CONSORT	Consolidated Standards of Reporting Trials	OSS	Oxford Shoulder Score
CRF	case report form	PPP	post-procedural physiotherapy
DMEC	Data Monitoring and Ethics Committee	PSSRU	Personal Social Services Research Unit
EQ-5D	EuroQol-5 Dimensions	QALY	quality-adjusted life-year
EQ-5D-5L	EuroQol-5 Dimensions, five-level version	QuickDASH	Quick Disabilities of the Arm, Shoulder and Hand
ESP	early structured physiotherapy	RCT	randomised controlled trial
GP	general practitioner	SAE	serious adverse event
HRQoL	health-related quality of life	SD	standard deviation
HTA	Health Technology Assessment	SE	standard error
ICER	incremental cost-effectiveness ratio	SUR	seemingly unrelated regression
ICF	<i>International Classification of Functioning, Disability and Health</i>	TMG	Trial Management Group
		TSC	Trial Steering Committee
		UK FROST	UK FROzen Shoulder Trial

Plain English summary

Frozen shoulder occurs when the soft tissue envelope around the shoulder joint becomes inflamed, scarred and contracted, making movement painful and stiff. It affects around 1 in 10 people and is more common in women. Most patients are treated in the community. Those who do not improve are offered treatments in hospital. This includes costly and invasive surgical options. It is unclear which treatment provides the best patient outcomes and is cost-effective.

UK FROST (UK FROzen Shoulder Trial) comprised 503 patients (from 35 UK hospitals) who randomly received one of three commonly offered treatments for frozen shoulder:

1. early physiotherapy to restore movement, including a steroid injection for pain relief
2. manipulation under anaesthesia, to stretch and tear the tight capsule to restore movement, and a steroid injection followed by physiotherapy
3. arthroscopic capsular release, which uses keyhole surgery, including manipulation, to restore movement, followed by physiotherapy with pain medication.

No important differences were found between the three treatments in shoulder function or pain at 12 months. Fewer patients who received arthroscopic capsular release required further treatment, and patients who received arthroscopic capsular release had slightly better shoulder function and pain outcomes than those who received the manipulation procedure or early physiotherapy. This improvement, however, was unlikely to be of clinical benefit to patients. Arthroscopic capsular release had slightly higher risks and substantially higher costs. Six serious complications were reported in patients who received arthroscopic capsular release (mostly owing to co-existing health problems) and two were reported in patients who received manipulation under anaesthesia. Physiotherapy was the least expensive treatment, but patients who received manipulation under anaesthesia had slightly better general health than those who received physiotherapy. Early physiotherapy with steroid injection could be accessed quicker than the surgical alternatives. Manipulation under anaesthesia cost more than physiotherapy but provided the best value for money. Patients in the study wanted early access to medical help to improve their shoulder problems.

Scientific summary

Background

Frozen shoulder occurs when the capsule, or soft tissue envelope, around the ball-and-socket shoulder joint becomes inflamed, scarred and contracted. This makes the shoulder very painful and stiff. Less invasive treatments, such as pain medication, are provided in primary care in the UK. When stiffness becomes more established, treatments include physiotherapy with a steroid injection, manipulation under anaesthesia, and arthroscopic capsular release. With the intention of facilitating quicker recovery, more invasive and costly surgical interventions (manipulation under anaesthesia and arthroscopic capsular release) are being used despite a lack of good evidence that these are effective.

Objectives

The objectives of UK FROST (UK FROzen Shoulder Trial) were to evaluate the clinical effectiveness and cost-effectiveness of early structured physiotherapy compared with manipulation under anaesthesia compared with arthroscopic capsular release for patients referred to secondary care for the treatment of primary frozen shoulder; to carry out a qualitative study to explore the acceptability of the different interventions to trial participants and health-care professionals; and to undertake a systematic review update to explore the trial findings in the context of existing evidence.

Methods

Randomised controlled trial

Design

This was a pragmatic, parallel-group, multicentre, open-label, three-arm, randomised superiority trial. The randomisation sequence was based on a computer-generated randomisation algorithm provided by a remote randomisation service. Individual patients were allocated to manipulation under anaesthesia, arthroscopic capsular release or early structured physiotherapy in the ratio of 2 : 2 : 1, stratified by the presence of diabetes, using random blocks sizes of 10 and 15.

Eligibility criteria

Adults aged ≥ 18 years presenting with a clinical diagnosis of frozen shoulder, characterised by restriction of passive external rotation in the affected shoulder to $< 50\%$ that of the contralateral shoulder, and radiographs to exclude other pathologies were eligible for inclusion. Exclusion criteria were a bilateral concurrent frozen shoulder; frozen shoulder secondary to trauma that required hospital care; frozen shoulder secondary to other causes; contraindication to any of the trial treatments; not resident in a catchment area of a trial site; or lack of mental capacity to understand the trial.

Setting

The orthopaedic departments of 35 NHS hospitals in the UK across a range of urban and rural areas (April 2015–December 2018).

Interventions

Early structured physiotherapy

Early structured physiotherapy consisted of up to 12 weekly sessions comprising essential 'focused physiotherapy' and optional supplementary physiotherapy. Focused physiotherapy comprised an information leaflet providing education and advice on pain management and function; an intra-articular steroid injection; and hands-on mobilisation techniques, increasingly stretching into the stiff zone of the shoulder as the condition improved. Participants received supervised exercises and instructions on a graduated home exercise programme.

Manipulation under anaesthesia

Manipulation under anaesthesia involved the affected shoulder being manipulated to stretch and tear the tight capsule and to improve the range of movement. An intra-articular corticosteroid injection to the glenohumeral joint was used while the patient was under anaesthesia, unless this was contraindicated. Post-procedural physiotherapy was provided.

Arthroscopic capsular release

Arthroscopic release of the contracted rotator interval and anterior capsule was performed, followed by manipulation under anaesthesia to complete the release of the inferior capsule. Steroid injections were permitted at the surgeon's discretion. Post-procedural physiotherapy was provided.

Outcome measures

The primary outcome was the Oxford Shoulder Score at 12 months post randomisation. The Oxford Shoulder Score is a 12-item patient-reported outcome measure with a score range from 0 (worst) to 48 (best). This was also completed at 3 and 6 months post randomisation. Secondary outcomes, gathered at 3, 6 and 12 months, were the QuickDASH (Quick Disabilities of the Arm, Shoulder and Hand); a Numeric Rating Scale for shoulder pain during the past 24 hours; extent of recovery using a visual analogue scale (0–100); and the EuroQol-5 Dimensions, five-level version. Expected and unexpected complications and adverse events were also recorded.

Sample size

The minimum clinically important difference on the Oxford Shoulder Score was defined as a 5-point difference (standard effect size 0.42) between surgery and no surgery and a 4-point (standard effect size 0.33) difference between manipulation under anaesthesia and arthroscopic capsular release. A total sample size of 500 patients was required to observe these effect sizes, with 90% power and 5% two-sided significance, adjusting for a moderate estimate ($r = 0.4$) of the correlation between Oxford Shoulder Score over 12 months and allowing for 20% attrition.

Analyses

Analyses were conducted for arthroscopic capsular release compared with early structured physiotherapy, manipulation under anaesthesia compared with early structured physiotherapy and arthroscopic capsular release compared with manipulation under anaesthesia using Stata[®] version 15 (StataCorp LP, College Station, TX, USA) and two-sided statistical significance at the 0.05 level. The intention-to-treat primary analysis was based on a linear mixed model incorporating the Oxford Shoulder Score at all available time points and using an unstructured covariance pattern to model the relationship of repeated measurements by the same individual. The model was adjusted for Oxford Shoulder Score at baseline and included as further fixed effects treatment arm, time, arm-by-time interaction, age, sex and diabetes, with recruitment site as a random effect. The model provided estimates for each of the three treatment comparisons at individual time points, including the primary end point of 12 months, as well as an overall treatment effect over 12 months. The estimates are reported as mean differences between treatment arms with 95% confidence intervals and associated p -values. Continuous secondary outcomes were analysed using the same method as the primary outcome and adjusting for the same covariates.

Prespecified sensitivity analyses explored the effect of non-compliance with early structured physiotherapy using complier average causal effect analysis; the effect of waiting times for interventions using additional data collected just before and 6 months following treatment; the impact of missing data; and the effect of questionnaire return outside the intended follow-up time. The Data Monitoring and Ethics Committee advised that employment status be included as a model covariate in a sensitivity analysis. Prespecified subgroup analyses explored possible treatment effect interactions with diabetes, previous receipt of physiotherapy and patient baseline treatment preference. The Trial Steering Committee advised on including a subgroup analysis for duration of symptoms at the time patient eligibility was confirmed.

Economic evaluation

Costs and health benefits were compared for the three treatment arms over the 12 months and hence discounting was not required. All costs were expressed in Great British pounds at a 2017–18 price base. Health outcomes were assessed in terms of quality-adjusted life-years, based on patients' health-related quality-of-life outcomes obtained from trial participants using the EuroQol-5 Dimensions, five-level version, at baseline and at 3, 6 and 12 months. Differences in mean costs and mean quality-adjusted life-years at 12 months were used to derive the incremental cost-effectiveness ratio for surgery and non-surgical treatment. The base-case analysis was conducted on an intention-to-treat basis, with multiple imputation for missing data, and using a UK NHS and Personal Social Services perspective. A secondary analysis took a broader perspective that included private care and productivity costs (i.e. days lost from work).

Qualitative study

This study explored the trial participants' experience and acceptability of the treatments and taking part in the trial, and surgeons' and physiotherapists' experience of the treatments they delivered in the trial. Face-to-face or telephone interviews were undertaken. Interviews were undertaken by a physiotherapy researcher trained in qualitative research methods who was not involved in the trial. Interviews were semistructured and used open questions, and they were audio-recorded and transcribed. The interviews were analysed using constant comparative methods. Transcripts were coded and categorised into themes using NVivo 11 qualitative data software (QSR International, Warrington, UK) and reviewed by a second researcher. Data from trial participants were mapped against the World Health Organization's *International Classification of Functioning, Disability and Health* framework.

Systematic review update

MEDLINE/PreMEDLINE, CENTRAL (Cochrane Central Register of Controlled Trials), EMBASE, PEDro, Science Citation Index, ClinicalTrials.gov and World Health Organization International Clinical Trials Registry were searched from January 2010 to December 2018, and studies reported prior to 2010 were obtained from our previous Health Technology Assessment review [Maund E, Craig D, Suekarran S, Neilson A, Wright K, Brealey S, *et al.* Management of frozen shoulder: a systematic review and cost-effectiveness analysis. *Health Technology Assessment* 2012;16(11)]. Randomised controlled trials evaluating manipulation under anaesthesia, arthroscopic capsular release, hydrodilatation or physiotherapy plus a steroid injection for treatment of primary frozen shoulder were compared with each other, no treatment or supportive care were eligible. The primary outcome was patient-reported function and disability at 12 months. Study selection was undertaken independently by two researchers. For continuous outcomes, the post-intervention mean (standard deviation, and number of participants) for each arm was extracted, where available. The standardised mean difference was calculated to allow comparison between studies. Data extraction and assessment using the Cochrane Risk of Bias tool was undertaken by one researcher and checked by a second. Narrative and tabular summaries of key study characteristics, results and quality assessment are provided. A pairwise meta-analysis using a random-effects model was undertaken for a single comparison only because of limited data and methodological and statistical heterogeneity.

Results

Randomised controlled trial

Of 914 patients screened, 503 were randomised to one of manipulation under anaesthesia ($n = 201$), arthroscopic capsular release ($n = 203$) and early structured physiotherapy ($n = 99$). Follow-up rates were between 85% and 89%, and no evidence was seen of differential dropout across the treatment arms. The primary analysis comprised all participants with Oxford Shoulder Score outcome data at one or more follow-ups (94%). Average shoulder function improved in all treatment arms, with many participants (24%) regaining function to the top Oxford Shoulder Score at 12 months.

At the primary end point at 12 months, participants randomised to arthroscopic capsular release had, on average, a statistically significantly higher (better) Oxford Shoulder Score than those randomised to manipulation under anaesthesia (2.01 points, 95% confidence interval 0.10 to 3.91 points) or to early structured physiotherapy (3.06 points, 95% confidence interval 0.71 to 5.41 points). There was no statistically significant difference between manipulation under anaesthesia and early structured physiotherapy (1.05 points, 95% confidence interval -1.28 to 3.39 points).

At the short-term follow-up at 3 months, arthroscopic capsular release had lower (worse) Oxford Shoulder Score than the other two interventions (vs. manipulation under anaesthesia -3.36 points, 95% confidence interval -5.27 to -1.45 points; vs. early structured physiotherapy -4.72 points, 95% confidence interval -7.06 to -2.39 points). There was no evidence of statistically significant differences in average Oxford Shoulder Score over the 12 months' follow-up (manipulation under anaesthesia vs. early structured physiotherapy 0.61 points, 95% confidence interval -1.31 to 2.53 points; arthroscopic capsular release vs. early structured physiotherapy -0.23 points, 95% confidence interval -2.15 to 1.70 points; arthroscopic capsular release vs. manipulation under anaesthesia -0.84 points, 95% confidence interval -2.41 to 0.72 points).

Mean differences were short of the minimal clinically important effect size of 4 (arthroscopic capsular release vs. manipulation under anaesthesia) to 5 (arthroscopic capsular release or manipulation under anaesthesia vs. early structured physiotherapy) Oxford Shoulder Score points. However, differences of that magnitude were included in the 95% confidence intervals for the benefit of manipulation under anaesthesia and early structured physiotherapy over arthroscopic capsular release at 3 months, and for the benefit of arthroscopic capsular release over early structured physiotherapy at 12 months. Sensitivity analyses did not substantially alter the results. There were no significant subgroup interactions.

Around 20% of all trial participants did not complete their treatment. The complexity of the multiple alternative pathways for each participant limited the analyses of the effect of compliance. At 12 months, the outcomes for early structured physiotherapy compliers remained lower (worse) than for those who complied in both surgery arms combined (-1.84 Oxford Shoulder Score points, 95% confidence interval -4.41 to 0.74 points; $p = 0.157$).

Among the secondary outcomes, QuickDASH and shoulder pain results followed a similar pattern to the Oxford Shoulder Score, in that, compared with those allocated to manipulation under anaesthesia or early structured physiotherapy, statistically significant poorer outcomes were observed for arthroscopic capsular release participants at 3 months but better outcomes were observed at 12 months. There were no statistically significant differences between the treatment arms in response to the global question on the extent of recovery.

In total, there were only 10 serious adverse events, reported for nine participants. All serious adverse events occurred in the surgical arms (arthroscopic capsular release, $n = 8$; manipulation under anaesthesia, $n = 2$), although one participant in the arthroscopic capsular release arm had a serious adverse event resulting from non-trial physiotherapy. The events mainly related to serious medical complications such as chest infection or stroke. Thirty-three non-serious adverse events were reported

for 31 participants, and these were mainly expected and often were related to persistent or worsening shoulder pain. There was no evidence of a statistical difference in the proportions of non-serious adverse events ($p = 0.186$).

Economic evaluation

The base-case economic analysis showed that, at 12 months, manipulation under anaesthesia was, on average, £276 more costly per participant (95% confidence interval £65.67 to £487.35) than early structured physiotherapy. Manipulation under anaesthesia was slightly more beneficial in terms of utilities (mean 0.0396 more quality-adjusted life-years per participant than early structured physiotherapy, 95% confidence interval -0.0008 to 0.0800 more quality-adjusted life-years). The resulting incremental cost-effectiveness ratio for manipulation under anaesthesia was £6984 per additional quality-adjusted life-year. Arthroscopic capsular release was substantially more costly than early structured physiotherapy [on average costing £1733.78 more per participant (95% confidence interval £1529.48 to £1938.06 more per participant)] for a slight benefit in utilities [mean 0.0103 more quality-adjusted life-years per participant than early structured physiotherapy (95% confidence interval -0.0304 to 0.0510 more quality-adjusted life-years)]; the incremental cost-effectiveness ratio was $> £100,000$ per additional quality-adjusted life-year. Arthroscopic capsular release was more expensive than manipulation under anaesthesia and resulted in slightly fewer quality-adjusted life-years. Manipulation under anaesthesia was the intervention most likely to be cost-effective at a threshold of £20,000 per quality-adjusted life-year (manipulation under anaesthesia, 86%; early structured physiotherapy, 14%; arthroscopic capsular release, 0%).

Qualitative study

Forty-four interviews (most of which were conducted over the telephone) were undertaken with trial participants, who were evenly distributed across the three interventions, and with eight surgeons and physiotherapists. Trial participants described how frozen shoulder had a major impact on all aspects of their life. They were keen to get their shoulder problems resolved, which motivated them to participate in the trial. They thought that seeking early medical help and having a quicker NHS care pathway were important. In general, trial participants were satisfied with the trial interventions and found them acceptable. They reported improvements in pain, shoulder movements and function. Participants who received arthroscopic capsular release described recovering quicker than they had expected. Surgeons and physiotherapists followed a stage-based treatment approach in their routine practice. Both groups felt that people with diabetes tend to have poorer outcomes. They suggested that hydrodilatation could have been a treatment arm of the trial. Both groups stated that some people who previously had received ineffective physiotherapy had not wanted to take part in the trial.

Systematic review

Nine studies were identified, including UK FROST, which provided by far the largest and most robust evidence. The number of participants in the other studies ranged from 26 to 136, and the studies were mostly single centre. All studies were rated as being at high risk of bias for blinding of participants and clinicians, and outcome assessment. Considerable heterogeneity of the interventions and generally limited evidence for many of the comparisons meant that only two studies could be pooled as part of a meta-analysis (UK FROST and one other trial) comparing long-term shoulder functioning for patients receiving either arthroscopic capsular release or physiotherapy. The pooled effect favoured arthroscopic capsular release; however, the second study provided little additional weighted information.

Conclusions

UK FROST has provided robust clinically relevant evidence that none of the three treatments was clearly superior in patient-reported shoulder pain and functioning at 12 months. Our specifically designed early structured physiotherapy pathway can be accessed quickly in the NHS and has lower costs. However, the likelihood of further treatment being required is higher with early structured

physiotherapy than with the surgical interventions. Manipulation under anaesthesia is the most cost-effective option, with an incremental cost-effectiveness ratio of £6984 per additional quality-adjusted life-year. Patients who receive arthroscopic capsular release are least likely to need further treatment, but arthroscopic capsular release is associated with relatively higher risks and costs.

To address the increasing popularity of hydrodilatation, and the paucity of rigorous evidence of hydrodilatation's effectiveness, a high-quality randomised controlled trial is recommended to compare hydrodilatation with early structured physiotherapy with steroid injection with manipulation under anaesthesia with steroid injection.

Trial registration

This trial is registered as ISRCTN48804508.

Funding

This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 24, No. 71. See the NIHR Journals Library website for further project information.

Chapter 1 Introduction

Frozen shoulder

Frozen shoulder (also known as adhesive capsulitis) occurs when the capsule, or soft tissue envelope, around the ball-and-socket shoulder joint becomes inflamed and then scarred and contracted. This makes the shoulder very painful, tight and stiff. It starts with pain, which increases in intensity as stiffness develops.¹ The exact cause of this condition is unknown. Reported associations include diabetes mellitus, cardiovascular disease, trauma, stroke, neurosurgery and thyroid disease.¹ In the absence of a known association, the condition is labelled by clinicians as 'idiopathic' or 'primary' frozen shoulder. The pathology of the capsule involves chronic inflammation, and proliferative fibrosis has been reported.² Myofibroblasts contribute to matrix deposition and fibrosis, with the underlying pathology considered similar to Dupuytren's contracture.^{2,3} The macroscopic appearance of these changes can be seen in the shoulder during arthroscopic visualisation of the rotator interval capsule. People with this condition may struggle with basic daily activities, suffer serious anxiety and have sleep disturbance due to shoulder pain. There is a tendency for spontaneous resolution, but recovery may be slow or incomplete. Even after an average of ≥ 4 years from onset, around 40% of patients can have mild to severe symptoms.⁴ Figure 1 illustrates the pathology of frozen shoulder.

Three clinical phases have historically been recognised for this condition,⁵ where the duration of each phase is indicative but varies considerably between patients:

- painful phase, which may last 3–9 months
- adhesive phase, with stiffness lasting for 4–6 months
- phase of resolution or 'thawing', lasting for 5–24 months.

These phases have considerable overlap, and therefore the current favoured terminology is 'pain predominant' and 'stiffness predominant' phases.⁶

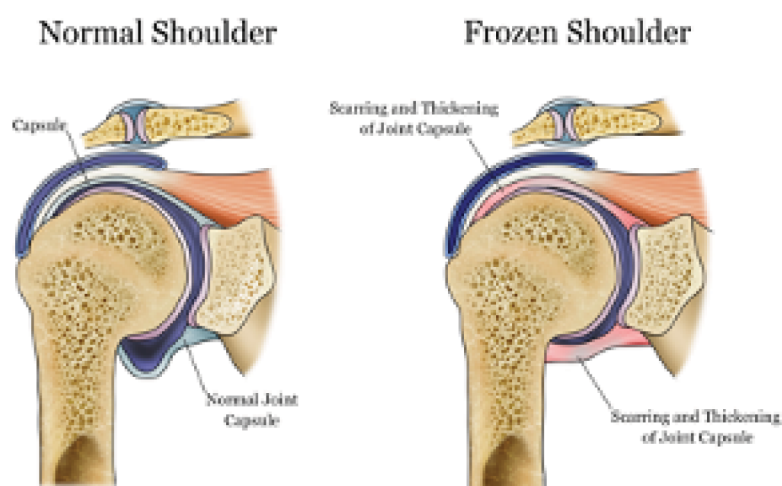


FIGURE 1 Diagram showing site of pathology of frozen shoulder. This figure has been re-used from www.local-physio.co.uk/articles/shoulder-pain/frozen-shoulder/ with permission from the copyright holders.

The cumulative incidence of frozen shoulder is estimated at 2.4 per 1000 population per year, based on a Dutch general practice sample.⁷ Frozen shoulder most commonly affects individuals in their sixth decade of life, and a large primary care-based study in the UK found that frozen shoulder affects 8.2% of men and 10.1% of women of working age.⁸ By contrast, an incidence of 1% was reported by a UK shoulder surgeon in his specialist hospital practice.⁹ This discrepancy in estimated incidence can be explained by the use of different populations as the denominators in the different studies.⁶ Although not clearly established, when associated with diabetes mellitus, frozen shoulder is considered to be more resistant to treatment.¹⁰

Diagnosis of frozen shoulder

A diagnosis of frozen shoulder is based on clinical criteria that include history of insidious onset deep-seated pain in the shoulder and upper arm with increasing stiffness, as well as clinical findings of limited active and passive external rotation in the absence of crepitus.¹¹ Radiographs are reported as not routinely required,⁶ but are usually performed in secondary care to exclude pathology like glenohumeral arthritis or posterior glenohumeral dislocation that could manifest with similar clinical signs. There is no reference standard for comparison, which explains the lack of diagnostic test accuracy data.¹¹ The key examination findings were originally described by Codman as restriction of elevation and external rotation.¹² As visual estimation of external rotation has fair to good reliability,¹³ restrictions (typically with pain) in both passive and active external rotation have been used as diagnostic criteria in clinical studies.^{11,14-17} It can be difficult, however, to correctly diagnose the problem, as highlighted in a qualitative study of patients' perceptions and priorities when living with primary frozen shoulder.¹⁸ This accords with other studies, which have found that general practitioners (GPs) in the UK and the USA lack confidence in making shoulder diagnoses.^{19,20}

Treatments for frozen shoulder

The aims of treating a patient with frozen shoulder are to provide advice, education and reassurance; achieve pain relief; improve shoulder mobility; reduce the duration of symptoms; and facilitate return to normal activities.²¹ Generally, less invasive treatments are provided in a primary care setting in the UK to those in the earlier phases of the disease, particularly for controlling pain. These may include oral analgesia, physiotherapy, acupuncture, and glucocorticoid (steroid) injection.²¹ The treatments utilised in secondary care, when stiffness has become more established, were confirmed in a UK survey of health professionals conducted in 2009 as physiotherapy, manipulation under anaesthesia (MUA) and arthroscopic capsular release (ACR).²²

Physiotherapy treatment includes combinations of advice, exercises, therapist-applied mobilisation techniques, and thermo- and electrotherapies. The modalities of treatment recommended for use are described in the UK national physiotherapy guidelines for frozen shoulder,⁶ which are based on a systematic review. These are provided either in isolation or as a supplement to other interventions, such as intra-articular corticosteroid injection or surgical interventions (MUA or ACR). Intra-articular corticosteroid injection helps reduce inflammation of the joint capsule and reduce pain, which may facilitate the performance of exercises and hence enhance the effects of physiotherapy. Intra-articular corticosteroid injection has been shown to provide short-term benefits, with better improvements in pain, function and range of movement (up to 6–7 weeks for all three improvements) than placebo¹³ and probably than isolated manual therapy and exercise.⁶

Manipulation under anaesthesia is a procedure the surgeon undertakes when the patient is under general anaesthesia. The affected shoulder joint is manipulated in a controlled fashion to stretch and tear the tight shoulder capsule. The joint is often injected with corticosteroid as part of this procedure. MUA is thought to facilitate recovery by releasing the tightness in the capsule, with the injection

helping control capsular inflammation and pain. This is followed by physiotherapy for mobilisation of the arm and shoulder to restore mobility and function.

Arthroscopic capsular release is a 'keyhole' surgical procedure performed under general anaesthesia. The keyholes are used to view the joint and divide (release) the contracted capsule using typically arthroscopic radiofrequency ablation. This is thought to allow more accurate and controlled release of the tight capsule. The procedure is completed by performing MUA to complete and confirm full release of the contracted capsule. ACR is also followed by physiotherapy of the arm and shoulder to restore mobility and function.

Rationale for the UK FROzen Shoulder Trial

It is unknown whether a combination of physiotherapy and steroid injection or either of the surgical interventions (MUA or ACR) followed by physiotherapy is more effective.¹³ Similarly, there is uncertainty about the benefits of MUA compared with other treatment options,^{23,24} and only limited evidence on ACR is available from randomised controlled trials (RCTs).^{13,25}

Systematic reviews have identified large gaps in the evidence base and uncertainty about the effectiveness of treatments for frozen shoulder and, therefore, a need for high-quality primary research.²⁶ In a systematic review¹³ commissioned by the National Institute for Health Research Health Technology Assessment (HTA) programme, 28 RCTs, one quasi-experimental study and two case series were included. The review found insufficient studies with a similar intervention and comparator to quantify effectiveness. Most studies were rated as having a high risk of bias, did not report adequate methods of randomisation, allocation concealment and outcome assessment, and seemed to be inadequately powered. Few studies reported collecting data on harms.

In view of the paucity of high-quality evidence to guide current practice, considerable uncertainties remain about the management of frozen shoulder. With the intention of facilitating quicker recovery, more invasive surgical interventions (MUA and ACR) are being used in spite of the lack of good evidence.¹³ There is a clear need for a well-designed, high-quality RCT to determine the clinical effectiveness and cost-effectiveness of commonly used interventions for the treatment of frozen shoulder.

The findings of a national survey²² of health-care professionals in the UK, conducted in 2009, were used to determine the most commonly used interventions that need to be tested in a RCT in a secondary care setting. Physiotherapy, MUA and ACR were those that health-care professionals recommended be compared in a RCT. Only 6% of respondents at the time suggested using hydrodilatation as a comparator in a trial, which did not make this a feasible intervention to test in a RCT. This survey informed our decision to compare early structured physiotherapy (ESP) combined with intra-articular steroid injection with the two most frequently used, invasive and costly surgical interventions, namely MUA and ACR.²² It is important to emphasise that, although physiotherapy is a common treatment in NHS practice, the ESP intervention was a specifically designed and standardised physiotherapy pathway to test the optimal delivery of physiotherapy in the NHS. As evidence about patients' experiences of frozen shoulder is also limited,¹⁸ participants were interviewed about their experience and the acceptability of treatment, as were health professionals (physiotherapists and surgeons).

Aim and objectives

The strategic aim of UK FROST (UK FROzen Shoulder Trial), underpinned by the key treatment uncertainties, was to provide evidence of the clinical effectiveness and cost-effectiveness of three

INTRODUCTION

common interventions currently provided in the UK NHS for the treatment of frozen shoulder in a hospital setting. The following objectives were defined to achieve this overarching aim:

1. The primary objective was to determine the effectiveness of ESP compared with MUA compared with ACR for patients referred to secondary care for the treatment of frozen shoulder. This was achieved using a parallel-group RCT, with the Oxford Shoulder Score (OSS) (a patient-reported outcome measure) as the primary outcome at 3, 6 and 12 months. The primary time point was 12 months after randomisation.
2. To compare the cost-effectiveness of the three interventions, to identify the most efficient provision of future care, and to describe the resource impact that the various interventions for frozen shoulder would have on the NHS.
3. To qualitatively explore the acceptability of different interventions for frozen shoulder to patients and health-care professionals and to provide important patient-centred insight to further guide clinical decision-making.
4. To update the HTA programme-funded systematic review examining the management of frozen shoulder by assessing current RCT evidence for the effectiveness of interventions used in secondary care. This would allow the trial findings to be considered in the context of the existing evidence for the interventions under evaluation.
5. To widely disseminate the findings of this study to all stakeholders through networks of health-care professionals, patients, health service managers and commissioning groups. This would be in addition to publishing the results of the study in key journals and publishing the National Institute for Health Research HTA report.

Chapter 2 Trial design and methods

This chapter describes the trial design and methods used to address the objectives about the clinical effectiveness of the health-care interventions being compared. The methods of the health economic evaluation and the nested qualitative study are described in the corresponding chapters. The trial protocol has been published.²⁷

Trial design

This was a pragmatic, multicentre, stratified (diabetes present or not) superiority trial comparing three parallel groups (MUA vs. ACR vs. ESP, with unequal allocation of 2 : 2 : 1) in adult patients referred to secondary care in England, Wales and Scotland for the treatment of primary frozen shoulder, and for whom surgery was being considered.

Participants

Patients with primary frozen shoulder were identified through clinical examination and plain radiographs.²⁸ To minimise diagnostic uncertainty, clinical examination included the key diagnostic assessment of restriction of passive external rotation in the affected shoulder,²⁹ for which there is evidence of good inter-rater agreement on whether or not restriction is present³⁰ and a high threshold (50% restriction of movement) for inclusion. Plain radiographs (anteroposterior and axillary projections) were obtained routinely for all patients to see whether or not these were normal and could allow the exclusion of glenohumeral arthritis and other pathology that could lead to similar clinical presentation (e.g. locked posterior dislocation).

Inclusion criteria

Patients, including those with diabetes, were eligible if:

- they were aged ≥ 18 years
- they presented with a clinical diagnosis of frozen shoulder characterised by restriction of passive external rotation in the affected shoulder to $< 50\%$ of that of the contralateral shoulder
- they had radiographs that excluded other pathologies.

Exclusion criteria

Patients were excluded if:

- they had a bilateral concurrent frozen shoulder
- their frozen shoulder was secondary to trauma that necessitated hospital care (e.g. fracture, dislocation, rotator cuff tear)
- their frozen shoulder was secondary to other causes (e.g. recent breast surgery, radiotherapy)
- any of the trial treatments (e.g. unfit for anaesthesia or corticosteroid injection) were contraindicated
- they were not resident in a catchment area of a trial site
- they lacked the mental capacity to understand the trial.

Setting

The trial recruited from the orthopaedic departments of 35 NHS hospitals in the UK across a range of urban and rural areas. This comprised 28 hospitals in England, six in Scotland and one in Wales. Two additional hospitals in England screened for patients but did not recruit into the trial.

Recruitment started in April 2015 and the final follow-up was in December 2018. All 37 participating hospitals are listed in *Appendix 1*.

Interventions

The components and standardisation of the surgical trial interventions were informed by a survey of 53 surgeons, who were the principal investigators of two multicentre shoulder surgical RCTs.^{31,32} The standalone physiotherapy and post-procedural physiotherapy programmes were developed using evidence from a systematic review,¹³ UK guidelines,⁶ previous surveys of UK physiotherapists^{33,34} and consensus of expert shoulder physiotherapists in secondary care derived from a Delphi survey specific to UK FROST.¹⁸ Ethics approval for the last of these was obtained from the School of Health and Social Care Research Governance and Ethics Committee of Teesside University on 23 May 2014 (Research Ethics Committee reference 069/14). The physiotherapy programmes developed are available online.³⁵ It is important to emphasise that, although physiotherapy is a common treatment in NHS practice, the ESP intervention was a specifically designed, standardised, new physiotherapy pathway to test the optimal delivery of physiotherapy in the NHS based on the best available evidence and expert consensus.

Participants assigned to either of the two surgical procedures were placed on the surgical waiting list and underwent routine preoperative screening. In keeping with NHS waiting time targets, both surgical procedures were expected to be performed within 18 weeks of randomisation. Participants would undergo these procedures under general anaesthetic and were expected to be admitted as day cases.

Physiotherapy was delivered by qualified physiotherapists (i.e. not students or assistants), and participating surgeons were familiar with the surgical procedure(s). There was no minimum number of surgical procedures that the surgeon had to have performed, and no grades of surgeon were excluded. Which surgeon operated on participants and whether or not the individual surgeon needed to be supervised by a consultant was at the discretion of the participating site, and followed normal care pathways and practices. The experience of physiotherapists and surgeons delivering the trial treatments was quantified and recorded in terms of their salary bands and the number of frozen shoulder patients they treated in a typical month.

Manipulation under anaesthesia with an intra-articular steroid

The affected shoulder was manipulated to stretch and tear the tight capsule and to improve range of movement. Intra-articular injection of corticosteroid to the glenohumeral joint was to be administered while the participant was under the same anaesthetic, unless the injection was contraindicated at the time of surgery. Postoperative analgesia, including nerve blocks, was provided as per usual care in the treating hospital. The details of MUA were collected prospectively using the MUA surgery form (see *Report Supplementary Material 1*). In the unlikely event that the MUA was judged to be incomplete, it was recommended that the surgeon should not cross over intraoperatively to capsular release. The need for this was to be reviewed at another clinic appointment to allow the outcome of the MUA to be assessed and the need for any further intervention to be decided. Details of any further intervention were collected prospectively.

Arthroscopic capsular release with manipulation under anaesthesia

Arthroscopic release of the contracted rotator interval and anterior capsule was performed, followed by MUA to complete the release of the inferior capsule. Additional procedures such as posterior capsular release or subacromial decompression were permitted at the operating surgeon's discretion. Steroid injections, which slightly increase the risk of infection and morbidity, were permitted at the surgeon's discretion.³⁶ Postoperative analgesia, including nerve blocks, was provided as per usual care in the treating hospital. The details of ACR were collected prospectively using the ACR surgery form (see *Report Supplementary Material 2*).

Nested shoulder capsular tissue and blood samples study

At six selected hospitals, 16 participants allocated to ACR who had not had a steroid injection within 6 weeks from the day of surgery were included in an exploratory nested capsular tissue and blood study. This was undertaken between January 2017 and December 2017 and had the following objectives:

1. to determine molecular and cellular abnormalities in tissue obtained during surgery from patients with frozen shoulder
2. to determine serum protein and cytokine signatures in patients with frozen shoulder
3. to correlate any tissue and serum abnormalities detected with clinical presentation and response to treatment.

When the date of surgery was known, the research nurse posted a letter about the nested study, a patient information leaflet and a consent form. Written informed consent was obtained at the participant's pre-surgery assessment. A tissue sample of capsule from the rotator interval, which is routinely incised or removed as part of ACR, and a venous blood sample were obtained for analysis. All samples were fresh frozen, stored on dry ice and transported securely by courier to the Oxford Musculoskeletal Biobank at the University of Oxford, and housed at the Botnar Research Centre for formal analysis. The biopsy material was small (2 mm by 2 mm) and obtained with the use of arthroscopic graspers, and biopsy was not expected to have any significant effect on patient outcomes. The results of this study have been published.³⁷

Early structured physiotherapy

Participants received up to 12 sessions of structured physiotherapy, comprising essential 'focused physiotherapy' and optional supplementary physiotherapy, over a period of up to 12 weeks. The focused physiotherapy package included an information leaflet (see *Report Supplementary Material 3*) providing education and advice on pain management and function; an intra-articular steroid injection; and hands-on mobilisation techniques, increasingly stretching into the stiff part of the range of movement as the condition improved.^{38,39} Participants received supervised exercises and were provided with instructions for a graduated home exercise programme (see *Report Supplementary Material 4*), progressing from gentle pendular exercises to firm stretching exercises according to stage, as is accepted good practice. All participants randomised to ESP underwent all elements of the focused physiotherapy package unless there was a specific clinical reason for them not to do so (e.g. a steroid injection might be withheld from a participant with currently uncontrolled diabetes, or from a participant with a stiff but painless and non-irritable shoulder).

Supplementary physiotherapy comprised those interventions that were non-essential but permissible additions, allowing physiotherapists some flexibility. These interventions, which may have been omitted from the national guidelines because they were outside their scope (e.g. acupuncture) and/or because there was a lack of evidence of their effectiveness in the primary academic literature (e.g. hydrotherapy, soft-tissue release techniques), were explored using a Delphi process.

Participants who did not improve with ESP were referred for further treatment in consultation with the treating clinician following a 12-week assessment. When further treatment after ESP involved surgical intervention, participants were placed on the normal surgical waiting list. Any further treatment provided was recorded. Participants allocated to ESP were offered reimbursement of their travel expenses. The ESP given during each session (e.g. injection, advice and education, gentle active exercise) was recorded in the structured physiotherapy log book (see *Report Supplementary Material 5*).

Post-procedural physiotherapy

Following MUA or ACR, participants underwent up to 12 weeks of physiotherapy, normally commencing within 24 hours of the procedure. The aim was to reduce pain and aid with regaining/maintaining the mobility achieved by the operation. The post-procedural physiotherapy (PPP) differed from ESP to suit its very different context.

As the research literature was uninformative, two essential 'focused physiotherapy' interventions were prespecified, based on established good practice. These were:

1. an information leaflet giving education and advice on pain management and function
2. instructions for a graduated home exercise programme.

All participants randomised to MUA or ACR were to undergo all elements of this focused physiotherapy package unless there was a specific clinical reason for them not to do so. The Delphi survey, which was interpreted as had been done for ESP, provided optional, supplementary interventions. A steroid injection was to be avoided where possible during PPP. The PPP log book (see *Report Supplementary Material 6*) was used to record the PPP given during each session.

Steroid injections

Steroid injections were administered with or without imaging guidance, depending on the usual practice of the hospital site. Current evidence does not support the superiority of either approach.⁴⁰

Modifications to interventions

There were no explicit criteria for modifying, discontinuing or crossing over from the assigned trial treatment. The clinician and participant discussed whether or not to continue with the assigned treatment for reasons such as the patient having poorly controlled diabetes or no longer requiring the treatment.

Adherence to interventions

Adherence to the trial treatments was explained in the trial site manual and during site initiation visits. A requirement of the internal pilot was to check the feasibility of delivering the ESP programme. This was expanded to include the surgical interventions and PPP. Every month, a designated trial co-ordinator extracted data from the hospital case report forms (CRFs) and updated a spreadsheet to record information about aspects of the treatments. The spreadsheet was reviewed for treatment adherence by the chief investigator, a consultant orthopaedic surgeon and the lead physiotherapist, who decided whether or not any action was required at a site. This was further monitored by the Trial Management Group (TMG), the independent Trial Steering Committee (TSC) and the Data Monitoring and Ethics Committee (DMEC).

Concomitant care

Analgesia for pain relief, general advice on care of the arm (e.g. axillary hygiene) and general advice to prevent further stiffness in the limb were all permitted at part of the management of a participant awaiting surgery. Specific home exercise programmes, such as that provided with the structured physiotherapy intervention, were not permitted. Steroid injections were avoided, as these were considered active interventions.

Outcomes

Primary outcome

The primary outcome was the OSS, a patient-reported measure of functional limitation following shoulder surgery. The development and validation of this measure included patients with frozen shoulder,⁴¹ and it has been used in the follow-up of these patients.⁴ The OSS is a 12-item measure with five response categories and a range of scores from 0 (worst) to 48 (best).⁴² It has been validated against the professionally endorsed Constant score⁴³ and the SF-36 (Short Form questionnaire-36 items),⁴⁴ and its responsiveness over a 6-month period following surgical intervention has been established.⁴⁵

Participants completed the OSS at baseline prior to randomisation. The questionnaire was then posted to the participants 3, 6 and 12 months after randomisation. The primary end point was 12 months after randomisation, allowing the interventions and co-treatment interventions to be delivered and

the majority of complications to be treated. The OSS was also completed at the hospital at the start of treatment. This was either the day of the operation or, for participants allocated to ESP, the day when the steroid injection was given or at the first visit to physiotherapy, whichever was first. The OSS was then posted to participants for them to complete 6 months from when treatment started.

Secondary outcomes

Secondary outcomes were collected at baseline and at 3, 6 and 12 months from randomisation, unless otherwise stated.

Quick Disabilities of the Arm, Shoulder and Hand

The DASH (Disabilities of the Arm, Shoulder and Hand) is a well-validated and reliable measure of symptoms and functional limitation in the upper extremity.⁴⁶ To minimise responder burden, the validated 11-item short version, the QuickDASH (Quick Disabilities of the Arm, Shoulder and Hand), was used.⁴⁷ This is scored from 0 to 100, and an 8-unit improvement in scores has been defined as the minimum clinically important difference for patients with shoulder problems.⁴⁸ Its validity with and responsiveness to frozen shoulder has been established.⁴⁹

EuroQol-5 Dimension, five-level version

The EuroQol-5 Dimensions (EQ-5D) is a validated, generic and health economic, self-completed, patient-reported outcome measure covering five health domains with three response options.^{50,51} The EuroQol-5 Dimensions, five-level version (EQ-5D-5L), consists of the same five domains as the original EuroQol-5 Dimensions, three-level version (EQ-5D-3L), but with five levels rather than three to help overcome problems with ceiling effects and to improve sensitivity.^{52,53} The EQ-5D-3L has been validated for use with a range of shoulder conditions.^{54,55} The EQ-5D-5L provides a simple descriptive profile of health status that can be used to estimate quality-adjusted-life-year (QALY) scores in economic evaluations.

Pain

Shoulder pain 'during the past 24 hours' was measured using the Numerical Rating Scale for pain,⁵⁶ a single 11-point numerical scale on which 0 represents 'no pain' and 10 represents 'worst possible pain'. This measure is considered the most valid for use in this population.⁵⁷

Extent of recovery

A simple subjective global question asked to what extent the participant's frozen shoulder symptoms in the past 24 hours had affected their assessment of needing treatment. This informed the extent to which symptoms resolved over time. Responses were measured using a visual analogue scale with anchors from 0 to 100 (e.g. 0, no need to ask for treatment; 100, definitely ask for treatment).

Complications

At 12 months, sites recorded all expected and unexpected complications on the 52-week complication forms (see *Report Supplementary Material 7*). Infection was defined as for the 'surgical site infection' audit.⁵⁸ Delayed wound healing was defined as any wound that had not healed by 2 weeks post surgery. Complex regional pain syndrome was defined as pain, swelling and stiffness of the affected shoulder, and arm and/or hand restrictions limiting the full tuck of the fingers. In addition, nerve, blood vessel, tendon or bone injury and complications related to steroid injection, including steroid flare and septic arthritis, were recorded.

Adverse events

A non-serious adverse event (AE) was defined as any untoward medical occurrence in a trial participant related to the affected shoulder up to 12 months from randomisation. A serious adverse event (SAE) was defined as any untoward medical occurrence that resulted in death, threat to life, hospitalisation or prolongation of existing hospitalisation, persistent or significant disability or incapacity, a congenital abnormality or birth defect, or any other medical condition that may require medical or surgical intervention to prevent any of these from occurring.

Sample size

The primary trial outcome was the OSS, and this was assessed for three treatment comparisons: ESP with MUA, ESP with ACR, and MUA with ACR.

Data suggest that a 5-point improvement can be found on the OSS (standard effect size of 0.42) between surgically and non-surgically treated patients,⁵⁹ with a stable standard deviation (SD) of 12 points across different populations. This larger effect size was required to justify the higher costs and potential risks associated with surgery when comparing ESP with MUA, and ESP with ACR.⁴² A smaller difference of 4 points on the OSS (standard effect size of 0.33) was expected to distinguish between MUA and ACR.

To observe the above effect sizes with 90% power and 5% two-sided significance, adjusting for a moderate estimate ($r = 0.4$) of the correlation between OSS over 12 months and allowing for 20% attrition, a total sample size of 500 patients was required (MUA, $n = 200$; ACR, $n = 200$; ESP, $n = 100$). The sample size calculation was not adjusted for multiple comparisons owing to the a priori specified sequence of treatment comparisons and the analysis of the primary outcome in a single analysis model.⁶⁰

There were no planned interim analyses for the trial or stopping guidelines. An internal pilot from which data contributed to the final analyses was performed to confirm the feasibility of the trial, and this is explained in the following section.

Internal pilot study

There were two phases of the internal pilot study.

Phase 1 (months 4–9)

It was important to critically test our assumptions after 6 months of recruitment by reviewing the number of sites set up and the number of eligible patients identified, approached and consented. This was to help inform the number of participating sites required to achieve the recruitment target. Secondary reasons for undertaking this phase of the pilot were to review (1) whether or not the participating sites were being provided with enough training and documentation; (2) the number of reasons why patients were not eligible for the trial; (3) the length of time it took to consent a patient and the reasons why patients did not take part; (4) whether or not all clinicians at a site were actively taking part in the trial, and, if not, why not; and (5) patient adherence to treatment allocation.

The independent oversight committees assessed the success of phase 1 based on the following objectives:

- to have a minimum of four sites recruiting during the 6 months that had recruited 24 patients (i.e. evidence that sites could recruit the expected one participant per month)
- to ensure that adequate progress was made with setting up other sites to recruit in order to have 12 sites set up (i.e. 50% of sites).

Phase 2 (months 10–27)

This phase of the internal pilot continued for a further 18 months and was reviewed at 6-monthly intervals with the independent oversight committees. Patients were likely to have already suffered from frozen shoulder for several months and received physiotherapy in primary care before their referral to hospital. There was concern that this could have an impact on patient consent and adherence to the ESP intervention, which were threats to both the feasibility and the validity of the trial. Evidence from simulation work was that, with 80% power, a true treatment effect size of 0.2 or 0.4 and 30% non-compliance, the power is reduced to 54%.⁶¹ In UK FROST, with a sample size that had 90% power and effect sizes of around 0.3–0.4, 20–30% non-compliance in the ESP arm was expected to reduce the power to between 60% and 70%. Therefore, if at the 24-month review, when 50% of the patients were

expected to have been recruited, the non-compliance in the ESP arm was between 20% and 30%, the oversight committees would advise on whether to continue with a three-arm trial or with the surgical comparisons only. The following were also monitored:

- reasons for patient non-consent to participate in the trial, their treatment preferences and a member of the trial team to informally discuss this with willing patients
- whether or not all 25 sites were set up and had recruited 250 patients (50% of our target)
- waiting times at sites from randomisation to intervention, with consideration of the need to replace sites that were not meeting the waiting time targets agreed in the protocol (i.e. the surgical procedure being performed within 18 weeks of randomisation).

Recruitment

Initial estimates for recruitment were based on Hospital Episode Statistics for NHS hospitals in England in 2009/10 and 2010/11. These excluded post-trauma or secondary referrals from other specialties, giving a stable rate of 210 per million patients treated for frozen shoulder. Assuming that 50% of frozen shoulder patients presenting in secondary care met the inclusion criteria and 40% of these consented, this left around 40 patients per million to be recruited into the trial. It was estimated that, to recruit 500 trial participants from trusts each serving catchment areas of around half a million people, 25 hospitals would be required to recruit for a minimum of 1 year. This assumed that there would be no delays in set-up or problems at any subsequent time point, that all surgeons at the sites would be willing to participate and that all potential participants would be screened for eligibility. Following the pilot phase, the number of sites required was increased to ensure that recruitment was achieved to target.

Patients who had been referred for a frozen shoulder to an outpatient hospital clinic were identified by the research nurse or assessing clinician. In the clinic, a designated individual within the shoulder team (e.g. surgeon or physiotherapist) completed the study eligibility form (see *Report Supplementary Material 8*) to confirm whether or not the patient was eligible and, when applicable, approached the patient about the study. The research nurse then provided the patient with an information sheet (see *Report Supplementary Material 9*) and answered any questions. The patient was able to consent at that time or they could take up to 1 week to decide. When a patient consented (see *Report Supplementary Material 10*), he or she was asked to complete the baseline form (see *Report Supplementary Material 11*). The research nurse completed the consent status form (see *Report Supplementary Material 12*) to confirm the patient's status.

When a patient did not consent, the research nurse recorded the reason and the treatment plan in another section of the consent status form. The patient was also offered an optional patient preference form (see *Report Supplementary Material 13*) to complete if they wanted to provide more information about why they chose not to take part.

Training in recruitment was provided to hospital staff as part of the site initiation visit, and a trial site manual was prepared that included guidance on consenting patients into the trial and how to answer questions that might arise during consent-taking. In addition, a poster was provided to publicise the trial to hospital staff and patients. During the trial, training and reminders were implemented using regular e-mail bulletins and face-to-face meetings with principal investigators and research nurses, with trial co-ordinators providing support and guidance to staff as required.

Randomisation

The randomisation sequence was based on a computer-generated randomisation algorithm provided by a remote randomisation service (telephone or online access) at York Trials Unit, University of York. The unit of randomisation was the individual patient, allocated to the trial interventions MUA,

ACR and ESP in the ratio of 2 : 2 : 1, stratified by the presence of diabetes,⁶² using random blocks sizes of 10 and 15. The research nurse used the remote randomisation service to register eligible and consenting patients before computer generation of the allocation. This ensured treatment concealment and immediate unbiased allocation.

The research nurse then informed the treating clinician and the patient of the treatment allocation.

Blinding

Given the nature of the trial treatments, comparing surgical and non-surgical treatment options, the blinding of participants and clinicians to treatment allocation was not possible or desirable in this pragmatic trial. Therefore, patients and clinicians were informed of the treatment allocation after randomisation. The statistician was blind to treatment allocation until after data were hard locked and no further changes could be made.

Statistical methods

Analyses were conducted for the three treatment comparisons of interest, ACR with ESP, MUA with ESP, and ACR with MUA, according to the principle of intention to treat (ITT). All analyses were conducted in Stata[®] version 15 (StataCorp LP, College Station, TX, USA) using two-sided statistical significance at the 0.05 level. The statistical analyses plan was completed prior to completion of data collection on 12 February 2019.

Trial progression

The characteristics (age, sex, diabetes, symptom duration, laterality and patient preferences) of ineligible and non-consenting patients were compared with the randomised patient population. Reasons for exclusion and non-consent were tabulated, including free-text entries summarised by the trial team. The agreed treatment for excluded patients was tabulated. The flow of participants from eligibility and randomisation to follow-up and analysis of the trial was presented in a Consolidated Standards of Reporting Trials (CONSORT) flow diagram.

Baseline characteristics

All participant baseline characteristics were summarised descriptively by trial arm both for participants 'as randomised' and for those 'as analysed'. The 'as analysed' population comprised all participants included in the primary analysis (i.e. patients who had complete data for the baseline covariates and outcome data for at least one post-randomisation time point). No formal statistical comparisons were undertaken between arms. Continuous measures were summarised using numbers, mean, SD, median, minimum and maximum, and categorical data were reported as counts and percentages.

Intervention delivery/fidelity

Details of the interventions as delivered were presented, including time to treatment, receipt of steroid injections, and optimal or suboptimal release achieved during surgery, as well as number and content of physiotherapy sessions. Fidelity was reported descriptively by trial arm, with baseline characteristics tabulated for each arm. Reasons for not receiving randomised treatment, alternative treatments and any further recorded treatments were tabulated by trial arm. Caseload by site and surgeon/physiotherapist were reported descriptively. The grades/bands and experience of treating surgeons and physiotherapists were presented.

Missing data

Item-level missing data for individual outcomes (OSS and QuickDASH) were managed according to the instrument scoring guidance, and patterns of missing items were reported by trial arm. As the follow-up

dates for the 6-month CRF and 6-month post-treatment CRF were in proximity for some participants, OSS data from these CRFs were used as a substitute if data were available for one and missing for the other, and if the two CRFs had been sent to the participant within 4 weeks (28 days). Missing baseline covariates for the primary analysis were imputed for the purpose of the analysis if participants provided follow-up data for at least one time point. Two participants with follow-up data had missing OSS baseline scores. Using the QuickDASH as a proxy, their scores were imputed as the median OSS of any participants with the same QuickDASH value.

Primary outcome (Oxford Shoulder Score) analysis

The OSS was summarised descriptively at each collected time point by trial arm, and mean scores and confidence intervals (CIs) were illustrated graphically.

The primary analysis was conducted on an ITT basis, including patients in the treatment arm to which they were randomised. The primary analysis compared the OSS between treatment arms at 12 months. The primary outcome, OSS, was analysed using a covariance pattern linear mixed model, including assessments at all available time points with reference to the date of randomisation (3, 6 and 12 months, thereby increasing power) and treating patients as a random effect. The model was adjusted for OSS at baseline and included as further fixed effects: treatment arm, time, arm-by-time interaction, age, sex and diabetes. Differences in local practice and expertise were accounted for by including recruitment site as a random effect in the model. Given the low individual practitioner caseload (designated surgeon or physiotherapist in the shoulder team) expected in this multicentre trial, surgeons or physiotherapists were not specifically adjusted for.

For the modelling of repeated measurements, the best-fitting (based on Akaike information criterion and Bayesian information criterion), simple (not significantly different from an unstructured pattern) covariance pattern was selected. For all three treatment comparisons, the model provided estimates at individual time points (the estimate at the 12-month time point served as the primary end point for each of the three treatment comparisons), as well as an overall treatment effect over 12 months. These are reported as mean differences between treatment arms, with 95% CIs and associated *p*-values.

Data were assumed missing at random (MAR). Model assumptions were checked, and, if they were in doubt, the data were transformed prior to analysis or alternative non-parametric analysis methods were explored.

Secondary analyses

Analysis adjusted for treatment compliance

To take account of an expected degree of participant non-compliance with the allocated treatment, a secondary complier-average causal effect (CACE) analysis was carried out. This retains the initial randomised assignments but overcomes the problems of a per-protocol analysis. Given the three active treatments under investigation with different adherence criteria and the multiple alternative treatment pathways for each participant, not all comparisons were suitable for CACE analysis. Therefore, only compliance with ESP (minimum of eight ESP sessions or participant/physiotherapist satisfied with progress) was assessed using instrumental variable regression, predicting OSS at the primary end point at 12 months. The analysis adjusted for covariates of the primary analysis model. Assuming that the same proportion of participants in the comparator arm would have adhered to the intervention if they had been offered it (which should be achieved by way of randomisation), the group differences from this model provided an estimate of the treatment effect among participants who adhered to the treatment.

Analysis adjusted for waiting times

A separate secondary ITT random intercept linear mixed-model analysis including pre-treatment OSS and OSS 6 months from the start of treatment in addition to the 3- and 6-month post-randomisation data was conducted, including the same covariates as the primary analysis. Time was included as a

continuous variable in order to explicitly model participant trajectories over time using all available data and thereby explore the influence of variable waiting times on the results of the study. Treatment effect estimates and p -values were derived at 3, 6 and 12 months post randomisation.

Missing data

The extent and pattern of missing outcomes over time were explored by trial arm. Logistic regression models were used to identify predictors of non-response and included all baseline data and primary outcome assessments before any missing values as potential predictors. Any variables found to be predictive of non-response were included in a repeat of the model specified for the primary analysis. Analysis by multiple imputation was considered if missing data exceeded the planned level of attrition (i.e. at least 20% of missing total OSS scores at 12 months).

Analysis using data close to intended follow-up points

If > 5% of all questionnaires were returned outside their intended time of follow-up [general follow-up: on or after 6 weeks, i.e. after the telephone reminder; pre-treatment form (see *Report Supplementary Material 14*); day of operation or the earlier of first day of physiotherapy or steroid injection], then the primary analysis and analysis adjusted for waiting times were repeated with data from such questionnaires excluded.

Analysis adjusting for baseline imbalances

The UK FROST DMEC observed an imbalance of employment status between randomised treatment arms during the monitoring of the trial. On its recommendation, a binary variable of working status (working vs. not working) was included as a covariate in the same model as the primary analysis if it was found to be associated with the OSS outcome.

Subgroup analyses

To explore differences in treatment response for different participant populations, three planned exploratory subgroup analyses were conducted: one exploring the influence of whether or not the participant was diabetic (yes/no), one exploring whether or not the participant had been in previous receipt of physiotherapy (yes/no) and one exploring patient treatment preferences as expressed at baseline (allocated to preferred treatment/not allocated to preferred treatment/had no preference). In addition, the TSC proposed a further subgroup analysis based on the duration of frozen shoulder symptoms at baseline (using the median of less than/more than 9 months as the cut-off point). For each analysis, a treatment-arm-by-subgroup interaction term was included in the primary analysis model, and the p -value of the interaction term was reported along with descriptives of the primary outcome for each subgroup-treatment arm pairing.

Analysis of secondary outcomes

QuickDASH, pain, extent of recovery

Continuous secondary outcomes were reported descriptively (unadjusted mean, SD, median, minimum and maximum). ITT linear mixed models were conducted for each outcome, adjusting for the same covariates as the primary analysis.

Pain or stiffness

As part of physiotherapy, the participant's predominant problem, pain or stiffness, was recorded at each session. Equal pain and stiffness was classified, managed and recorded as pain. The proportion of each category at the first and last recorded physiotherapy session for each participant was presented by treatment arm.

Complications/adverse events

Based on the overlap between recorded complications and AE data, these data sets were reviewed, and a single list of serious and non-serious AEs was compiled to avoid duplication in reporting. These events

were then summarised by type for each treatment arm. A logistic regression model was used to determine treatment arm differences in having experienced at least one AE if the number of participants with one or more events exceeded 10 in each arm. The same covariates as those used in the primary analysis were adjusted for.

Other analyses

Treatment preferences

Patient and clinician treatment preferences were explored for non-consenting patients where this information was provided.

Baseline patient preferences and expectations of randomised patients were explored descriptively by trial arm as well as for patients who had and patients who had not received prior physiotherapy and for patients who did and patients who did not receive their allocated intervention. Any change in preferences was explored by tabulating participants' preferences at the 12-month follow-up against their baseline preferences and against their allocated treatment.

Oxford Shoulder Score change scores

Patients' comparative shoulder assessment at 12 months (e.g. slightly better or much better) was matched with their change in OSS between baseline and 12 months in order to explore the magnitude of meaningful differences in the outcome in the study population.

Oxford Shoulder Score subdomains

Exploratory and confirmatory factor analysis of OSS from a population of patients with rotator cuff tears in the UKUFF trial⁶³ identified reliable OSS subdomains of pain (items 1, 8, 11 and 12) and function (items 2, 3, 4, 5, 6, 7, 9 and 10). To further explore the nature of shoulder outcomes, descriptive statistics and associated graphs were presented for OSS subdomains of pain and function by treatment arm at each time point.

Outcomes for participants receiving no treatment

The OSS and QuickDASH scores were summarised descriptively at baseline and at all follow-up points for participants who did and participants who did not receive any treatment as indicated on their change in status form (see *Report Supplementary Material 15*). Where available, the average time to the decision of no treatment was reported for this group.

Update of systematic review

To place the trial findings in the context of current evidence, the HTA systematic review about the management of frozen shoulder was updated.¹³ MEDLINE/PreMEDLINE, CENTRAL, EMBASE, PEDro, Science Citation Index, Clinicaltrials.gov and WHO International Clinical Trials Registry were searched from January 2010 to December 2018 and studies reported prior to 2010 obtained from our previous HTA review. The updated review focused only on evidence from RCTs and the interventions and outcomes collected in UK FROST. Hydrodilatation, however, was also included, as its popularity has increased since a survey was undertaken to inform the design of UK FROST.²² Moreover, during the qualitative interviews with health-care professionals in the nested study, some surgeons and physiotherapists commented that hydrodilatation could have been a treatment option in the trial. The review protocol has been registered (PROSPERO CRD42019122999).

Data management

A central database at York Trials Unit was used to manage data collection, including the sending and return of participant questionnaires (see *Report Supplementary Material 16*) and hospital CRFs. This included automated e-mail reminders to participating sites to help ensure the timely return of

hospital CRFs. Participant questionnaires and hospital CRFs were designed using TeleForm software (version 10; Cardiff Software, Cambridge, UK) and marked up with variable names and appropriate scoring. To maximise data quality, when hospital CRFs were returned to York Trials Unit the key variables required for the statistical analysis and checking adherence in the delivery of the treatments were reviewed for completion and accuracy by a research data administrator, who resolved any queries with the research nurse at the site. The hospital site was reimbursed for the completion of all CRFs up to a maximum value of £124.00. This was agreed by the trust and trial sponsor using a clinical trial agreement during the site set-up. No checks of the quality of data in the postal questionnaires were made on return of the questionnaires to York Trials Unit, although a trial co-ordinator checked whether the participant had given extreme responses to the last EQ-5D-5L question and/or given a free-text response that indicated they could be at harm. When either of these occurred, the principal investigator, research nurse and chief investigator were notified by e-mail. After this initial check, all postal questionnaires and hospital CRFs passed through a process of scanning in the TeleForm software, second checking and validating against predetermined rules.

Active and systematic follow-up of all randomised participants by post included pre-notification reminders, 2- and 4-week letter reminders and the option to complete an abridged questionnaire (a minimum of the OSS and EQ-5D) over the telephone after 6 weeks. At 12 months, the primary end point, an unconditional incentive of £5 was included. If the patient agreed at the time of consent, text messages were sent on the day the participant was sent the postal questionnaire⁶⁴ and newsletters were circulated to trial participants.⁶⁵ Trial participants could withdraw entirely from the study at any time for any reason, but any data collected up to that point were included in the analysis. The participant could agree to being withdrawn from only postal questionnaire collection or from only hospital CRF collection.

Essential trial documentation were kept with the trial master file and investigator site files, allowing the conduct of the trial and quality of the data produced to be evaluated. The documentation will be retained for a minimum of 5 years after the conclusion of the trial. The postal questionnaires and hospital CRFs will be stored for a minimum of 5 years after the conclusion of the trial as paper records, and for a minimum of 20 years in electronic format.

Adverse event management

All AEs and SAEs were recorded by the site principal investigator or delegated clinician and returned to the trial office on a CRF (see *Report Supplementary Materials 17* and *18*). In accordance with good clinical practice, SAEs were reported within 24 hours and AEs were reported within 5 days of the investigator becoming aware.

Once this information was received, the chief investigator determined causality and expectedness. The Research Ethics Committee was notified of SAEs that were unexpected and related to the trial within 15 days for a non-life-threatening event and within 7 days for a life-threatening event. For non-serious AEs, the central office was notified within 5 days of the event being known. All AEs and SAEs were reported to the DMEC, TSC and TMG. Expected AEs for this shoulder condition included infection; bleeding; delayed wound healing; conversion of a planned day-case procedure to an overnight stay for control of pain; post-procedural worsening of shoulder pain; injury to adjacent structures such as nerve, tendon, bone or joint; recurrent stiffness requiring further treatment; transient hyperglycaemia, steroid flare or joint sepsis following corticosteroid injection; and injuries related to the heating or cooling of tissues. The chief investigator reviewed follow-up reports 1 month later (see *Report Supplementary Material 19*) to ensure that adequate action had been taken and progress had been made.

Ethics approval and monitoring

Ethics committee approval and any changes to the project protocol

National Research Ethics Service Committee North East – Newcastle and North Tyneside 2 approved the study on 18 November 2014 (Research Ethics Committee reference 14/NE/1176). Health Research Authority approval for the study with an existing UK-wide review was granted on 15 June 2016. A summary of the changes made to the protocol since the original Research Ethics Committee approval is in *Appendix 2*.

Trial Management Group

The day-to-day management of the trial was overseen by the TMG, which met quarterly. A representative of the sponsor attended when available. These meetings monitored progress with recruitment (e.g. enrolment, consent, eligibility), allocation to study groups, adherence of the trial interventions to the protocol, retention of trial participants, monitoring of AEs/SAEs and reasons for participant withdrawal. The review of progress was undertaken at a participating site level and, as necessary, feedback was given to the principal investigator and research nurses at each site.

Trial Steering Committee

A TSC was appointed by the funding body to provide overall supervision of the trial and to advise on its continuation. The membership of the TSC is listed in the *Acknowledgements*.

Data Monitoring and Ethics Committee

The DMEC was appointed by the funding body and had access to the unblinded comparative data as provided by a statistician at York Trials Unit who was independent of the trial team. The DMEC monitored the data and made any recommendations about (dis)continuation of the trial to the independent TSC. The membership of the DMEC is listed in the *Acknowledgements*.

Patient and public involvement

Two patients who had previously received treatment for frozen shoulder at the lead site (James Cook University Hospital) and the independent patient representative member of the TSC were invited to comment on the patient information leaflet, the patient-facing data collection forms and the consent process for trial participation. The need to develop a leaflet to provide general information about frozen shoulder was identified following a qualitative study of patients with frozen shoulder using semistructured interviews.¹⁸ The two patient representatives were invited to attend the TMG during the early stages of the study and it was later agreed to seek their opinion outside the meetings when necessary. Recruitment was steady and the target was met on time. The retention of participants also went well and the target was exceeded. Therefore, there was little further contact with the two patient representatives during the trial, although they did advise on the newsletters to be sent to trial participants.

Following the initial analyses of the study results, we sought the advice of the two patient representatives and a wider group of seven patients with frozen shoulder at the lead site. Study results, associated risks for individual trial treatments and their health economic impact were discussed. Members shared their thoughts on their preferred choice of treatment based on the study results and agreed to support the trial team with dissemination to various platforms. This included contributing to the *Plain English summary* of this report, journal publications and web-based outputs, such as updating the entry about management of frozen shoulder on Wikipedia and helping to develop content for other appropriate web pages. These patients will also meet with local (shoulder research users group) and national shoulder patient groups (British Elbow and Shoulder Society patient liaison group) to ensure that the current evidence base for treatment options is available and disseminated appropriately to patients and the wider public.

Chapter 3 Trial results

This chapter begins with a summary of the findings of the internal pilot study and the nested shoulder capsular tissue and blood samples study. It then summarises recruitment, the flow of participants through the trial, the characteristics of participants at baseline and the results of analyses of the primary and secondary outcomes, as well as the integration of the findings into the existing literature.

Summary findings of the internal pilot

The objective of phase 1 of the internal pilot (months 4–9) was to have a minimum of four sites recruiting during the 6 months that had recruited 24 patients (i.e. evidence that they could recruit the expected one participant per month). To ensure that adequate progress had been made with setting up sites to recruit, 12 sites were to be set up (i.e. 50% of the total sites). At the end of month 9, we had recruited 20 patients (i.e. 83% of the target). This was in spite of not starting recruitment until month 7 and having only three of the four pilot sites set up. There were, however, two sites at which we were waiting on approval, and 16 out of 26 sites with which we had held a preliminary meeting.

We also reviewed other aspects of the study. In summary, of the 34 patients who had a clinically confirmed frozen shoulder, four met the exclusion criteria, 10 did not consent and 20 consented. Early data showed that treatment preference was the reason for all non-consent, rather than patients being too busy or not wanting to be involved in research. The time taken to consent ranged from 15 minutes to 1 hour. Participating sites had confirmed that they had received sufficient training and supporting documentation. Except at one of the four pilot sites, all surgeons were supportive of the study. At this one site, one surgeon was taking part, another surgeon felt that he did not see a sufficient number of patients to take part, and a third surgeon lacked equipoise to consent patients. All three surgeons, however, had agreed to deliver the surgical interventions to which patients were allocated. No patient non-compliance with treatment had yet been reported.

Although we had not met our patient recruitment or site set-up targets, both oversight committees were satisfied with the overall progress made during phase 1.

The primary objective of phase 2 was to review the feasibility of the ESP intervention and whether non-compliance in the ESP arm did not exceed 20–30%. At the end of this phase (month 27), of the 65 trial participants who had been allocated to ESP, 37 had ended their treatment and could be assessed for non-compliance. The remaining 28 participants either had started their treatment or were waiting to start treatment. Of the 37 participants who had ended their treatment, 29 (78%) met our criteria for completing the intervention as had been agreed with the trial team and independent committees: the participant had attended eight sessions or more ($n = 19$); or had attended fewer than eight sessions but both the participant and the physiotherapist were satisfied with their progress ($n = 9$); or had attended fewer than eight sessions and declined to attend more because they were satisfied with their progress, their ability to manage independently, or both ($n = 1$). Therefore, non-compliance with the ESP intervention applied to 22% of participants, which was within the threshold of 20–30%. The oversight committees agreed that this was an acceptable degree of non-compliance and that the trial should continue with all three treatment arms.

Another aspect of trial feasibility that was reviewed at month 27 was non-consent into the trial. This was because there was concern that patients would often have already received physiotherapy in primary care and that this could affect their decision to take part in the trial, given that the ESP intervention was one of the treatment options. It was found that 55% ($n = 72$) of the 131 reasons

for patients not taking part was because they either 'want surgery' or 'do not want physiotherapy'. This compared with 18% ($n = 24$) of patients who 'want physiotherapy' or 'do not want surgery'. Other reasons for non-consent were infrequent. The main treatment that non-consenting patients went on to have was keyhole surgery (45%, $n = 67$). Of patients randomised into the trial, the majority had no treatment preference (53%, $n = 159$), over one-third preferred surgery (39%, $n = 116$) and the rest preferred physiotherapy (8%, $n = 25$). Despite the preferences for surgery, this did not have an impact on the feasibility of the trial, with 36 sites set up, compared with the target of 25 sites, and 325 participants recruited against the target of 250 participants. We also reviewed the timing of the delivery of interventions, which confirmed that only one site had regularly failed to deliver surgery on time because of local pressures. The local trial team and principal investigator were very engaged and responsive to the trial team's concerns and prioritised the trial participants for the surgical procedures.

In short, UK FROST was being delivered on time and to target, with an acceptable degree of non-compliance with the ESP intervention. The oversight committees were satisfied with the progress of all aspects of the trial and for it to continue as planned.

Summary findings of the shoulder capsule tissue and blood samples study

The primary aim of this nested study was to determine the key molecular processes and changes seen in the shoulder capsular tissue of patients with frozen shoulder in order to better understand these processes; and to determine the relationship between tissue changes, serum biomarkers and clinical symptoms and signs at presentation. This was done by determining the molecular and cellular abnormalities in shoulder capsular tissue obtained during surgery, by determining serum protein and cytokine signatures, and by correlating any tissue and serum abnormalities detected with the clinical presentation.

Following research ethics approval from the Oxford Musculoskeletal Biobank (09/H0606/11) and National Research Ethics Service Committee Newcastle and North Tyneside (14/NE/1176), appropriate informed consent was sought from UK FROST participants randomised to receive the ACR intervention. For a small sample of 16 patients who consented to the study, the shoulder capsular tissue and a venous blood sample were collected. The findings from the analysis of the capsular tissue samples were then compared with data available in the Oxford Tissue Biobank of findings from both healthy and diseased rotator cuff tendon tissues.

Inflammation signatures differed between tissue from frozen shoulder and that from tendon tears. Compared with tendon tear tissue, frozen shoulder capsular tissue showed reduced expression of nuclear factor- κ B response genes, including *TNFA*, *IL6* and *IL8*, and increased expression of *IL10*, *CD14*, *CD163* and *C1QA* messenger RNA. The fibroblast activation markers podoplanin (*PDPN*), *CD106* (*VCAM1*) and *CD248* and the fibroblast activation protein were highly expressed in adhesive capsulitis and torn tendons, compared with healthy tendons. However, fibroblast activation marker *CD90* was significantly reduced in adhesive capsulitis compared with healthy and diseased tendon tissue. Proresolving receptors mediating resolution of inflammation, including *ALX/FPR2*, *CMKLR1* and *GPR32*, were highly expressed in frozen shoulder capsular tissue.³⁷

This study in patients of a similar age has provided some insight into why inflammation ultimately resolves in frozen shoulder but persists in tendon tears. This study suggests that the phenotypes of fibroblast subsets populating diseased shoulder tissues differ between conditions with self-limiting and those with persistent inflammation. *CD90* therefore represents an important pathogenic marker and possible molecular checkpoint regulating persistent stroma-mediated inflammation in common soft tissue diseases of the shoulder. Proresolving proteins were highly expressed in frozen shoulder tissue compared with established shoulder tendon tears. These findings have provided a novel insight into the disease mechanism of frozen shoulder, which points towards a resolving inflammatory environment.

Further studies to better understand the biological mechanisms governing successful resolution of inflammation should inform new therapeutic strategies to accelerate disease resolution in frozen shoulder.

Recruitment into UK FROST

A total of 37 sites screened patients for the UK FROST trial, of which 35 randomised at least one patient. *Appendix 3* presents the number of patients screened and randomised at each site, as well as the number of participants who withdrew before the end of the study.

Flow of participants

The flow of participants from screening to randomisation, treatment, follow-up and analysis is illustrated in the CONSORT flow diagram (*Figure 2*). Of 914 screened patients, 503 were randomised into the UK FROST trial. The reasons for exclusion were not meeting eligibility criteria ($n = 95$), non-consent ($n = 295$) and other reasons ($n = 21$). The most frequent reason for exclusion was having frozen shoulder symptoms secondary to trauma that required hospital care. Most patients who provided information about why they were not willing to join the trial said that this was because they had already had physiotherapy and wanted to have surgery (*Table 1*).

Treatment allocations were 2 : 2 : 1 to MUA with steroid injection ($n = 201$), ACR with MUA ($n = 203$), and ESP with steroid injection ($n = 99$). Follow-up rates at 3, 6 and 12 months post randomisation were between 85% and 89%, above the target of 80% assumed in the sample size and with no evidence of differential dropout in any of the treatment arms. The primary analysis at 12-month follow-up included all participants with OSS outcome data at one or more follow-ups, and therefore 94% of participants could be included in the analysis.

Baseline characteristics

Eligible patients who did and eligible patients who did not consent to participate in the trial were comparable in their baseline characteristics (*Table 2*). The demographic and clinical characteristics of participants at baseline are presented in *Table 3*, comparing the profile of the total patients randomised ($n = 503$) with that of participants included in the primary analysis ($n = 473$). No systematic differences between the two populations were evident. The characteristics of patients in the three randomised arms were broadly similar, with the exception of a greater number of participants in the MUA arm being currently in paid work and some group imbalance in having had a similar shoulder problem on the opposite side to the reference shoulder.

Intervention delivery

The criteria for having completed each of the three trial interventions were agreed and documented in the statistical analysis plan. For MUA and ACR, this was the completed surgical procedure, regardless of the completion of any PPP. In the ESP arm, completion of the intervention was defined as receipt of a minimum of eight physiotherapy sessions, unless the patient was discharged as satisfied with their progress earlier.

Table 4 shows that 82% of patients completed MUA, 80% of patients completed ACR and 81% of patients completed ESP. Overall, 16 participants (3%) crossed over to a different trial treatment, and 17 (3%) received an alternative non-trial treatment. Participants who did not start or complete any trial or non-trial treatment were classed as 'no treatment recorded' ($n = 64$, 13%).

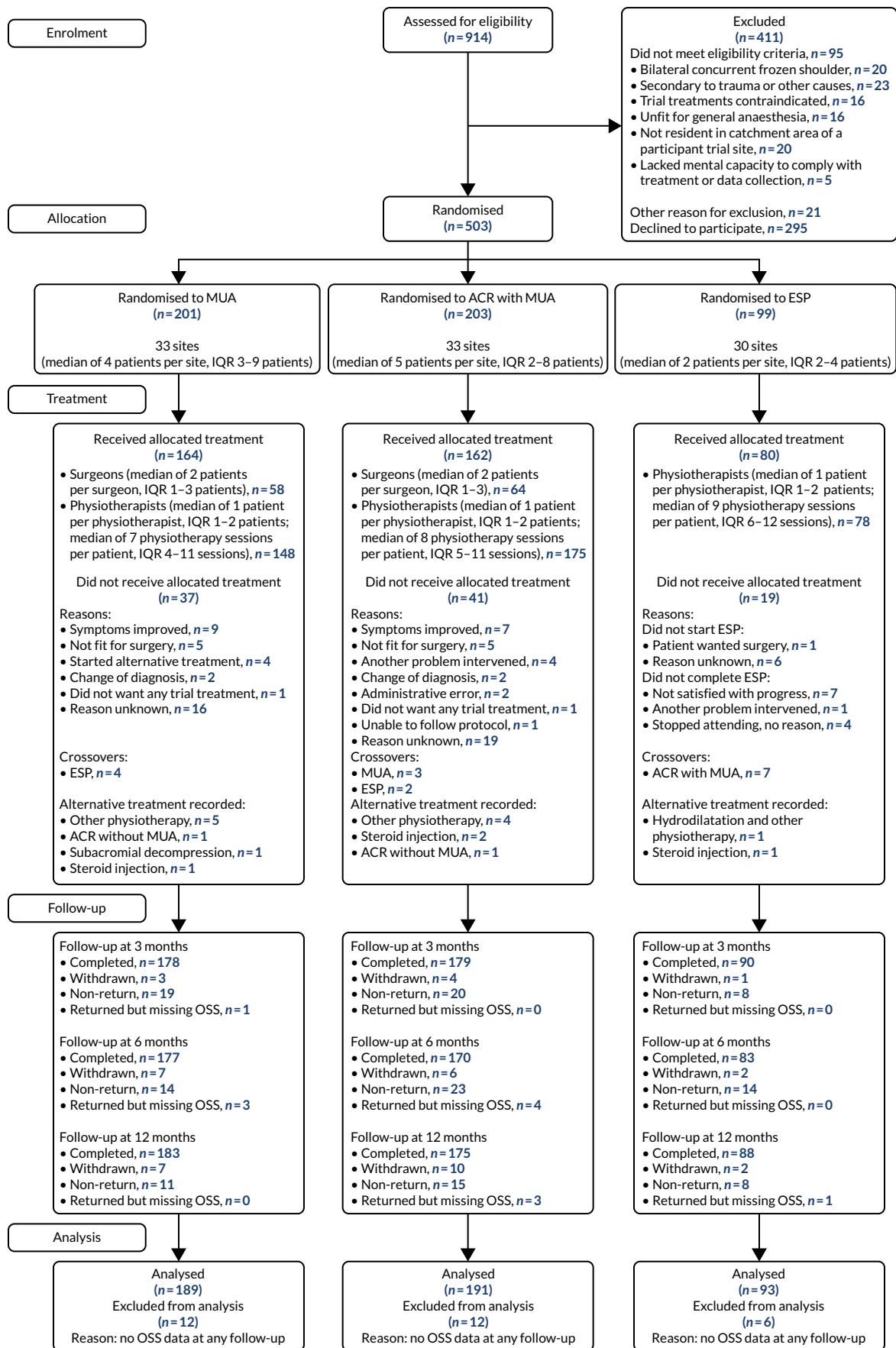


FIGURE 2 The CONSORT flow diagram. IQR, interquartile range.

TABLE 1 Reasons for exclusion

Reason	Number excluded	Per cent of total excluded (N = 411)
<i>Trial exclusion criteria (n = 116; more than one reason possible)</i>		
Bilateral concurrent frozen shoulder	20	4.9
Frozen shoulder secondary to trauma (i.e. trauma to the shoulder that required hospital care, e.g. fracture, dislocation, rotator cuff tear)	23	5.6
Frozen shoulder secondary to other causes (e.g. recent breast surgery or radiotherapy)	16	3.9
Any of the trial treatments are contraindicated (e.g. patient is unfit for anaesthesia or corticosteroid)	16	3.9
Not resident in a catchment area of a participating site	20	4.9
Lacks mental capacity and unable to understand the trial or instructions for treatment	5	1.2
Other reason	21	5.1
<i>Patient non-consent (n = 295; grouped free-text information from screening form)</i>		
Wanted surgery	79	19.2
Did not want surgery	40	9.7
Wanted physiotherapy	22	5.4
Did not want physiotherapy	48	11.7
Wanted steroid injection	2	0.5
Wanted clinician to decide	3	0.7
Wanted no further treatment	2	0.5
Could not travel to trial site	1	0.2
Was too busy to take part	7	1.7
Too many questionnaires	1	0.2
Did not want to take part	29	7.1
Unclear/no reason given	61	14.8
<i>Patient non-consent (n = 295; selection of possible reasons from list of preference form if agreed to complete, more than one reason possible)</i>		
I wanted the treating clinician to make a decision for me	13	3.2
I have already had physiotherapy	84	20.4
I do not want physiotherapy	38	9.2
I do not want surgery	36	8.8
I do want physiotherapy	28	6.8
I do want surgery	75	18.2
I am too busy to take part in research	10	2.4
I do not want to be involved in research	6	1.5
I thought there were too many questionnaires to complete	1	0.2
I just did not want to take part	10	2.4
Other	29	7.1
Did not agree to complete preference form	109	26.5

TABLE 2 Baseline characteristics of different populations

Characteristic	Eligible but non-consenting (N = 295)	Eligible and randomised (N = 503)
Sex, n (%)		
Female	200 (68)	319 (63)
Age (years)		
<i>n</i>	293	503
Mean (SD)	53.7 (8.0)	54.3 (7.7)
Median (minimum, maximum)	53 (32, 82)	54 (30, 77)
Diabetic, n (%)		
No	219 (74)	353 (70)
Type 1	23 (8)	29 (6)
Type 2	51 (17)	121 (24)
Missing	2 (1)	0 (0)
Affected shoulder, n (%)		
Left	181 (61)	304 (60)
Right	110 (37)	196 (39)
Missing	4 (1)	3 (1)
Duration of symptoms (months)		
<i>n</i>	288	495
Mean (SD)	10.5 (7.0)	10.9 (9.2)
Median (IQR)	9 (6–12)	8 (6–12)
Minimum, maximum	1, 48	0, 96
Duration of symptoms (grouped), n (%)		
< 9 months	135 (46)	249 (50)
≥ 9 months	153 (52)	246 (49)
Missing	7 (2)	8 (2)

IQR, interquartile range.

TABLE 3 Baseline characteristics of randomised participants

Characteristic	As randomised (N = 503)			As analysed (N = 473)		
	MUA	ACR	ESP	MUA	ACR	ESP
Sex, n (%)						
Female	129 (64)	126 (62)	64 (65)	121 (64)	117 (61)	62 (67)
Age (years)						
<i>n</i>	201	203	99	189	191	93
Mean (SD)	54.5 (7.7)	53.9 (7.7)	54.5 (7.8)	54.4 (7.3)	54.4 (7.6)	54.8 (7.8)
Median (minimum, maximum)	54 (30, 75)	54 (33, 76)	53 (39, 77)	54 (30, 75)	55 (33, 76)	53 (39, 77)
Diabetic, n (%)						
No	141 (70)	143 (70)	69 (70)	131 (69)	135 (71)	66 (71)
Type 1	12 (6)	12 (6)	5 (5)	12 (6)	11 (6)	5 (5)
Type 2	48 (24)	48 (24)	25 (25)	46 (24)	45 (24)	22 (24)

TABLE 3 Baseline characteristics of randomised participants (continued)

Characteristic	As randomised (N = 503)			As analysed (N = 473)		
	MUA	ACR	ESP	MUA	ACR	ESP
Affected shoulder, n (%)						
Left	127 (63)	121 (60)	56 (57)	119 (63)	114 (60)	54 (58)
Right	73 (36)	80 (39)	43 (43)	69 (37)	75 (39)	39 (42)
Missing	1 (1)	2 (1)	0 (0)	1 (1)	2 (1)	0 (0)
Duration of symptoms (months)						
n	196	201	98	185	190	92
Mean (SD)	10.5 (8.6)	11.3 (10.0)	10.8 (8.8)	10.7 (8.7)	11.3 (10.1)	11.0 (9.0)
Median (IQR)	8 (6–12)	9 (6–12)	8 (6–12)	8 (6–12)	9 (6–12)	8 (6–12)
Minimum, maximum	2, 60	0, 96	2, 72	2, 60	2, 96	2, 72
Duration of symptoms (grouped), n (%)						
< 9 months	103 (51)	95 (47)	51 (52)	96 (51)	90 (47)	48 (52)
≥ 9 months	93 (46)	106 (52)	47 (47)	89 (47)	100 (52)	44 (47)
Missing	5 (2)	2 (1)	1 (1)	4 (2)	1 (1)	1 (1)
Radiographs, n (%)						
Anteroposterior view	200 (100)	201 (99)	99 (100)	188 (99)	190 (99)	93 (100)
Axillary view	174 (87)	179 (88)	86 (87)	163 (86)	169 (88)	80 (86)
Modified axillary	29 (14)	24 (12)	14 (14)	27 (14)	24 (13)	14 (15)
Ethnicity summary, n (%)						
White British	187 (93)	185 (91)	84 (85)	176 (93%)	175 (92)	80 (86)
Other	13 (6)	17 (8)	15 (15)	12 (6)	15 (8)	13 (14)
Missing	1 (0.5)	1 (0.5)	0 (0)	1 (0.5)	1 (0.5)	0 (0)
Education, n (%)						
Left school before 16 years old	33 (16)	28 (14)	15 (15)	31 (16)	26 (14)	14 (15)
Left school at 16 years old	75 (37)	74 (37)	37 (37)	70 (37)	71 (37)	34 (37)
Left education at 18 years old	27 (13)	28 (14)	14 (14)	25 (13)	26 (14)	12 (13)
Degree-level education	28 (14)	36 (18)	18 (18)	27 (14)	33 (18)	18 (19)
Other vocational/work-related qualifications	23 (11)	19 (9)	6 (6)	22 (12)	18 (9)	6 (6)
Other	11 (5)	16 (8)	9 (9)	11 (6)	15 (8)	9 (10)
Missing	4 (2)	2 (1)	0 (0)	3 (2)	2 (1)	0 (0)
Employment status summary, n (%)						
In paid work	129 (64)	118 (58)	53 (54)	124 (66)	111 (58)	50 (54)
Not in paid work	69 (34)	82 (40)	46 (46)	62 (33)	78 (41)	43 (46)
Missing	3 (1)	3 (1)	0 (0)	3 (2)	2 (1)	0 (0)

continued

TABLE 3 Baseline characteristics of randomised participants (continued)

Characteristic	As randomised (N = 503)			As analysed (N = 473)		
	MUA	ACR	ESP	MUA	ACR	ESP
Type of employment, n (%)						
Unskilled manual	17 (8)	15 (7)	8 (8)	16 (8)	13 (7)	7 (8)
Skilled manual	21 (10)	18 (9)	18 (18)	19 (10)	16 (8)	17 (18)
Unskilled non-manual	19 (9)	17 (8)	4 (4)	19 (10)	17 (9)	4 (4)
Skilled non-manual	41 (20)	37 (18)	13 (13)	40 (21)	37 (19)	12 (13)
Professional	13 (6)	19 (9)	10 (10)	13 (7)	18 (9)	10 (11)
Other	20 (10)	17 (8)	10 (10)	18 (10)	15 (8)	10 (11)
Currently taking steroids for affected shoulder, n (%)						
Yes	2 (1)	7 (3)	0 (0)	2 (1)	7 (4)	0 (0)
No	196 (98)	195 (96)	99 (100)	184 (97)	183 (96)	93 (100)
Missing	3 (1)	1 (<0.5)	0 (0)	3 (2)	1 (1)	0 (0)
Had steroid injection for affected shoulder, n (%)						
Yes	97 (48)	117 (58)	55 (56)	93 (49)	112 (59)	53 (57)
No	102 (51)	86 (42)	44 (44)	94 (50)	79 (41)	40 (43)
Missing	2 (1)	0 (0)	0 (0)	2 (1)	0 (0)	0 (0)
If yes						
Number of injections, median (IQR)	1 (1-2)	1 (1-2)	1 (1-2)	1 (1-2)	1 (1-2)	1 (1-2)
Weeks since last injection, median (IQR)	12 (8-24)	12 (6-20)	10 (6-20)	12 (8-24)	12 (6-20)	10 (6-20)
Delivered by GP, n (%)	59 (29)	74 (36)	36 (36)	56 (30)	72 (38)	35 (38)
Delivered by physiotherapist, n (%)	26 (13)	27 (13)	11 (11)	26 (14)	25 (13)	11 (12)
Other delivery, n (%)	5 (2)	6 (3)	4 (4)	4 (2)	5 (3)	4 (4)
Previous physiotherapy for affected shoulder, n (%)						
Yes	125 (62)	124 (61)	59 (60)	117 (62)	117 (61)	58 (62)
No	76 (38)	77 (38)	39 (39)	72 (38)	73 (38)	35 (38)
Missing	0 (0)	2 (1)	1 (1)	0 (0)	1 (1)	0 (0)
If yes						
General practice	31 (15)	25 (12)	13 (13)	28 (15)	23 (12)	13 (14)
Hospital	60 (30)	58 (29)	30 (30)	56 (30)	54 (28)	29 (31)
Home	6 (3)	5 (2)	2 (2)	6 (3)	5 (3)	2 (2)
Other	22 (11)	35 (17)	15 (15)	21 (11)	33 (17)	14 (15)
Number of physiotherapy sessions, median (IQR)	5 (3-8)	5 (3-6)	4 (2-6)	5 (3-8)	5 (3-6)	4 (2.5-6)
Number of weeks had physiotherapy, median (IQR)	6 (4-12)	6 (4-12)	7.5 (5-10)	6 (4-12)	6 (4-12)	7.5 (5-10)

TABLE 3 Baseline characteristics of randomised participants (continued)

Characteristic	As randomised (N = 503)			As analysed (N = 473)		
	MUA	ACR	ESP	MUA	ACR	ESP
Dominant arm affected, n (%)						
Yes	81 (40)	82 (40)	39 (39)	77 (41)	76 (40)	36 (39)
No	115 (57)	120 (59)	59 (60)	107 (57)	114 (60)	56 (60)
Ambidextrous	0 (0)	1 (<0.5)	0 (0)	0 (0)	1 (1)	0 (0)
Missing	5 (2)	0 (0)	1 (1)	5 (3)	0 (0)	1 (1)
Number of weeks had shoulder problem, median (IQR)	32 (24–52)	35 (24–52)	32 (24–48)	34 (24–52)	35.5 (24–52)	32 (24–48)
Similar shoulder problem on the same side, n (%)						
Yes	19 (9)	26 (13)	12 (12)	17 (9)	24 (13)	12 (13)
No	178 (8)	177 (87)	87 (88)	168 (89)	167 (87)	81 (87)
Missing	4 (2)	0 (0)	0 (0)	4 (2)	0 (0)	0 (0)
Similar shoulder problem on the opposite side, n (%)						
Yes	62 (31)	53 (26)	13 (13)	59 (31)	51 (27)	12 (13)
No	132 (66)	146 (72)	85 (86)	124 (66)	136 (71)	80 (86)
Missing	7 (3)	4 (2)	1 (1)	6 (3)	4 (2)	1 (1)
OSS (0–48)						
n	200	202	99	188	190	93
Mean (SD)	20.5 (8.9)	19.1 (7.7)	20.3 (8.0)	20.6 (8.9)	19.2 (7.5)	20.3 (8.1)
Median	20	19	20	20	19	20
Minimum, maximum	2, 48	1, 37	2, 42	2, 48	1, 37	2, 42
QuickDASH (0–100)						
n	192	197	96	181	187	90
Mean (SD)	57.0 (21.0)	61.7 (18.5)	59.4 (19.7)	56.8 (21.1)	61.3 (18.5)	59.1 (20.0)
Median	59	64	60	59	64	59.5
Minimum, maximum	0, 100	14, 100	14, 98	0, 100	14, 100	14, 98
Numeric Rating Scale for Pain (0–10)						
n	199	201	99	187	190	93
Mean (SD)	6.8 (2.2)	7.0 (1.9)	6.9 (2.4)	6.7 (2.3)	7.0 (1.9)	6.8 (2.4)
Median	7	7	7	7	7	7
Minimum, maximum	0, 10	0, 10	0, 10	0, 10	0, 10	0, 10
Symptom severity (0–100)						
n	198	201	99	186	189	93
Mean (SD)	83.8 (21.8)	86.2 (20.1)	89.2 (15.4)	83.9 (22.1)	86.0 (20.4)	89.0 (15.5)
Median	90	95	100	90	95	100
Minimum, maximum	0, 100	0, 100	50, 100	0, 100	0, 100	50, 100
IQR, interquartile range.						

TABLE 4 Completed treatment

Treatment received	Randomised treatment, n (%)		
	MUA (N = 201)	ACR (N = 203)	ESP (N = 99)
Trial treatment			
MUA ^a	164 (82)	3 (1)	-
ACR ^b	-	162 (80)	7 (7)
ESP ^c	4 (2)	2 (1)	80 (81)
Alternative treatment ^d			
Other physiotherapy	5 (2)	4 (2)	-
ACR without MUA	1 (0.5)	1 (0.5)	-
Steroid injection	1 (0.5)	2 (1)	1 (1)
Subacromial decompression	1 (0.5)	-	-
Hydrodilataion and other physiotherapy	-	-	1 (1)
No treatment recorded ^e	25 (12)	29 (14)	10 (10)

a Patient had trial MUA (regardless of release status, receipt of steroid injection or PPP).
b Patient had trial ACR and MUA (regardless of release status or receipt of PPP).
c Patient completed eight or more ESP sessions, or fewer if patient and/or physiotherapist were satisfied with progress (regardless of receipt of steroid injection).
d Patient did not receive any trial treatment as defined under a, b and c, but alternative treatment was recorded.
e No trial or alternative treatment as defined under a, b, c and d was recorded for patient.

The profile of surgeons and physiotherapists treating patients who completed their randomised intervention is presented in *Appendix 4*. Based on the available data, operating surgeons were predominantly consultants who had experience of routinely performing the trial operations up to once per month. Physiotherapists delivering ESP or PPP were most frequently band 6, treating between two and three frozen shoulders per month.

Waiting times to the start of each randomised intervention varied considerably. *Table 5* shows that ESP patients received their first physiotherapy session or steroid injection within a median of 14 days, whereas patients waited a median of 56.5 days for MUA and a median of 71.5 days for ACR. Seventy patients in the ESP arm received a steroid injection on average within 12.8 days of randomisation. Nearly half (46%, $n = 32$) of these injections were administered on the day of randomisation.

Following completion of their randomised treatment, a number of patients received further treatment, as detailed in *Table 6*. Most commonly, this was ACR for patients randomised to MUA and further physiotherapy for patients randomised to ESP. Patients in the ACR arm received fewest further treatments.

As part of the surgical treatments, optimal release was reported as achieved in 92% of MUA procedures and in 98% of ACR procedures (*Table 7*). Steroid injection was delivered for all completed MUAs and 28% of ACRs. Steroid injection was also given to 80% of patients randomised to ESP (*Table 8*).

The number of delivered physiotherapy sessions is presented in *Table 9*. Participants who completed the ESP intervention attended a median of 9 sessions, whereas PPP following surgical procedures attended slightly fewer sessions (median of 7 for MUA and 8 for ACR). Individual therapeutic elements delivered as part of ESP and PPP sessions are summarised in *Appendix 5*.

TABLE 5 Time to start/end of treatment

Day started	Treatment arm		
	MUA	ACR	ESP
<i>Days from the date of randomisation to ...</i>	<i>... the day of operation</i>	<i>... the day of operation</i>	<i>... the first day of physiotherapy/injection</i>
<i>n</i>	164	162	80
Mean (SD)	63 (39.3)	82 (52.2)	20 (21.2)
Median	56.5	71.5	14
Minimum, maximum	4, 244	1, 249	0, 140
<i>Days from the date of randomisation to ...</i>	<i>... the first day of PPP</i>	<i>... the first day of PPP</i>	
<i>n</i>	158	156	–
Mean (SD)	64 (39.6)	83 (52.0)	–
Median	57.5	71.5	–
Minimum, maximum	4, 245	1, 249	–
<i>Days from the date of surgery to ...</i>	<i>... the first day of PPP</i>	<i>... the first day of PPP</i>	
<i>n</i>	158	156	–
Mean (SD)	3 (7.0)	3 (8.6)	–
Median	1	1	–
Minimum, maximum	–6, 40	–40, 76	–
<i>Days from the first day of physiotherapy to ...</i>	<i>... the last day of PPP</i>	<i>... the last day of PPP</i>	<i>... the last day of ESP</i>
<i>n</i>	158	156	80
Mean (SD)	86 (53.1)	91 (46.0)	100 (46.5)
Median	78	85.5	92
Minimum, maximum	1, 243	1, 285	15, 246

TABLE 6 Further treatment (any treatment following completion of trial treatment)

Further treatment	Subgroup: randomised and completed treatment		
	MUA	ACR	ESP
Further surgical treatment			
ACR	4	0	3
ACR without MUA	3	0	0
ACR plus injection to opposite shoulder	0	0	1
Arthroscopic arthrolysis and decompression	0	0	1
MUA	1	1	3
Further non-surgical treatment			
Steroid injection	3	3	3
Glenohumeral joint injection	2	0	0
Ultrasound guided injection	0	1	1
Other/further physiotherapy	2	3	6
Rheumatology clinic	0	0	1
Total number of further treatments	15	8	19
Total number (%) of patients having one or more further treatments	14 (7)	8 (4)	15 (15)

TABLE 7 Fidelity (surgery and injection)

Fidelity	MUA			ACR			ESP	
	<i>n</i>	% of patients randomised to MUA (N = 201)	% of patients randomised to and completed MUA (N = 164)	<i>n</i>	% of patients randomised to ACR (N = 203)	% of patients to and completed ACR (N = 162)	% of patients randomised to ESP (N = 99)	% of patients randomised to and completed ESP (N = 80)
Surgery delivered	164	82	100	162	80	100	-	-
Optimal release achieved	151	75	92	158	78	98	-	-
Steroid injection received	164	82	100	45	22	28	80	86

TABLE 8 Fidelity (physiotherapy)

Number of physiotherapy sessions	MUA		ACR		ESP	
	Randomised to and completed	Randomised	Randomised to and completed	Randomised	Randomised to and completed	Randomised
<i>n</i>	164	201	162	203	80	99
Mean (SD)	7.7 (4.39)	6.3 (4.93)	8.1 (4.00)	6.5 (4.78)	8.7 (3.26)	7.6 (3.95)
Median	7	6	8	6	9	8
Minimum, maximum	0, 18	0, 18	0, 18	0, 18	2, 15	0, 15

TABLE 9 Unadjusted OSS by treatment arm: follow-up since randomisation

Time point	Treatment arm			
	MUA	ACR	ESP	Total
Baseline				
<i>N</i>	200	202	99	501
Mean (SD)	20.5 (8.88)	19.1 (7.72)	20.3 (7.97)	19.9 (8.26)
Median	20	19	20	20
Minimum, maximum	2, 48	1, 37	2, 42	1, 48
<i>n</i> (%) maximum score (48)	1 (1)	0 (0)	0 (0)	1 (< 0.5)
3 months				
<i>N</i>	178	179	90	447
Mean (SD)	31.7 (10.41)	27.4 (11.12)	32.7 (10.95)	30.2 (11.03)
Median	34	28	35	32
Minimum, maximum	5, 48	2, 48	4, 48	2, 48
<i>n</i> (%) maximum score (48)	2 (1)	1 (1)	2 (2)	5 (1)

TABLE 9 Unadjusted OSS by treatment arm: follow-up since randomisation (continued)

Time point	Treatment arm			Total
	MUA	ACR	ESP	
6 months				
N	177	170	83	430
Mean (SD)	38.6 (9.70)	36.5 (9.96)	36.5 (11.08)	37.3 (10.11)
Median	41	39	40	40
Minimum, maximum	3, 48	7, 48	6, 48	3, 48
n (%) maximum score (48)	23 (13)	11 (6)	10 (12)	44 (10)
12 months				
N	183	175	88	446
Mean (SD)	39.4 (9.87)	40.7 (9.99)	38.9 (10.49)	39.8 (10.05)
Median	43	45	42.5	43
Minimum, maximum	4, 48	2, 48	4, 48	2, 48
n (%) maximum score (48)	44 (24)	45 (26)	17 (19)	106 (24)

Primary outcome

Descriptives

The OSS was the trial primary outcome and was collected using questionnaires at baseline and at 3, 6 and 12 months post randomisation. When OSS data from either the 6-month post-randomisation or the 6-month post-treatment questionnaire were available, and the two questionnaires had been sent to patients within 28 days, the available responses were used to complete any missing OSS outcomes. A summary of descriptive statistics of OSS scores is presented in Table 9 and Figure 3 (see Appendix 6 for the scores split by pain and function subdomains). By 12-month follow-up, many participants (24%) had regained function up to the top OSS score of 48, and a ceiling effect of OSS scores for all three arms could be observed. This restricted variability of scores at the top end meant that the ability of the trial to detect clinically meaningful differences at the primary end point was reduced.

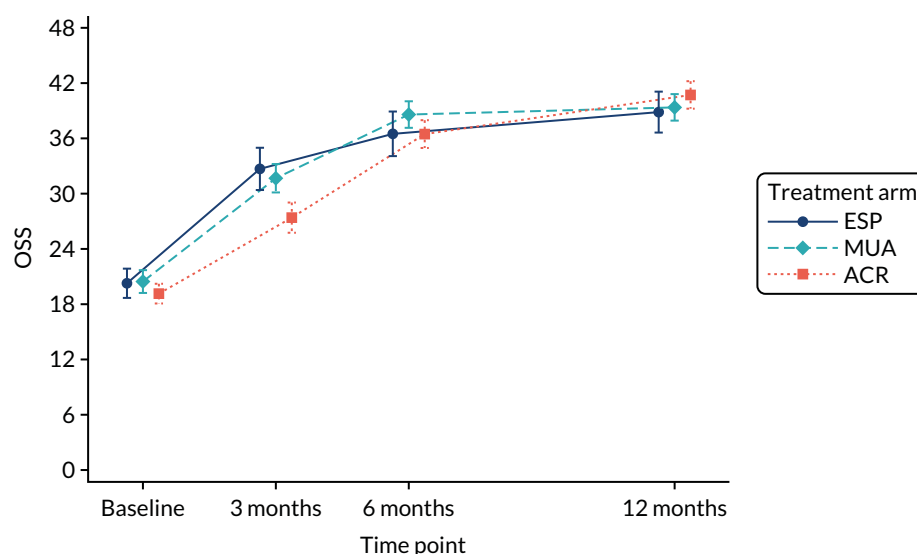


FIGURE 3 Unadjusted mean OSS and 95% CIs by treatment arm.

Primary analysis

The ITT primary analysis was based on a linear mixed model incorporating all time points and using an unstructured covariance pattern to model the relationship of repeated measurements by the same individual. The model adjusted for age, sex, diabetes and OSS at baseline and incorporated a random effect for site. The results in *Table 10* present adjusted estimates of group means and mean differences for each treatment comparison. At the primary end point at 12 months, participants randomised to ACR were shown to have, on average, statistically significantly higher (better) OSS scores than participants randomised to MUA (2.01 points, 95% CI 0.10 to 3.91 points) and ESP (3.06 points, 95% CI 0.71 to 5.41 points). Although statistically significant, mean estimates were short of the sought minimal clinically important effect size of 4–5 OSS points (the trial was powered for differences of 4 points for comparing MUA with ACR and of 5 points for comparisons with ESP).

For the short-term follow-up at 3 months post randomisation, ACR was shown to result in lower (worse) outcomes than the other two interventions. Mean differences for all treatment comparisons are in *Appendix 7*. There was no evidence of statistically significant differences in average OSS scores over the 12 months' follow-up. Differences of clinically important magnitude, as defined above, were included in the 95% CIs for the benefit of MUA and ESP compared with ACR at 3 months, and ACR compared with ESP at 12 months. Clinically meaningful group differences may therefore exist for these comparisons in the population.

TABLE 10 Estimated mean OSS differences by treatment arm (estimates from primary analysis model^a)

Time point	Treatment arm		Difference, mean (95% CI)	p-value
	Mean (95% CI)	Mean (95% CI)		
	MUA	ESP		
3 months	30.2 (28.8 to 31.6)	31.6 (29.7 to 33.5)	-1.36 (-3.70 to 0.98)	0.25
6 months	37.1 (35.7 to 38.4)	34.9 (33.0 to 36.8)	2.15 (-0.12 to 4.42)	0.06
12 months ^b	38.3 (36.9 to 39.7)	37.2 (35.3 to 39.2)	1.05 (-1.28 to 3.39)	0.38
Average	35.2 (34.0 to 36.4)	34.6 (33.0 to 36.2)	0.61 (-1.31 to 2.53)	0.53
	ACR	ESP		
3 months	26.9 (25.5 to 28.3)	31.6 (29.7 to 33.5)	-4.72 (-7.06 to -2.39)	< 0.01
6 months	35.9 (34.6 to 37.3)	34.9 (33.0 to 36.8)	0.98 (-1.31 to 3.26)	0.40
12 months ^b	40.3 (38.9 to 41.7)	37.2 (35.3 to 39.2)	3.06 (0.71 to 5.41)	0.01
Average	34.4 (33.2 to 35.5)	34.6 (33.0 to 36.2)	-0.23 (-2.15 to 1.70)	0.82
	ACR	MUA		
3 months	26.9 (25.5 to 28.3)	30.2 (28.8 to 31.6)	-3.36 (-5.27 to -1.45)	< 0.01
6 months	35.9 (34.6 to 37.3)	37.1 (35.7 to 38.4)	-1.17 (-3.02 to 0.67)	0.21
12 months ^b	40.3 (38.9 to 41.7)	38.3 (36.9 to 39.7)	2.01 (0.10 to 3.91)	0.04
Average	34.4 (33.2 to 35.5)	35.2 (34.0 to 36.4)	-0.84 (-2.41 to 0.72)	0.29

a Linear mixed covariance pattern model adjusted for age, sex, diabetes, OSS at baseline (fixed effects) and site (random effect).
b Primary end point for each treatment comparison.

Secondary analyses

Analysis incorporating different waiting times

In addition to questionnaires completed at post-randomisation follow-ups, participants were asked to complete the OSS just before and 6 months following receipt of treatment in order to account for the differential waiting times for each trial treatment. Descriptive results for these outcomes at these two points are given in *Table 11* and are presented together with OSS scores at baseline and at 12 months' post-randomisation follow-up in *Figure 4*. The OSS scores appeared to stay stable between baseline and the start of any of the treatments, which was later for the surgical arms (95% CI of mean difference ESP vs. MUA -2.0 to 2.7 , ESP vs. ACR -1.5 to 3.2 , MUA vs. ACR -1.4 to 2.5). Six months following treatment, scores had improved to a greater extent in the surgical arms than in the ESP arm (95% CI of mean difference ESP vs. MUA -5.5 to -2.2 , ESP vs. ACR -6.1 to -0.6 , MUA vs. ACR -2.6 to 1.6) and were similar to final follow-up scores by 8 months.

TABLE 11 Unadjusted OSS by treatment arm: pre treatment and 6 months post treatment

Time point	Treatment arm			Total
	MUA	ACR	ESP	
Pre treatment				
<i>n</i>	159	157	77	393
Mean (SD)	21.5 (8.79)	21.0 (8.92)	21.8 (8.02)	21.4 (8.68)
Median	21	21	22	21
Minimum, maximum	3, 46	1, 42	6, 42	1, 46
Post treatment (6 months)				
<i>n</i>	157	152	81	390
Mean (SD)	39.0 (9.03)	39.4 (9.68)	36.1 (10.67)	38.5 (9.70)
Median	42	43	39	42
Minimum, maximum	6, 48	2, 48	6, 48	2, 48

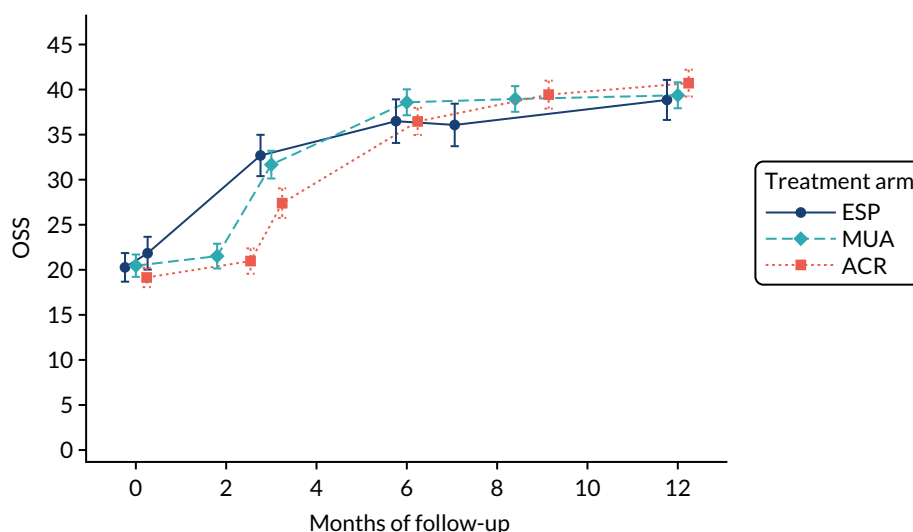


FIGURE 4 Unadjusted mean OSS and 95% CIs by treatment arm (using scores at baseline, follow-up before treatment and 6 months after treatment; and at 3, 6 and 12 months post randomisation).

A linear mixed random intercept model incorporated time as a continuous variable, included data from all available time points for each patient (up to five measurements) and adjusted for the same covariates as the primary analysis model. OSS estimates at 3, 6 and 12 months post-randomisation follow-up were derived from the model and are presented in *Table 12*. Compared with the primary analysis model, group differences tended to be of smaller magnitude, with the exception of the difference between ACR and ESP at 12 months (3.26 points in favour of ACR, 95% CI 1.18 to 5.35 points). The 95% CI interval still included the minimal clinically important difference for this comparison of 5 OSS points.

Analysis incorporating treatment compliance

Baseline characteristics for participants who did and participants who did not complete their randomised treatment according to the trial definitions are presented in *Table 13*. The profile of non-completers in each treatment arm tended to be different.

Owing to the three active treatments under investigation and multiple alternative treatment pathways for each patient, the scope for conducting CACE analysis was limited, as assumptions of the analysis did not hold. Only one treatment comparison was conducted at the primary end point at 12 months, that of compliance with ESP as defined in the fidelity section of this report (*Table 8*).

Instrumental variable regression was implemented predicting OSS at the primary end point at 12 months in order to quantify the effect of compliance with ESP. From the model, outcomes for ESP compliers remained lower than for patients in the other treatment arms (-1.84 OSS points, 95% CI -4.41 to 0.74 OSS points; $p = 0.157$); however, the difference was not statistically significant. Based on *Appendix 8*, patients tended to have better outcomes if they completed their randomised treatment.

TABLE 12 Estimated mean OSS differences by treatment arm (estimates from model incorporating follow-ups before and after treatment in addition to post-randomisation outcomes^a)

Time point	Treatment arm		Difference, mean (95% CI)	p-value
	Mean (95% CI)	Mean (95% CI)		
	MUA	ESP		
3 months	28.2 (27.1 to 29.3)	29.4 (27.8 to 30.9)	-1.18 (-3.10 to 0.73)	0.23
6 months	32.5 (31.5 to 33.5)	32.7 (31.2 to 34.1)	-0.15 (-1.90 to 1.60)	0.87
12 months	41.1 (40.0 to 42.3)	39.2 (37.5 to 40.9)	1.92 (-0.16 to 4.00)	0.07
	ACR	ESP		
3 months	26.0 (24.9 to 27.2)	29.4 (27.8 to 30.9)	-3.33 (-5.25 to -1.40)	< 0.01
6 months	31.5 (30.5 to 32.5)	32.7 (31.2 to 34.1)	-1.13 (-2.88 to 0.62)	0.21
12 months	42.5 (41.3 to 43.7)	39.2 (37.5 to 40.9)	3.26 (1.18 to 5.35)	< 0.01
	ACR	MUA		
3 months	26.0 (24.9 to 27.2)	28.2 (27.1 to 29.3)	-2.14 (-3.71 to -0.57)	0.01
6 months	31.5 (30.5 to 32.5)	32.5 (31.5 to 33.5)	-0.98 (-2.40 to 0.44)	0.18
12 months	42.5 (41.3 to 43.7)	41.1 (40.0 to 42.3)	1.35 (-0.33 to 3.02)	0.12

^a Linear mixed random intercept model adjusted for age, sex, diabetes, OSS at baseline (fixed effects) and site (random effect).

TABLE 13 Comparison of baseline characteristics by treatment compliance

Characteristic	MUA		ACR		ESP	
	Completed treatment (N = 164)	Did not complete treatment (N = 37)	Completed treatment (N = 162)	Did not complete treatment (N = 41)	Completed treatment (N = 80)	Did not complete treatment (N = 19)
Sex, n (%)						
Male	54 (33)	18 (49)	63 (39)	14 (34)	29 (36)	6 (32)
Female	110 (67)	19 (51)	99 (61)	27 (66)	51 (64)	13 (68)
Age (years)						
n	164	37	162	41	80	19
Mean (SD)	54.0 (7.4)	56.8 (8.8)	54.3 (7.5)	52.3 (8.3)	55.6 (7.7)	49.8 (6.3)
Median (minimum, maximum)	54 (30, 57)	56 (36, 73)	54 (33, 76)	52 (34, 71)	55 (39, 77)	50 (39, 69)
Diabetic, n (%)						
No	115 (70)	26 (70)	118 (73)	25 (61)	55 (69)	14 (74)
Type 1	12 (7)	-	8 (5)	4 (10)	4 (5)	1 (5)
Type 2	37 (23)	11 (30)	36 (22)	12 (29)	21 (26)	4 (21)
Employment status summary, n (%)						
In paid work	113 (69)	16 (43)	95 (59)	23 (56)	44 (55)	9 (47)
Not in paid work	48 (29)	21 (57)	65 (40)	17 (41)	36 (45)	10 (53)
Missing	3 (2)	-	2 (1)	1 (2)	-	-
Duration of symptoms (months)						
n	160	36	161	40	79	19
Mean (SD)	10.6 (8.5)	10.3 (8.9)	11.5 (10.5)	10.56 (7.5)	10.3 (6.1)	13.0 (15.8)
Median (minimum, maximum)	8 (2, 60)	7.5 (2, 48)	9 (2, 96)	9 (0, 36)	9 (3, 36)	8 (2, 72)
Previous physiotherapy for affected shoulder, n (%)						
Yes	100 (61)	25 (68)	99 (61)	25 (61)	49 (6)	10 (53)
No	64 (39)	12 (32)	63 (39)	14 (34)	31 (39)	8 (42)
Missing	-	-	-	2 (5)	-	1 (5)
OSS, points (0-48)						
n	163	37	161	41	80	19
Mean (SD)	20.4 (8.9)	20.8 (8.9)	19.0 (7.6)	19.9 (8.4)	20.7 (7.8)	18.3 (8.4)
Median (minimum, maximum)	20 (2, 48)	20 (3, 36)	19 (1, 37)	19 (4, 35)	20 (2, 42)	18 (4, 34)

Missing data

Possible predictors of missing OSS data at 3-, 6- or 12-month follow-up are presented in *Table 14*. Only age (younger participants being more likely to have missing data) and OSS outcomes prior to the time of missing data (participants with poorer outcomes being more likely to have missing data) were significant predictors of missingness. As these are already covariates in the primary analysis model, no model adjustments were undertaken.

TABLE 14 Comparison of patient characteristics by missingness of OSS over time

	Not missing	Missing	p-value
3-month follow-up	N = 447	N = 56	
Age (years), mean (SD)	54.6 (7.6)	51.5 (8.4)	0.01 ^a
Male, n (%)	161 (36%)	23 (41%)	0.46
Diabetic, n (%)	128 (29%)	22 (39%)	0.10
In employment, n (%)	270 (60%)	30 (54%)	0.45
Duration of symptoms (months), mean (SD)	11.1 (9.4)	9.6 (7.1)	0.26
Previous physiotherapy, n (%)	281 (63%)	27 (48%)	0.07
Baseline OSS (points), mean (SD)	20.1 (8.2)	18.5 (8.6)	0.19
6-month follow-up	N = 430	N = 73	
Age (years), mean (SD)	54.7 (7.6)	51.6 (8.1)	< 0.01 ^a
Male, n (%)	155 (36%)	29 (40%)	0.55
Diabetic, n (%)	122 (28%)	28 (38%)	0.09
In employment, n (%)	257 (60%)	43 (59%)	0.90
Duration of symptoms (months), mean (SD)	11.0 (9.1)	10.2 (9.9)	0.50
Previous physiotherapy, n (%)	266 (62%)	42 (58%)	0.65
Baseline OSS (points), mean (SD)	20.2 (8.2)	18.2 (8.4)	0.06
3-month OSS (points), mean (SD)	30.4 (10.9)	26.2 (11.9)	0.05 ^a
12-month follow-up	N = 446	N = 57	
Age (years), mean (SD)	54.5 (7.5)	52.1 (9.0)	0.03 ^a
Male, n (%)	164 (37%)	20 (35%)	0.80
Diabetic, n (%)	136 (30%)	14 (25%)	0.36
In employment, n (%)	266 (60%)	34 (60%)	0.82
Duration of symptoms (months), mean (SD)	11.0 (9.4)	10.4 (7.8)	0.68
Previous physiotherapy, n (%)	278 (62%)	30 (53%)	0.26
Baseline OSS (points), mean (SD)	20.1 (8.2)	18.3 (8.7)	0.11
3-month OSS (points), mean (SD)	30.4 (10.8)	25.9 (13.4)	0.05
6-month OSS (points), mean (SD)	37.4 (10.1)	36.5 (10.7)	0.72

a Statistically significant at the 5% level.

Based on the low drop-out rate at the primary end point at 12 months (11%), and the fact that nearly all patients could be included in the primary analysis (94%), further adjustments for missing data, such as multiple imputation, were not implemented.

Other secondary analyses

Further secondary analyses excluded responses received beyond 6 weeks of each intended follow-up and adjusted for the observed baseline imbalance in employment status (see *Appendix 9*). The results were similar to those observed in the primary analysis.

Subgroup analyses

The possibility of differential treatment effects were explored for subgroups based on diabetes status, receipt of previous physiotherapy and baseline treatment preference, and, in addition, length of frozen shoulder symptoms at baseline following advice from the trial oversight committee. Interaction terms between treatment allocation and subgroups were added to the primary analysis model and *p*-values for interactions for each treatment comparison were derived (Table 15). None of the interaction terms was statistically significant, although the study was not powered to detect such interactions and the number of participants in some of the subgroups in each treatment arm was very small.

Possible trends are illustrated in Figures 5–8 and descriptive tables are in Appendix 10. Diabetic patients tended to have poorer outcomes than non-diabetic patients at all time points, and especially at the 3-month follow-up for patients in the ACR arm. Patients who had previous physiotherapy tended to have worse outcomes if they were randomised to ESP, especially at the 3- and 6-month follow-ups, whereas patients who had indicated a prior preference for physiotherapy tended to have better outcomes if they were randomised to ESP and worse outcomes if they were randomised to either surgical treatment. Participants who reported frozen shoulder symptoms for ≥ 9 months before entering the trial tended to have worse outcomes at 3 months if they were randomised to ACR and better outcomes at 3 months if they were randomised to ESP.

TABLE 15 Subgroup analyses summary

Subgroup	Treatment arm (n)			Treatment comparison	Contrast	95% CI	p-value of allocation interaction with subgroup
	MUA	ACR	ESP				
Diabetes							
Diabetic (n = 150)	60	60	30	MUA vs. ESP	-0.34	-4.58 to 3.90	0.88
				ACR vs. ESP	0.09	-4.16 to 4.34	0.97
Not diabetic (n = 353)	141	143	69	ACR vs. MUA	0.43	-2.98 to 3.85	0.80
Previous physiotherapy							
Had previous physiotherapy (n = 308)	125	124	59	MUA vs. ESP	-2.08	-6.04 to 1.89	0.30
				ACR vs. ESP	-0.87	-4.86 to 3.12	0.67
Did not have previous physiotherapy (n = 192)	76	77	39	ACR vs. MUA	1.21	-2.02 to 4.44	0.46
Patient treatment preference							
Allocated to preferred treatment (n = 131)	56	64	11	MUA vs. ESP	2.10 ^a	-4.32 to 8.52	0.81
				ACR vs. ESP	1.28 ^b	-4.66 to 7.22	
Allocated to non-preferred treatment (n = 105)	40	27	38	ACR vs. ESP	3.11 ^a	-3.50 to 9.73	0.65
				ACR vs. MUA	2.18 ^b	-3.73 to 8.09	
Had no treatment preference (n = 263)	103	111	49	ACR vs. MUA	1.01 ^a	-3.84 to 5.87	0.87
				MUA vs. ESP	0.90 ^b	-2.70 to 4.50	
Duration of symptoms at baseline							
< 9 months (n = 249)	103	95	51	MUA vs. ESP	-2.41	-6.29 to 1.46	0.22
				ACR vs. ESP	-2.00	-5.85 to 1.85	0.31
≥ 9 months (n = 246)	93	106	47	ACR vs. MUA	0.41	-2.73 to 3.56	0.80

a Allocated to non-preferred vs. allocated to preferred treatment.

b No treatment preference vs. allocated to preferred treatment.

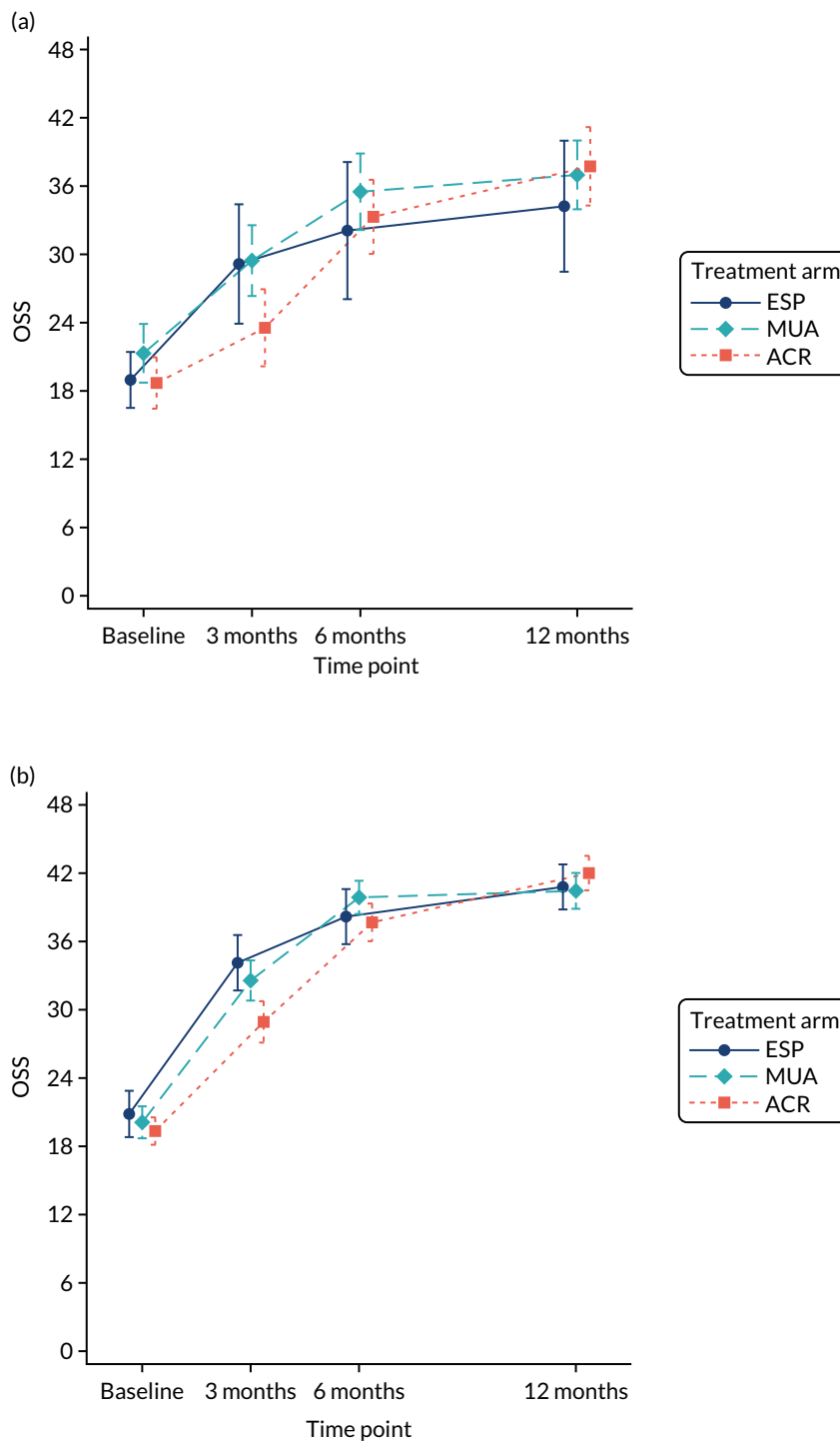


FIGURE 5 Unadjusted mean OSS function items and 95% CIs by treatment arm and diabetes: (a) diabetic; and (b) not diabetic.

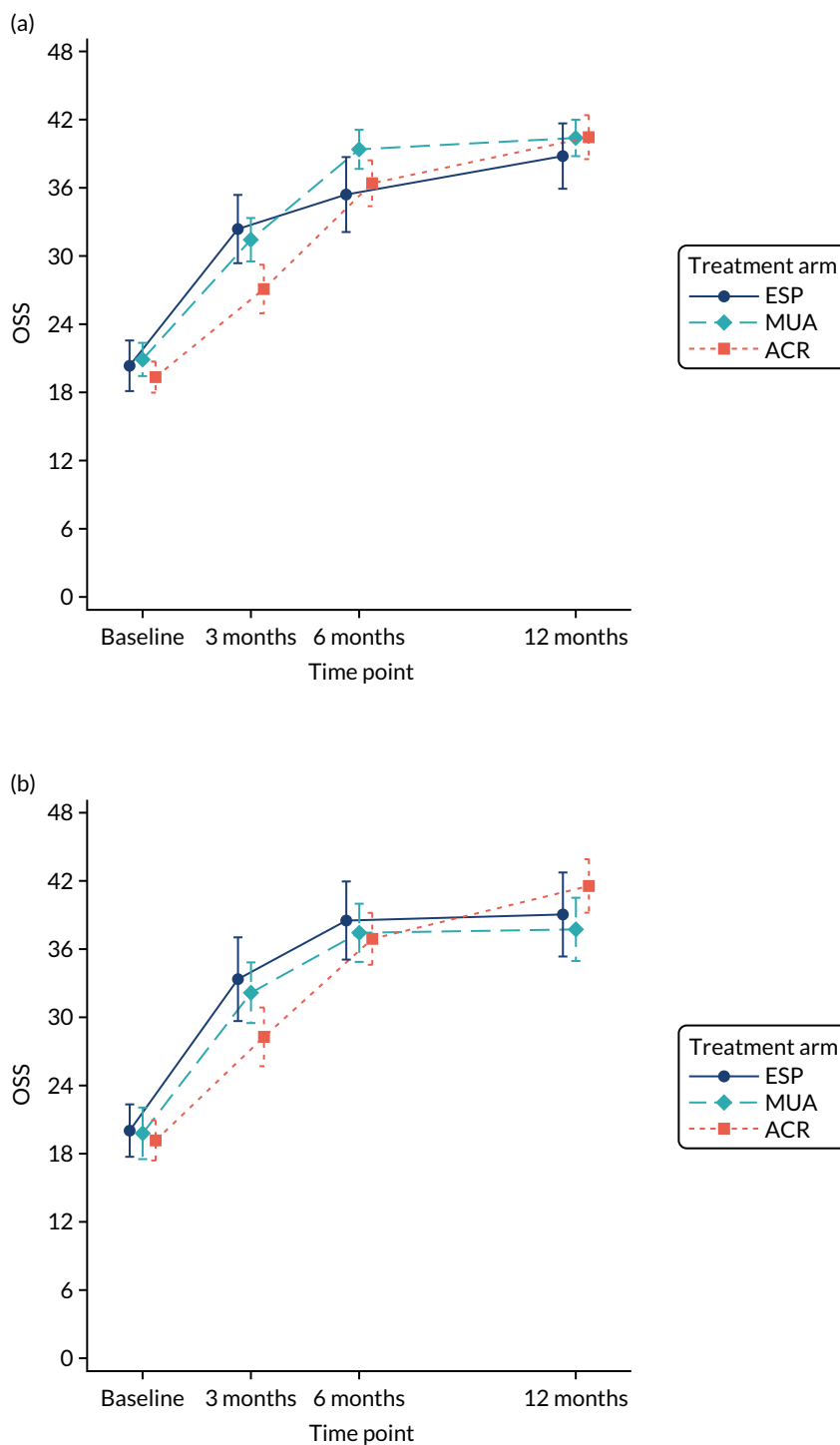


FIGURE 6 Unadjusted mean OSS function items and 95% CIs by treatment arm and previous physiotherapy: (a) had previous physiotherapy; and (b) did not have previous physiotherapy.

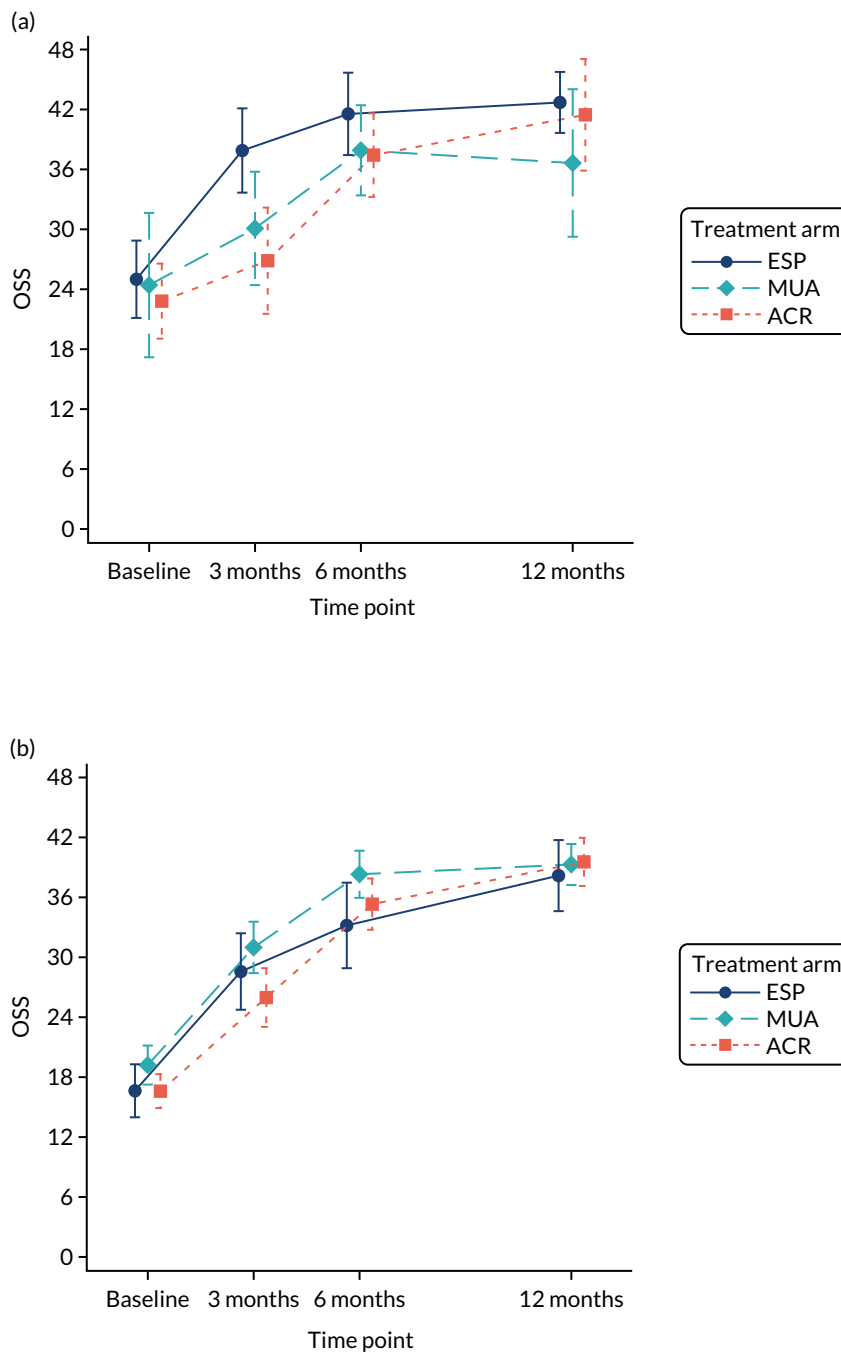


FIGURE 7 Unadjusted mean OSS function items and 95% CIs by treatment arm and baseline preference: (a) preference for physiotherapy; (b) preference for surgery; and (c) no preference. (continued)

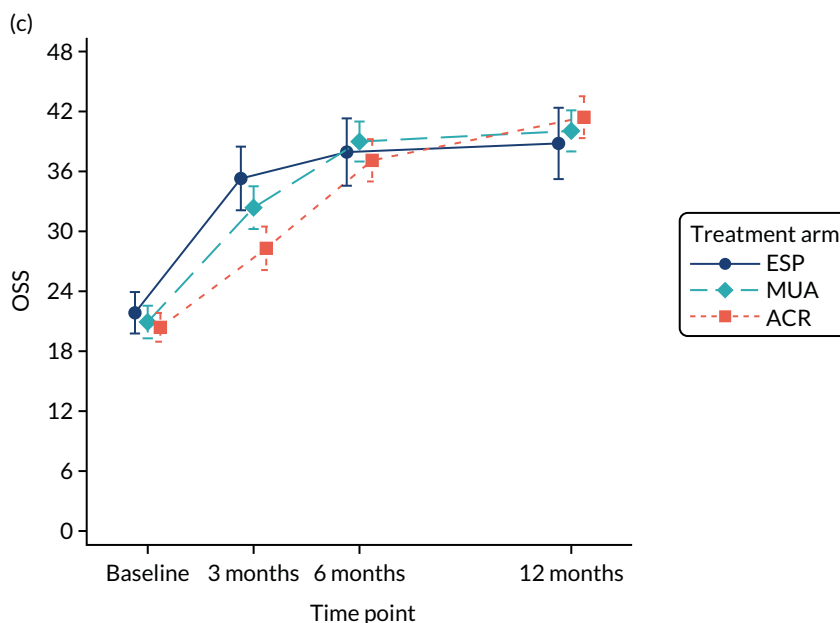


FIGURE 7 Unadjusted mean OSS function items and 95% CIs by treatment arm and baseline preference: (a) preference for physiotherapy; (b) preference for surgery; and (c) no preference.

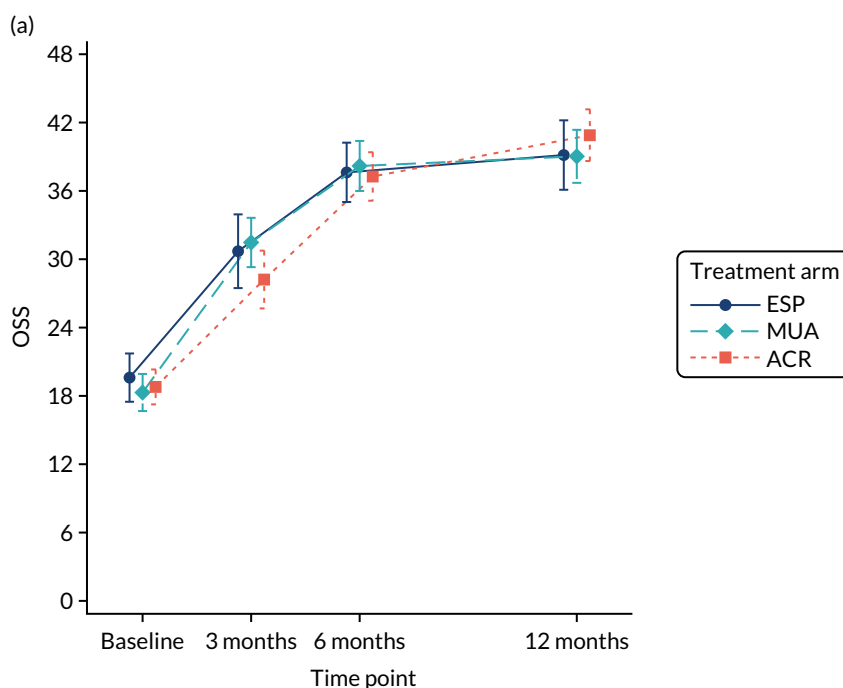


FIGURE 8 Unadjusted mean OSS function items and 95% CIs by treatment arm and length of time with symptoms: (a) symptoms for < 9 months; and (b) symptoms for ≥ 9 months. (continued)

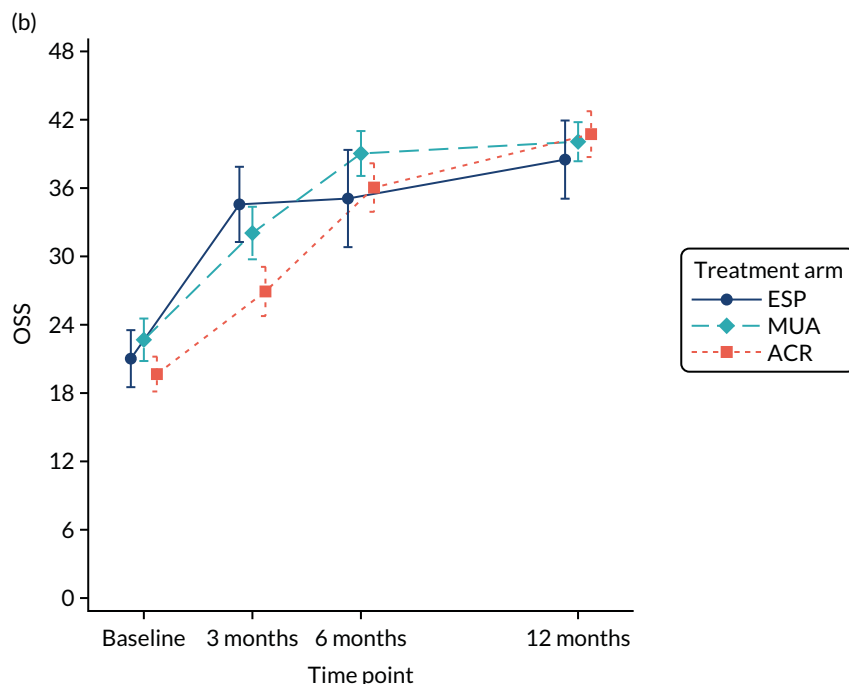


FIGURE 8 Unadjusted mean OSS function items and 95% CIs by treatment arm and length of time with symptoms: (a) symptoms for < 9 months; and (b) symptoms for ≥ 9 months.

Secondary outcomes

Among the secondary outcomes, QuickDASH and shoulder pain followed a similar pattern to the OSS, in that ACR patients were observed to have significantly poorer outcomes at 3 months (note that many patients had only recently had or were still waiting for their surgery at this point) but better outcomes at 12 months post randomisation than those allocated to MUA or ESP (*Tables 16 and 17*). Unadjusted means are presented and illustrated in the tables and figures below (see *Appendix 11 and Figures 9 and 10*).

QuickDASH

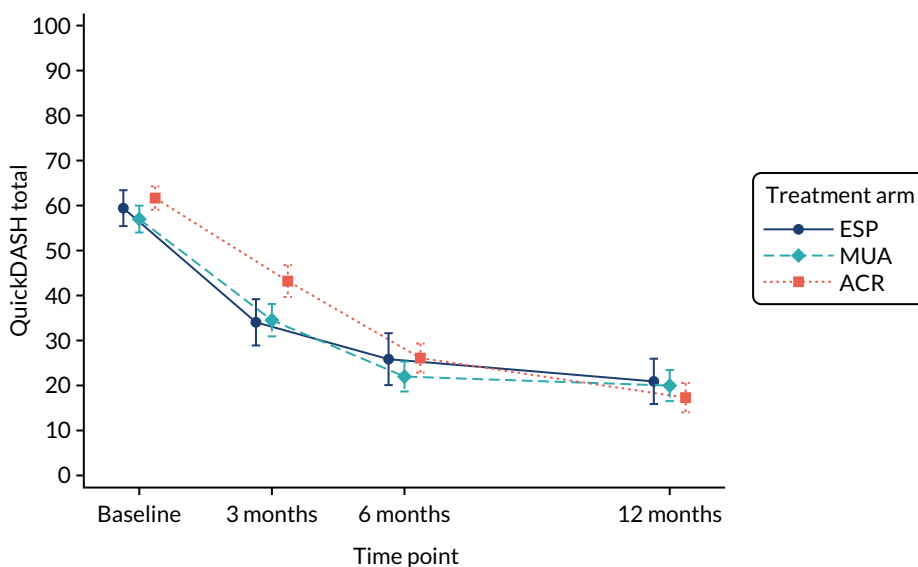


FIGURE 9 Unadjusted mean QuickDASH scores and 95% CIs by treatment arm.

TABLE 16 Estimated mean QuickDASH differences by treatment arm^a

Time point	Mean (95% CI)	Mean (95% CI)	Mean difference (95% CI)	p-value
	MUA	ESP		
3 months	38.8 (35.7 to 42.0)	37.1 (32.7 to 41.4)	1.77 (-3.41 to 6.96)	0.50
6 months	25.7 (22.6 to 28.7)	29.2 (24.9 to 33.5)	-3.55 (-8.68 to 1.58)	0.18
12 months	22.9 (19.8 to 26.0)	23.4 (19.0 to 27.8)	-0.50 (-5.70 to 4.70)	0.85
	ACR	ESP		
3 months	44.4 (41.3 to 47.5)	37.1 (32.7 to 41.4)	7.33 (2.16 to 12.49)	< 0.01
6 months	27.4 (24.4 to 30.4)	29.2 (24.9 to 33.5)	-1.82 (-6.94 to 3.31)	0.49
12 months	18.2 (15.1 to 21.3)	23.4 (19.0 to 27.8)	-5.20 (-10.42 to 0.02)	0.05
	ACR	MUA		
3 months	44.4 (41.3 to 47.5)	38.8 (35.7 to 42.0)	5.55 (1.32 to 9.78)	0.01
6 months	27.4 (24.4 to 30.4)	25.7 (22.6 to 28.7)	1.73 (-2.39 to 5.86)	0.41
12 months	18.2 (15.1 to 21.3)	22.9 (19.8 to 26.0)	-4.71 (-8.91 to -0.50)	0.03

a Linear mixed covariance pattern model adjusted for age, sex, diabetes, QuickDASH at baseline (fixed effects) and site (random effect).

Shoulder pain

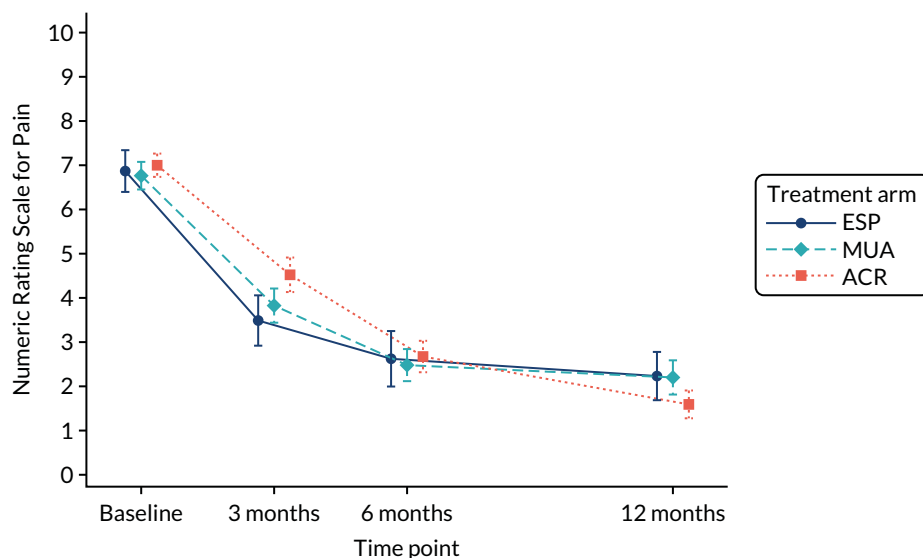


FIGURE 10 Unadjusted mean shoulder Numeric Rating Scale for Pain and 95% CIs by treatment arm.

TABLE 17 Estimated shoulder Numeric Rating Scale for Pain differences by treatment arm^a

Time point	Mean (95% CI)	Mean (95% CI)	Mean difference (95% CI)	p-value
	MUA	ESP		
3 months	4.1 (3.8 to 4.5)	3.7 (3.2 to 4.2)	0.43 (-0.17 to 1.03)	0.16
6 months	2.8 (2.4 to 3.1)	3.0 (2.5 to 3.5)	-0.19 (-0.78 to 0.40)	0.53
12 months	2.4 (2.1 to 2.8)	2.5 (2.0 to 3.0)	-0.08 (-0.66 to 0.50)	0.78
	ACR	ESP		
3 months	4.7 (4.3 to 5.1)	3.7 (3.2 to 4.2)	1.02 (0.42 to 1.61)	< 0.01
6 months	2.8 (2.5 to 3.2)	3.0 (2.5 to 3.5)	-0.14 (-0.74 to 0.45)	0.63
12 months	1.7 (1.4 to 2.0)	2.5 (2.0 to 3.0)	-0.81 (-1.39 to -0.23)	< 0.01
	ACR	MUA		
3 months	4.7 (4.3 to 5.1)	4.1 (3.8 to 4.5)	0.59 (0.10 to 1.07)	0.02
6 months	2.8 (2.5 to 3.2)	2.8 (2.4 to 3.1)	0.05 (-0.43 to 0.52)	0.85
12 months	1.7 (1.4 to 2.0)	2.4 (2.1 to 2.8)	-0.73 (-1.20 to -0.25)	< 0.01

a Linear mixed covariance pattern model adjusted for age, sex, diabetes, Numeric Rating Scale for Pain at baseline (fixed effects) and site (random effect).

Extent of recovery

There was no evidence of statistically significant differences between treatment arms in the reduction in frozen shoulder symptoms as measured by the extent of recovery ('To what extent would your frozen shoulder symptoms in the past 24 hours prompt you to ask for further treatment?'; response on visual analogue scale of 0–100). Descriptives are presented in *Appendix 11* (see *Table 53*) and illustrated in *Figure 11*, and the results of the analysis are in *Table 18*.

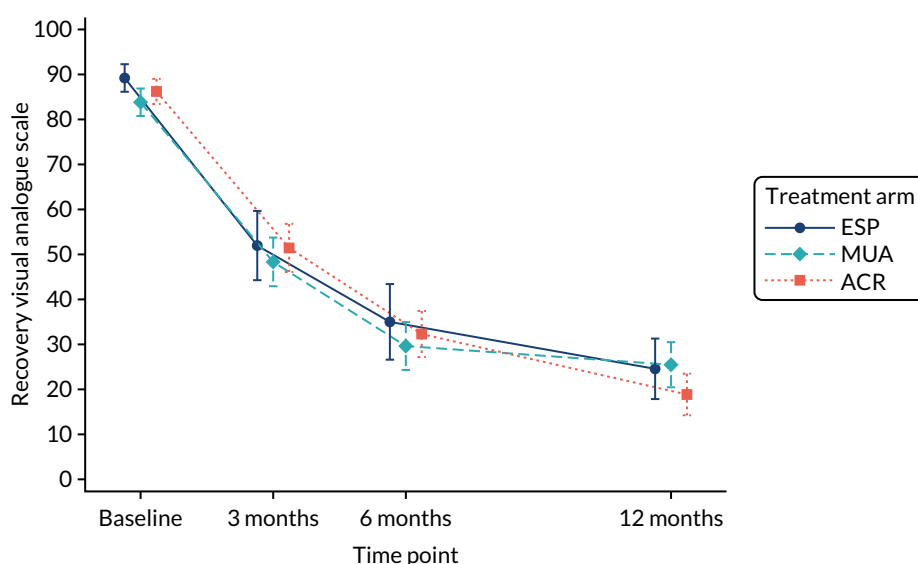


FIGURE 11 Unadjusted mean extent of recovery visual analogue scale and 95% CIs by treatment arm. Scale is 0–100, where 100 is equivalent to maximum belief that symptoms require further treatment.

TABLE 18 Estimated mean extent of recovery^a visual analogue scale differences by treatment arm^b

Time point	Mean (95% CI)	Mean (95% CI)	Mean difference (95% CI)	p-value
	MUA	ESP		
3 months	51.4 (45.8 to 56.9)	53.9 (46.3 to 61.5)	-2.55 (-11.68 to 6.58)	0.58
6 months	31.9 (26.5 to 37.2)	38.6 (30.9 to 46.3)	-6.71 (-15.83 to 2.42)	0.15
12 months	27.3 (22.4 to 32.3)	26.9 (20.0 to 33.8)	0.46 (-7.79 to 8.70)	0.91
	ACR	ESP		
3 months	54.0 (48.5 to 59.5)	53.9 (46.3 to 61.5)	0.11 (-9.02 to 9.23)	0.98
6 months	34.7 (29.3 to 40.0)	38.6 (30.9 to 46.3)	-3.93 (-13.06 to 5.21)	0.40
12 months	21.2 (16.3 to 26.2)	26.9 (20.0 to 33.8)	-5.65 (-13.91 to 2.61)	0.18
	ACR	MUA		
3 months	54.0 (48.5 to 59.5)	51.4 (45.8 to 56.9)	2.66 (-4.84 to 10.15)	0.49
6 months	34.7 (29.3 to 40.0)	31.9 (26.5 to 37.2)	2.78 (-4.50 to 10.06)	0.45
12 months	21.2 (16.3 to 26.2)	27.3 (22.4 to 32.3)	-6.11 (-12.86 to 0.64)	0.08

a Scale is 0–100, where 100 is equivalent to the maximum belief that symptoms require further treatment.

b Linear mixed covariance pattern model adjusted for age, sex, diabetes, symptom severity at baseline (fixed effects) and site (random effect).

Stiffness

Although stiffness was not collected as a separate outcome, it was of interest whether the trial interventions differentially addressed pain or stiffness associated with frozen shoulder. The proportion of predominant pain or stiffness reported by patients at the first and last physiotherapy session of their treatment is presented in *Appendix 11* (see *Table 54*). Patients in the ESP arm had relatively lower levels of predominant pain by the end of physiotherapy, whereas patients in the ACR arm had relatively lower levels of predominant stiffness than those in the other arms.

Complications/adverse events

Any reported complications were reconciled with recorded AEs by two senior surgeons (independently initially and by consensus following any disagreement) to arrive at a single record of untoward occurrences. Some variables recorded as standard through the AE reporting process (e.g. expectedness and event severity following clinical review) were not available or relevant for any events identified from the complications form or change in status form alone. This is why some of the information appears as missing in the events listed in *Table 19*. Only possible relatedness to the trial treatments was recorded retrospectively, where this information was missing.

In total, there were only 10 SAEs, reported for nine patients (summarised in *Appendix 12*; an itemised list is in *Table 19*). All SAEs occurred for patients randomised to the surgical arms (ACR, $n = 8$; MUA, $n = 2$). However, one SAE in the ACR arm was experienced by a participant who had non-trial-specific physiotherapy. The events mainly related to serious medical complications such as chest infection or stroke, some of which may be related to having received surgery in general, rather than being specifically related to the trial surgical procedures. Numbers were insufficient to allow formal analysis.

Thirty-three non-serious AEs were reported for 31 patients, with comparable rates in the three arms (7% of MUA patients, 6% of ACR patients and 5% of ESP patients). These events were mainly expected and often related to persistent or worsening shoulder pain (summarised in *Appendix 12*; an itemised list is in *Table 20*). Sufficient numbers of patients experienced one or more AE to allow for a valid statistical comparison between the two surgical arms, which confirmed no evidence of statistical differences in the proportion of non-serious AEs ($p = 0.186$).

TABLE 19 Serious adverse events (itemised)

Source	Treatment		Description	Type	Relatedness	Potentially long-lasting consequences
	Allocated	Received				
SAE CRF	MUA	None	Attended A&E for numbness of right arm and heaviness with kaleidoscope vision and headache	Medically important	Not related	No
SAE CRF	ACR	ACR	Elevated blood sugar levels	Prolonged hospitalisation	Probably related	No
SAE CRF	ACR	ACR	Decreased oxygen saturation	Prolonged hospitalisation	Not related	No
SAE CRF	ACR	ACR	Hypoglycaemic seizure while under anaesthetic	Prolonged hospitalisation	Unlikely to be related	No
SAE CRF	ACR	ACR	Patient noticed facial drooping/ weakness after surgery	Medically important	Definitely related	No
Review	MUA	MUA	Septic joint arthritis	– ^a	Definitely related	Yes
Review	ACR	ACR	Stroke	– ^a	Not related	Yes
Review	ACR	MUA	Likely anterior dislocation	– ^a	Definitely related	Yes
Review	ACR	Other	Deep-vein thrombosis	– ^a	Not related	Yes
Review	ACR	ACR	Chest infection	– ^a	Unlikely to be related	No

A&E, accident and emergency.

^a Event was identified following review rather than from the AE CRF; not all information is available.

TABLE 20 Non-serious AEs (itemised)

Source	Treatment		Description	Related	Expectedness	Severity
	Allocated	Received				
AE CRF	ESP	ESP	Persistent pain	Not related	Expected	Unknown
AE CRF	ESP	ESP	Long head of the biceps tendon pain and rupture	Not related	Unexpected	Mild
AE CRF	ESP	ACR	Post-procedural worsening of shoulder pain	Possibly related	Expected	Mild
AE CRF	ESP	ACR	Recurrent stiffness requiring further treatment	Not related	Expected	Moderate
AE CRF	MUA	MUA	Transient hyperglycaemia, steroid flare or joint sepsis following corticosteroid injection	Possibly related	Expected	Mild
AE CRF	MUA	MUA	Additional diagnosis requiring further treatment	Not related	Expected	Severe
AE CRF	MUA	MUA	Post-procedural worsening of shoulder pain	Possibly related	Expected	Mild
AE CRF	MUA	MUA	Transient hyperglycaemia, steroid flare or joint sepsis following corticosteroid injection	Probably related	Expected	Unknown
AE CRF	MUA	MUA	Ipsilateral face swelling, face flushed and neck and face hot	Possibly related	Unexpected	Moderate

TABLE 20 Non-serious AEs (itemised) (continued)

Source	Treatment		Description	Related	Expectedness	Severity
	Allocated	Received				
AE CRF	MUA	MUA	Neuropathic symptoms	Not related	Unexpected	Moderate
AE CRF	MUA	MUA	Post-procedural worsening of shoulder pain	Unlikely to be related	Expected	Moderate
AE CRF	MUA	MUA	Injury to adjacent structures such as nerve, tendon, bone or joint	Possibly related	Expected	Severe
AE CRF	MUA	MUA	Post-procedural worsening of shoulder pain	Not related	Expected	Moderate
AE CRF	MUA	MUA	Persistent pain requiring further treatment	Unlikely to be related	Expected	Moderate
AE CRF	MUA	MUA	Persistent stiffness and pain requiring treatment	Not related	Unexpected	Unknown
AE CRF	MUA	ESP	Transient hyperglycaemia, steroid flare or joint sepsis following corticosteroid injection	Definitely related	Expected	Mild
AE CRF	ACR	ACR	Infection	Possibly related	Expected	Mild
AE CRF	ACR	ACR	Persistent pain	Possibly related	Expected	Mild
AE CRF	ACR	ACR	Post-procedural worsening of shoulder pain	Definitely related	Expected	Mild
AE CRF	ACR	ACR	Persistent pain requiring further treatment	Possibly related	Expected	Moderate
AE CRF	ACR	ACR	Neuropathic symptoms	Unlikely to be related	Expected	Mild
AE CRF	ACR	ACR	Adverse reaction to concurrent medication	Possibly related	Unexpected	Severe
AE CRF	ACR	ACR	Allergic reaction to dressing	Definitely related	Unexpected	Mild
AE CRF	ACR	ACR	Post-procedural worsening of shoulder pain	Possibly related	Expected	Mild
AE CRF	ACR	MUA	Injury to adjacent structures such as nerve, tendon, bone or joint	Definitely related	Expected	Severe
AE CRF	ACR	Other	Neuropathic symptoms	Unlikely to be related	Unexpected	Moderate
Review	ESP	ESP	Supraspinatus tendinopathy	— ^a	— ^a	— ^a
Review	MUA	MUA	Episode of inflammation	— ^a	— ^a	— ^a
Review	MUA	MUA	Pins and needles in hand	— ^a	— ^a	— ^a
Review	MUA	MUA	Chest infection	— ^a	— ^a	— ^a
Review	ACR	ACR	Post-procedural worsening of shoulder pain	— ^a	— ^a	— ^a
Review	ACR	ACR	Patient being investigated for neck problems	— ^a	— ^a	— ^a
Review	ACR	Other	Surgical site infection	— ^a	— ^a	— ^a

^a Event was identified following review rather than from the AE CRF; not all information is available.

Other analyses

Treatment preferences

Summaries of treatment preference data are presented in *Appendix 13*. Non-consenting patients tended to have a preference for ACR, which they expected to be more effective. Although randomised patients also expected ACR to be the most effective treatment, they were more likely than non-consenting patients to be undecided about the effectiveness of any of the treatments. At the end of 12 months' follow-up, many patients had changed their preference to the treatment they had received. ACR remained the most popular, especially among participants who had received this as their preferred treatment at baseline.

Oxford Shoulder Score change scores

Details of participants' comparative assessment of their symptoms at baseline and 12 months with reference to their change OSS are presented in *Appendix 14*. Unfortunately, this analysis was not able to reveal a more nuanced understanding of minimal clinically meaningful differences using the OSS, as symptoms of the majority of participants improved substantially over the course of the trial, which was associated with very large increases in OSS scores.

Outcomes for patients receiving no treatment

The OSS and QuickDASH scores for participants who did ($n = 441$) and participants who did not ($n = 62$) receive any treatment for their frozen shoulder are presented in *Appendix 15*. Patients for whom no treatment was recorded tended to have progressively lower rates of improvement by 6- and 12-month follow-up; however, the proportion of participants with available data was also much smaller in this group (e.g. 66% of valid OSS scores at 12 months vs. 92% of participants who did receive treatment).

Systematic review: integrating the new evidence

A systematic review was undertaken to assess the effectiveness of MUA, ACR, hydrodilatation and physiotherapy with steroid injection in the management of patients with a primary frozen shoulder in order to place the findings of UK FROST in the context of existing evidence for these treatments. Nine relevant studies were identified, including UK FROST, which provided by far the most, and the most robust, evidence.

Owing to considerable heterogeneity of the interventions and study populations, only two studies could be pooled as part of a meta-analysis, comparing long-term shoulder functioning for patients receiving either ACR or physiotherapy in UK FROST and one other trial. The pooled effect favoured ACR (standard effect size 0.32, 95% CI 0.08 to 0.56), which was largely determined by the UK FROST results, as the second trial was much smaller ($n = 44$). The pooled effect was of smaller magnitude than the clinical threshold of 5 OSS points, equivalent to a standard effect size of approximately 0.42.

Full details are presented in *Report Supplementary Material 20*.

Chapter 4 Economic evaluation

Objective

The objective of this economic evaluation is to assist decision-making to identify the most efficient provision of future care for the management of frozen shoulder in secondary care in the NHS.

Overview

A prospective economic evaluation was conducted alongside the UK FROST trial, the aim being to estimate the cost-effectiveness of the three most commonly used interventions for the management of frozen shoulder in secondary care. The three interventions in the study were ESP with an intra-articular steroid injection compared with MUA with a steroid injection or ACR followed by manipulation. Both surgical interventions were followed with a programme of PPP.

Costs and health benefits were compared for the three groups over 12 months, and hence discounting was not required. All costs were expressed in Great British pounds at a 2017–18 price base. Health benefits were expressed in terms of QALYs based on patients' health-related quality of life (HRQoL) assessed using the EQ-5D-5L.^{50,66} Differences in mean costs and mean QALYs at 1 year were used to derive an estimate of the cost-effectiveness of surgery and of non-surgical treatment.

The base-case analysis was conducted on an ITT basis as with the statistical analyses in *Chapter 3*. The perspective of the UK NHS and Personal Social Services was adopted for the analysis, and hence costs incurred by families and informal carers were excluded from the base case. A secondary analysis was undertaken from a broader perspective. The National Institute for Health and Care Excellence (NICE) guidelines were applied to all methods used in this economic analysis.⁶⁷

Owing to the impact of missing data, the base-case analysis was conducted as an imputed analysis;⁶⁸ the choice of method for handling missing data (multiple imputation) was grounded in the assumed missing data mechanism (MAR), which in turn was supported by the UK FROST data set. The impact of alternative assumptions about the missing data mechanism was carefully assessed in sensitivity analyses.

Methods

Data sources

The data required for the analyses were collected from both participants (self-reported using postal questionnaires) and health-care professionals (via hospital forms) during the 12 months' follow-up.

Data relating to surgical care were collected using surgical forms that were specifically designed for the trial. Similarly, physiotherapy logs were completed by physiotherapists providing patient care. These logs were used to record the essential components of physiotherapy at each session for each participant, and they were also used to estimate the cost of ESP and the costs of PPP following MUA or ACR.

Data on resource use in primary health-care consultations were collected using participant questionnaires only. All resource use data recorded by participants were split into 'shoulder related' and 'non-shoulder related' and were collected at 3, 6 and 12 months. The base-case analysis was based on shoulder-related resource use. Hospital stays and hospital outpatient appointments were recorded on two sources (patient questionnaires and hospital forms). Our health economic analysis plan indicated that when data could be sourced from patient questionnaires and hospital forms, hospital forms would be used as the main source for calculating resource use. Two main hospital data sources were available for the analysis: (1) complication form and (2) change in status form. On these forms was recorded any hospitalisation from discharge after initial treatment up to 12 months.

Sensitivity analysis explored the impact of including both shoulder-related and non-shoulder-related resource use in the results. As stated before, hospital cost data were available from two sources (self-reported questionnaires and hospital forms). To avoid estimation bias from using multiple sources for the analysis of the same cost, the sensitivity analysis of non-shoulder-related costs was restricted to primary care data, as these data were collected using patient questionnaires exclusively.

Broader resource use data (i.e. private care and productivity costs) were collected during the period between randomisation and 12 months after enrolment into the study, which was also analysed as per the broader sensitivity analysis.

Data on health benefits, expressed in terms of HRQoL, were elicited from participants using the EQ-5D-5L at baseline and at 3, 6 and 12 months.

Measurement of resource use and costs

There are two main cost components in the analysis: (1) the cost of both non-surgical (i.e. ESP) and surgical (i.e. MUA and ACR) interventions; and (2) the costs of health-care usage at both primary and secondary care level.

Surgery (manipulation under anaesthesia and arthroscopic capsular release)

An accurate record of procedures at hospital level (e.g. centres in the trial) was put in place of per-patient information on surgical procedures and complications related to surgery. Data extracted from surgical forms, which include the main items of resource use relating to each operation, were used to calculate the costs of MUA and ACR. The PPP form was used to cost post-surgical physiotherapy care for participants receiving MUA or ACR.

Non-surgery (early structured physiotherapy)

The structured physiotherapy form was used for patients receiving ESP. This form was used to record information on the physiotherapy sessions (i.e. the duration of the session and the staff band of the physiotherapist delivering the session). Information on physiotherapy visits was also available from participant questionnaires at 3, 6 and 12 months. As stated in our health economic analysis plan, PPP and structured physiotherapy forms were used as the primary sources of the base-case analysis. As part of ESP, patients were offered an intra-articular steroid injection at the earliest opportunity. To cost injections, we collected information on the type of steroid used (e.g. methylprednisolone acetate, triamcinolone acetonide, triamcinolone hexacetonide or another steroid), the dose of steroid used (i.e. 20 mg, 30 mg, 40 mg or another dose), whether local anaesthetic or image guidance was used, and the job title of the person administering the injection (i.e. specialist registrar, associate specialist, consultant, band 6 physiotherapist or band 7 physiotherapist).

Health-care consultations and hospital care

Data on health-care consultations and hospital care were used to assess whether or not participants allocated to MUA and ACR experienced different levels of resource use from those in the ESP group. The costs of health-care consultations consist of all costs of visits to both primary and secondary health-care professionals. Participant questionnaires were used to estimate the number of visits to

primary care facilities (e.g. contacts with a GP and general practice nurse), visits to the physiotherapist and use of community care (occupational therapist). Data on resource use were collected at 3, 6 and 12 months. As stated before, hospital forms were used to calculate the number of hospital stays and hospital outpatient appointments because of additional treatments (i.e. treatments the patient received before or during receipt of the randomised treatment), further treatments (i.e. treatments the patients received after completing the randomised treatment), other treatments (i.e. any non-trial treatments the patient received if they did not start or did not complete their randomised treatment) or medical complications.

Resource use items were summarised by trial allocation group and follow-up period.

Estimation of costs

The cost for each trial participant was calculated by multiplying the health-care resource use by the associated unit costs.

Costs relating to both surgical interventions were based on time in theatre, staff time, consumables and length of stay. The staff cost per minute was estimated using PSSRU 2018 (Personal Social Services Research Unit) data⁶⁹ for hospital-based health staff. These unit cost estimates were inclusive of staff salaries, salary on-costs, overheads and capital overheads. To cost length of stay, we used *NHS Reference Costs 2017 to 2018*,⁷⁰ taking the weighted average inpatient bed-day for all major and intermediate shoulder procedures (footnote to Healthcare Resource Group codes). Drug tariffs per milligram for medications (i.e. anaesthesia, antibiotics and steroid injections) were obtained from the *British National Formulary*.⁷¹

Costs relating to the ESP intervention were based on staff time. The cost of a physiotherapist per hour was estimated using PSSRU 2018⁶⁹ based on data for hospital-based health staff (bands 5–8). The full course of ESP was up to 12 sessions; exceptionally, the physiotherapist could decide that more than 12 sessions were needed. The costs relating to the ESP intervention comprised the costs of the physiotherapy sessions and the cost of the steroid injection, the latter of which was obtained from the *British National Formulary*.⁷⁰

The use of other hospital-based care was valued by applying unit costs extracted from national tariffs.^{69,71} Similarly, costs of the primary care and community-based services were estimated by applying unit costs from national tariffs^{69,71} to resource volumes. Other costs included lost productivity measured as missed work; the costs of time taken off work were estimated by applying costs from the Office for National Statistics⁷² to occupational information derived from self-reported work status information.

Costs were estimated in Great British pounds and based on the financial year 2017–18. *Table 21* details the unit costs used in the analysis. The total cost has three main components: (1) the cost of the initial intervention (i.e. ESP, MUA or ACR); (2) the cost of hospital stays and hospital outpatient appointments after the initial intervention; and (3) the cost of visits to primary and community health-care professionals (GP, practice/community nurse, physiotherapist and occupational therapist).

The total costs for the base-case analysis included only shoulder-related resource use, except for hospital stay, which included both shoulder and general medical complications that could apply to the affected shoulder-specific and general medical complications. Sensitivity analyses were used to explore the impact of a broader perspective (i.e. private care costs and productivity costs) on the cost-effectiveness results.

The mean [standard error (SE)] costs by cost category and the mean (SE) total cost were estimated for each treatment group, using regression analysis to control for patients' covariates [i.e. age, sex, treatment group, baseline OSS and diabetes (yes/no)].

TABLE 21 Resource use and costs related to initial surgery: MUA and ACR

Characteristic of the surgical procedure	MUA (N = 168)	ACR (N = 170)
Theatre time (minutes), mean (SD)	25.11 (14.20)	76.64 (24.22)
Number of staff in operation, mean (SD)	6.41 (1.42)	6.36 (1.40)
Patients had injection during operation, n (%)	162 (97)	46 (27)
Intervention delivered as day case, n (%)	163 (97)	153 (90)
Intervention delivered as inpatient, n (%)	5 (3)	17 (10)
Length of stay (nights), mean (SD)	1.2 (0.45)	2.8 (7.31)
Patients had PPP in their allocated group	160 (80)	159 (78)
Number of physiotherapy sessions (PPP), mean (SD); maximum	6.42 (4.95); 18	6.65 (4.81); 18
Patients had injection, n (%)	162 (97)	46 (27)
Costs of the components of the surgical procedure (£), mean (SD)		
Staff in theatre	106.45 (79.47)	360.50 (140.21)
Anaesthesia and steroid injection	99.84 (39.34)	219.15 (89.12)
Hospital stay (day case/length of stay)	218.52 (8.03)	1590.81 (398.29)
Total cost of surgical procedure	424.81 (115.55)	2170.46 (431.11)
<i>Cost of surgical procedure – sensitivity analysis</i>	428.57 (242.45)	1308.26 (413.43)
Cost of PPP	213.61 (157.13)	209.44 (152.95)
Total cost of surgical procedure including PPP	638.42 (204.75)	2379.90 (457.88)

Health-related quality of life and quality-adjusted life-years

This economic evaluation took the form of a cost-utility analysis, whereby health outcomes were assessed in terms of QALYs. HRQoL was expressed in terms of utilities, which were obtained from trial patients using the EQ-5D-5L at baseline and at 3, 6 and 12 months. The EQ-5D-5L has two principal measurement components. The first is a descriptive system that defines HRQoL in terms of five dimensions: 'mobility', 'self-care', 'usual activities', 'pain/discomfort' and 'anxiety/depression'. Responses to each dimension are divided into five ordinal levels, coded (1) no problems, (2) slight problems, (3) moderate problems, (4) severe problems and (5) extreme problems/unable to perform. We evaluated the raw EQ-5D-5L scores by domain to examine the movements between levels for each domain by trial arm. The second component is the visual analogue scale, which is used to record a patient's self-rated health on a vertical visual analogue scale, on which the end points are labelled 'the best health you can imagine' and 'the worst health you can imagine'.

According to the responses to the EQ-5D classification system, a health status can be defined and a single index utility can be assigned. A value set for the EQ-5D-5L is now available that reflects the preference of members of the public in England for health states that are defined by the EQ-5D-5L descriptive system.⁷³ However, at the time of this analysis, the most recent guidance issued by NICE regarding the EQ-5D-5L⁷⁴ recommended the use of the mapping function (i.e. crosswalk) developed by van Hout *et al.*⁶⁶ to derive utilities. Therefore, this crosswalk was applied to each set of responses to generate an EQ-5D utility score (preference weight) for each trial participant. The resulting utility scores range from -0.281 to 1.0, with 0 representing death and 1.0 representing full health; values of < 0 indicate health states worse than death.

Differences in the baseline utility values between groups may lead to biases in the results even if these differences are not statistically significant.⁷⁵ Therefore, utility values were adjusted using a univariate generalised linear model, including group as a fixed factor and baseline EQ-5D-5L score as a covariate. Models were estimated separately for each of the time points at which utility data were collected.

We converted the utilities derived from the EQ-5D-5L into QALYs for each patient using the area under the curve method, following the trapezium rule, which assumes linear interpolation between follow-up points.⁷⁶

Incremental mean QALYs between treatments groups were estimated with regression models according to treatment allocation. Despite the randomisation process, which ensures that baseline variables are balanced between trial arms, in practice (regardless of sample size) it is normal to find an imbalance in mean baseline utility. As baseline utility is likely to be correlated with patients' QALYs gained over time, there are robust reasons to control for baseline utilities when estimating QALYs. Therefore, QALYs were analysed both (1) adjusting for baseline EQ-5D-5L and (2) adjusting for baseline EQ-5D-5L plus the same set of covariates used in the clinical effectiveness analysis, which included baseline utility, that is age, sex, treatment group, baseline OSS and diabetes (yes or no).

Missing data

Missing data occur frequently in RCTs, irrespective of how well designed the data collection is. This is a major concern for within-trial cost-effectiveness analyses, as costs and QALYs, the main outcomes in cost-effectiveness analyses, are cumulative measures collected over trial follow-up. Therefore, missing data at one follow-up time point (e.g. one dimension response missing on the EQ-5D at one time point) result in missing aggregate data (e.g. total QALYs over the trial) for that participant. This problem is common in economic evaluations, as the analysis has to draw on all aspects of the study, including resource use and health outcomes. Non-response to questionnaires and returned but incomplete questionnaires reduce, often considerably, the number of data on resource use that are available for analysis. The problem is amplified when there are frequent assessments, as in UK FROST.

Different methods of handling missing data can yield different results and decisions about the value for money of the assessed interventions. Complete-case assessment and available case analysis are proposed as useful preliminary estimations for economic evaluation but should not constitute the base case for within-trial economic evaluation.⁷⁷ Therefore, it was decided prior to the analysis that complete-case assessment would be presented only for comparison purposes. Additionally, the analysis of the missing pattern of the UK FROST data set would support this decision, as the results suggest that data are not missing completely at random (the assumption driving the complete-case mechanism).

An alternative method for addressing missing data in cost-effectiveness analyses alongside RCTs is multiple imputation,⁷⁷ which has been recommended as the appropriate method to reflect the uncertainty in the results of economic evaluations attributable to missing data.⁶⁷ The main assumption that drives the multiple imputation mechanism is that the data are 'missing at random' (MAR). That is, the missing values in the data set may depend on the value of other observed variables in the data set, but, conditional on those values, the data are MAR. A major concern is that the chance that data are missing may be linked directly to the unobserved value itself [missing not at random (MNAR)]; for example, patients with poorer health may be less likely to complete EQ-5D questionnaires. Therefore, it remains important that the choice of method be grounded in the assumed missing data mechanism, which, in turn, should be informed by the available evidence.

Following methodological recommendations for handling missing data in cost-effectiveness analyses conducted within RCTs,^{78,79} we conducted descriptive analyses of the missing data to explore whether or not the MAR assumption is plausible given the actual missing data mechanism of the UK FROST data set. We assessed the number of missing data by trial arm at each follow-up period, explored

missing data patterns using graphical tools and investigated the association between missingness and baseline variables/observed outcomes using logistic regressions.

Based on the results of the descriptive analyses, we could conclude that MAR is a plausible assumption fitting the UK FROST data set. Therefore, multiple imputation was selected to handle missing data for the base-case analysis. Multiple imputation using chained equations⁸⁰ and predicted mean matching were carried out on the EQ-5D-5L at 3, 6 and 12 months, as well as on the total cost estimates. Predicted mean matching is a semiparametric imputation approach that ensures that observed data are used to estimate a predictive model (using the specified covariates) but, instead of replacing missing values with the model predicted values, the nearest observed value is used to fill the missing one. This guarantees that the imputed values are sampled from values in the original data set, and, therefore, no imputed values will lie outside the bounds of the original data distribution. The multiple imputation model was validated by comparing the distribution of the observed UK FROST data with the imputed data using graphical plots to visualise whether or not the distribution of imputed data resembles the distribution of original data. Age, sex, baseline OSS score and diabetes (yes/no) at baseline were included as explanatory variables in the imputation models. In addition, the baseline EQ-5D-5L utility score and all predictors of missingness were included as an explanatory variable in the models. Multiple imputation by chained equations was performed for a total of 60 imputations. The estimates obtained from each imputed data set are combined to generate mean estimates of costs and QALYs, variances and CIs using Rubin's rules.⁸¹

Finally, as it is impossible to know whether data are MNAR or MAR from the observed data, we explored possible departures from the MAR assumption by means of sensitivity analyses, evaluating the impact of assuming that the data are MNAR rather than MAR. In addition, a mixed model, which does not require an imputation process, is also presented as per sensitivity analysis.

Cost-effectiveness analyses

The main cost-effectiveness analyses were conducted following multiple imputations of all cost and outcomes data. The mean difference in costs and QALYs for the base-case analysis was estimated using regression methods for data on costs and QALYs, adjusting for baseline characteristics.

A bivariate regression model – seemingly unrelated regression (SUR) – of costs and QALYs was used to calculate incremental estimates, using conventional decision rules and estimating incremental cost-effectiveness ratios (ICERs) when appropriate.⁸² SUR allows outcomes to be estimated jointly and so brings efficiency gains over ordinary least squares.⁸³ In the bivariate model, incremental costs and QALYs are estimated simultaneously from two separate ordinary least squares regressions, assuming correlation between the error terms in both regressions.⁸⁴ The SUR model used the same set of covariates as the mixed-effect regression model used for the clinical effectiveness analysis [age, sex, baseline EQ-5D score, baseline OSS score and diabetes (yes/no)].

The cost-effectiveness results were expressed in terms of ICERs. These were estimated as the difference in mean costs divided by the difference in mean QALYs between the trial comparators. The ICER is estimated to inform decision-makers about the optimal use of NHS resources. According to standard cost-effectiveness decision rules, four different eventualities are plausible when comparing incremental costs and QALYs. If the new intervention provides better outcomes (positive incremental QALYs) at lower costs (negative incremental costs), it is considered a dominant intervention and, hence, cost-effective. If the new intervention achieves poorer outcomes (negative incremental QALYs) at higher costs (positive incremental costs), it is considered a dominated option and, hence, not cost-effective. Thus, the ICER is considered only if either intervention does not dominate, that is both incremental costs and incremental QALYs are positive (or negative). In these last two situations, to determine whether or not the incremental health gain is worth the incremental cost, the ICER needs to be

compared against a threshold value. For positive incremental costs and QALYs (the most frequent situation in HTA), an intervention will be considered cost-effective only if the ICER is lower than the threshold. According to NICE, the willingness-to-pay threshold for an additional QALY ranges from £20,000 to £30,000.⁶⁷ This threshold has been used by NICE for more than a decade; however, it has recently been suggested that the threshold should be decreased to £13,000 per QALY gained.^{85,86} According to the current established decision rules, if the result of this cost-utility analysis, namely the estimated cost per QALY, is below the £30,000 threshold, the intervention would be considered cost-effective in terms of QALYs gained.

To compute the probability that each intervention is cost-effective at a given cost-effective threshold, the SUR was conducted in a bootstrapping approach on five imputed data sets to generate 10,000 replicates of incremental costs and benefits. These replicates were represented graphically as cost-effectiveness acceptability curves (CEACs). The probability that each intervention is cost-effective is reported at the cost-effectiveness threshold of £20,000–30,000 per QALY applied by NICE and at a threshold of £13,000 per QALY as suggested by recent research.

Sensitivity analyses and uncertainty

The uncertainty around the cost-effectiveness results was explored using sensitivity analyses that explored the robustness of the results to base-case assumptions. This involved re-estimating the main cost-effectiveness outcomes under different scenarios for costs and missing data. We conducted two sensitivity analysis around costs that implied recalculating costs: (1) including non-shoulder costs (ITT); and (2) adopting a broader perspective that included productivity costs and private care costs. A further number of sensitivity analyses were conducted to explore the impact of missing data on cost-effectiveness estimates: (3) restricting the analyses to complete cases following ITT; (4) imputing QALY data at the aggregated level rather than at the index-score level; (5) using a mixed-model approach; and (6) under a MNAR scenario.

Results

Study population

The baseline study population for the economic analysis was 503 patients. In total, 99 patients were allocated to ESP, 201 were allocated to MUA and 203 were allocated to ACR. Nineteen participants fully withdrew from the trial; for those participants we used multiple imputation techniques to impute missing economic data. As mentioned in the clinical section, 16 participants crossed over from their initial randomisation. This involved patients crossing from ESP to ACR ($n = 7$), from MUA to ESP ($n = 4$), from ACR to ESP ($n = 2$) and from ACR to MUA ($n = 3$).

Health-care resource use and costs

Costs of delivering surgery (manipulation under anaesthesia and arthroscopic capsular release)

Detailed resource use and costs of both surgical interventions are given in *Table 21*. Costs relating to surgical procedures are based on the time in theatre, the delivery of anaesthesia and injections, and the length of hospital stay. To estimate the cost of MUA and ACR, this included participants who had these interventions across any of the treatment groups. MUA surgical information was available for 168 participants: patients allocated to MUA ($n = 164$), patients who withdrew from treatment but still consumed surgical resources ($n = 2$) and patients allocated to ACR who crossed over to MUA and for whom a surgical form was available ($n = 2$). ACR surgical information was available for 170 participants: patients allocated to MUA ($n = 162$), patients who withdrew from treatment but still consumed surgical resources ($n = 3$) and patients allocated to ESP who crossed over to ACR and for whom a surgical form was available ($n = 5$).

The mean cost of MUA was £424.81 (SD £115.55). For 97% of the patients, MUA was delivered as a day-case procedure; only 3% of the patients required hospitalisation (only 1 night). The average duration of MUA was 25.11 minutes (SD 14.20 minutes).

The mean cost of ACR was £2170.46 (SD £431.11). For 90% of the patients ACR was delivered as a day-case procedure; 10% of the patients required hospitalisation of, on average, 2.8 nights (median 1, minimum 1, maximum 31 nights). The average duration of ACR was 76.61 minutes (SD 24.22 minutes).

The cost of PPP was similar for both groups: £213.61 (£157.13) for MUA and £209.44 (£152.95) for ACR.

Non-surgery (early structured physiotherapy intervention)

The total cost of ESP (*Table 22*) includes the cost of the injection and physiotherapy that patients received. The mean cost of ESP was £279.46 (SD £148.56).

The hospital costs related to complications and additional/further/other treatments patients had from discharge after initial treatment up to 12 months are shown in *Table 23*.

Descriptive statistics (mean, median and amount missing) of health-care resource use related to primary and community care, by resource category and by follow-up, are shown in *Table 24*. The results presented are based on the available data set. Although resource use was slightly higher for the ACR group, differences between the groups in resource use in the primary setting appeared small. In terms of dispersion of the results, the median estimates are smaller than the means for all resource use, which suggests that the distributions were skewed to the right.

TABLE 22 Costs related to ESP

Cost	ESP (n = 92)
Cost of steroid injection, mean (SD)	42.96 (31.82)
Cost of physiotherapy, mean (SD)	217.11 (146.85)
Mean (SD) cost – ESP intervention	260.07 (155.07)

TABLE 23 Cost related to complications and additional treatments by trial arm

Cost	ESP (N = 99)		MUA (N = 201)		ACR (N = 203)	
	n (%)	Mean cost (£) (SD)	n (%)	Mean cost (£) (SD)	n (%)	Mean cost (£) (SD)
Randomised patients: cost of additional treatments	2 (2.02)	3.39 (23.75)	2 (1)	1.67 (16.71)	5 (2.47)	3.69 (23.43)
Randomised patients: cost of further treatments	15 (15.1)	89.77 (285.23)	14 (6.96)	53.24 (246.32)	6 (2.95)	6.10 (39.05)
Withdrawals: cost of alternative treatments	2 (2.02)	8.01 (68.17)	8 (3.98)	5.97 (38.43)	9 (4.43)	7.96 (48.48)
Crossovers: cost of other treatments after crossover	2 (2.02)	2.52 (17.67)	0 (0)	0 (0)	0 (0)	0 (0)
Cost of complications (hospital inpatient)	7 (7.07)	9.27 (47.79)	9 (4.47)	42.84 (360.62)	5 (2.46)	34.46 (334.47)
Cost of complications (hospital outpatient)	11 (11.1)	34.09 (112.69)	16 (7.96)	19.26 (83.92)	11 (5.42)	12.30 (60.90)

TABLE 24 Average primary care and community care resource use (shoulder related) and days missed off work per treatment group

Resource type	MUA (N = 201)				ACR (N = 203)				ESP (N = 99)			
	n	Mean (SD)	Median	Missing, n (%)	n	Mean (SD)	Median	Missing, n (%)	n	Mean (SD)	Median	Missing, n (%)
GP surgery total	137	1.61 (3.04)	0	64 (31.8)	138	1.73 (3.23)	0	65 (32.0)	62	0.90 (1.89)	0	37 (37.4)
3 months	168	0.82 (1.64)	0	33 (16.42)	171	1.05 (1.97)	0	32 (15.76)	84	0.58 (1.44)	0	15 (15.15)
6 months	162	0.30 (1.25)	0	39 (19.40)	163	0.49 (1.60)	0	40 (19.70)	76	0.35 (0.89)	0	23 (23.23)
12 months	169	0.34 (1.20)	0	64 (31.84)	162	0.24 (0.76)	0	65 (32.02)	80	0.25 (0.88)	0	37 (37.37)
GP telephone total	136	0.54 (2.05)	0	65 (32.3)	134	0.44 (1.1)	0	69 (33.9)	61	0.10 (0.47)	0	38 (38.4)
3 months	168	0.28 (1.24)	0	3 (16.42)	165	0.32 (0.99)	0	28 (18.72)	82	0.06 (0.33)	0	17 (17.17)
6 months	162	0.16 (1.13)	0	39 (19.40)	161	0.09 (0.41)	0	42 (20.69)	74	0.03 (0.16)	0	25 (25.25)
12 months	168	0.05 (0.17)	0	33 (16.42)	162	0.03 (0.22)	0	41 (20.20)	83	0.01 (0.011)	0	16 (16.16)
Physiotherapist	135	0.83 (2.8)	0	66 (32.8)	136	1.25 (3.8)	0	67 (33.0)	64	1.17 (4.0)	0	35 (35.3)
3 months	167	0.66 (2.26)	0	34 (16.92)	167	0.64 (2.95)	0	36 (17.73)	83	0.42 (1.72)	0	16 (16.16)
6 months	161	0.14 (0.79)	0	40 (19.90)	161	0.31 (1.24)	0	42 (20.69)	77	0.49 (2.25)	0	22 (22.22)
12 months	170	0.71 (0.92)	0	31 (15.42)	162	0.31 (1.32)	0	41 (20.20)	83	0.24 (0.22)	0	16 (16.16)
Nurse surgery	132	0.07 (0.3)	0	69 (34.3)	129	0.39 (0.8)	0	74 (36.4)	59	0.05 (0.3)	0	40 (40.4)
3 months	166	0.2 (0.15)	0	35 (17.41)	165	0.34 (1.09)	0	38 (18.72)	79	0.05 (0.32)	0	20 (20.20)
6 months	160	0.01 (0.08)	0	41 (20.40)	156	0.08 (0.30)	0	47 (23.15)	75	0.04 (0.26)	0	24 (24.24)
12 months	165	0.05 (0.29)	0	36 (17.91)	160	0.02 (0.14)	0	43 (21.18)	79	0 (0)	0	20 (20.20)

continued

TABLE 24 Average primary care and community care resource use (shoulder related) and days missed off work per treatment group (continued)

Resource type	MUA (N = 201)				ACR (N = 203)				ESP (N = 99)			
	n	Mean (SD)	Median	Missing, n (%)	n	Mean (SD)	Median	Missing, n (%)	n	Mean (SD)	Median	Missing, n (%)
Community nurse	135	0 (0)	0	66 (32.8)	136	0.12 (0.9)	0	67 (33.0)	62	0 (0)	0	37 (37.4)
3 months	168	0 (0)	0	33 (16.42)	168	0.07 (0.51)	0	35 (17.24)	83	0 (0)	0	16 (16.16)
6 months	160	0 (0)	0	41 (20.40)	161	0.07 (0.79)	0	42 (20.69)	75	0 (0)	0	24 (24.24)
12 months	170	0.01 (0.15)	0	31 (15.42)	161	0 (0)	0	42 (20.69)	82	0 (0)	0	17 (17.17)
Occupational therapy	137	0.09 (0.7)	0	64 (31.8)	137	0.06 (0.7)	0	66 (32.5)	63	0 (0)	0	36 (36.4)
3 months	168	0.03 (0.46)	0	33 (16.42)	167	0 (0)	0	36 (17.73)	83	0 (0)	0	16 (16.16)
6 months	161	0 (0)	0	40 (19.90)	162	0.01 (0.08)	0	41 (20.20)	76	0 (0)	0	23 (23.23)
12 months	171	0.05 (0.48)	0	32 (15.92)	162	0.05 (0.63)	0	41 (20.20)	82	0 (0)	0	19 (19.19)
Lost days off work	105	17.5 (26.4)	6	96 (47.8)	92	32.8 (44.2)	14	111 (54.)	34	11.5 (27.8)	0	65 (65.6)
3 months	138	12.5 (22.0)	2	63 (31.34)	125	13.3 (23.6)	0	78 (38.42)	61	7.2 (20.6)	0	38 (38.38)
6 months	132	3.5 (10.5)	0	69 (34.32)	125	10.9 (23.2)	0	78 (38.42)	50	5.2 (18.8)	0	49 (49.49)
12 months	138	2.8 (13.3)	0	63 (31.34)	129	3.1 (13.1)	0	74 (36.45)	57	3.9 (13.1)	0	42 (42.42)

Over the entire follow-up period, a higher proportion of participants in the ACR group incurred a loss of earnings as a result of their problems with their shoulders than participants in the other two groups. On average, the number of missed days of work was 11.5 (SD 27.8; median 0; minimum 0, maximum 115) days in the ESP group, 17.5 (SD 26.4; median 6; minimum 0, maximum 120) days in the MUA group and 32.8 (SD 44.2; median 14; minimum 0, maximum 195) days in the ACR group. The difference between the groups is large, and this is reflected in the productivity costs shown in *Table 26*.

Resource use was multiplied by unit costs (*Table 25*) to estimate the economic costs of each resource category. Costs for patients with complete data are in *Table 26*, by trial group and cost category. Over the entire follow-up period, the mean (SE) total NHS and Personal Social Services costs, inclusive of the costs of the allocated index intervention, were £599.06 (£359.23) in the ESP arm, £834.20 (£752.66) in the MUA arm and £2271.09 (£902.50) in the ACR arm.

TABLE 25 Unit costs used for the analysis (£, 2017–18 prices)

Item	Unit cost (£)	Source
Primary and community care		
GP visit at GP practice ^a	37.40	<i>Unit Costs of Health and Social Care 2018</i> ⁶⁹
GP visit at home ^a	93.60	<i>Unit Costs of Health and Social Care 2018</i> ⁶⁹
GP by telephone ^a	15.20	<i>Unit Costs of Health and Social Care 2018</i> ⁶⁹
Nurse visit at GP practice	10.85	<i>Unit Costs of Health and Social Care 2018</i> ⁶⁹
District/community nurse	38.00	<i>Unit Costs of Health and Social Care 2018</i> ⁶⁹
Occupational therapist visit	47.00	<i>Unit Costs of Health and Social Care 2018</i> ⁶⁹
Physiotherapist visit ^b	57.25	<i>NHS Reference Costs 2017 to 2018</i> ⁷⁰
Hospital care		
Inpatient stay (shoulder) ^c	258.00–449.00	<i>NHS Reference Costs 2017 to 2018</i> ⁷⁰
Inpatient stay (non-shoulder)	384.22	<i>NHS Reference Costs 2017 to 2018</i> ⁷⁰
Day-case visit (shoulder) ^c	420.00–2512.00	<i>NHS Reference Costs 2017 to 2018</i> ⁷⁰
Outpatient visit (shoulder)	125.01	<i>NHS Reference Costs 2017 to 2018</i> ⁷⁰
Outpatient visit (non-shoulder)	123.93	<i>NHS Reference Costs 2017 to 2018</i> ⁷⁰
Hospital physiotherapy visit	54.91	<i>NHS Reference Costs 2017 to 2018</i> ⁷⁰
Other health service visit	74.11	<i>NHS Reference Costs 2017 to 2018</i> ⁷⁰
Consultant surgical	108.00	<i>Unit Costs of Health and Social Care 2018</i> ⁶⁹
Associate specialist	105.00	<i>Unit Costs of Health and Social Care 2018</i> ⁶⁹
Specialty registrar	43.00	<i>Unit Costs of Health and Social Care 2018</i> ⁶⁹
Foundation doctor FY1	32.00	<i>Unit Costs of Health and Social Care 2018</i> ⁶⁹
Foundation doctor FY2	28.00	<i>Unit Costs of Health and Social Care 2018</i> ⁶⁹
Physiotherapist band 5	35.00	<i>Unit Costs of Health and Social Care 2018</i> ⁶⁹
Physiotherapist band 6	46.00	<i>Unit Costs of Health and Social Care 2018</i> ⁶⁹
Physiotherapist band 7	55.00	<i>Unit Costs of Health and Social Care 2018</i> ⁶⁹
Physiotherapist band ≥ 8 ^d	72.00	<i>Unit Costs of Health and Social Care 2018</i> ⁶⁹
Nurse band 5	37.00	<i>Unit Costs of Health and Social Care 2018</i> ⁶⁹
Nurse band 6	45.00	<i>Unit Costs of Health and Social Care 2018</i> ⁶⁹
Nurse band 7	54.00	<i>Unit Costs of Health and Social Care 2018</i> ⁶⁹

continued

TABLE 25 Unit costs used for the analysis (£, 2017–18 prices) (continued)

Item	Unit cost (£)	Source
Medications		
Depomedrone 40 mg	3.44	BNF ⁷¹
Depomedrone 80 mg	6.88	BNF ⁷¹
Triamcinolone 40 mg	17.88	BNF ⁷¹
Triamcinolone 80 mg	35.76	BNF ⁷¹
Bupivacaine 0.5% (10 ml)	0.915	BNF ⁷¹
General anaesthesia	30.99	BNF ⁷¹
Antibiotics	6.11	BNF ⁷¹
Private care		
Private non-NHS physiotherapist	50.00	www.capitalphysio.com (accessed 1 November 2020)
Private osteopath	42.50	www.nhs.uk/conditions/osteopathy (accessed 1 November 2020)
Private chiropractitioner	55.00	www.nhs.uk/conditions/chiropractic (accessed 1 November 2020)
Community care service	49.00	Averaged of three above
Private hospital – night	337.00	NHS Reference Costs 2017 to 2018 ⁷⁰

BNF, *British National Formulary*.

a Durations sourced from PSSRU 2015.

b Community health services, physiotherapist, adult, one to one (currency code A08A1).

c Sum of total expenditure on excess bed-days (elective and non-elective) divided by total activity for Healthcare Resource Group codes relating to shoulder: MUA (HD24E; non-inflammatory, bone or joint disorders, with CC score 8–11); ACR (HN53A, HN53B, HN53C, HN54A, HN54B, HN54C; major and intermediate procedures for non-trauma with CC score 4+, 2–3 and 0–1).

d PPP form is featured to record staff at or above band 8. Hence, the unit cost for a physiotherapist at or above band 8 is estimated as the average of 8a (£66) and 8b (£78).

TABLE 26 Costs for patients with complete data by treatment arm and cost category (£, 2017–18 prices)

Costs	Treatment arm, mean (£) (SE)		
	MUA	ACR	ESP
MUA ^a	349.46 (191.91)	5.55 (55.78)	0
ACR ^a	0	1762.32 (934.61)	113.42 (495.67)
ESP ^a	6.90 (59.08)	1.25 (13.02)	260.07 (155.07)
Physiotherapy (hospital)	175.88 (163.90)	174.63 (161.73)	6.98 (36.55)
Physiotherapy (community)	43.80 (146.04)	66 (201.67)	61.87 (211.11)
Further treatments	60.27 (248.06)	17.77 (66.72)	103.70 (290.25)
Hospital inpatient	42.84 (360.62)	34.46 (334.46)	9.27 (47.79)
Hospital outpatient	19.26 (83.92)	12.30 (60.90)	34.09 (112.69)
GP at the surgery	60.33 (113.86)	64.77 (120.87)	33.78 (70.61)
GP over the telephone	8.15 (31.22)	6.69 (16.93)	1.49 (7.18)
Nurse at the surgery	0.74 (3.34)	4.20 (9.13)	0.55 (3.14)
Community nurse	0.00 (0.00)	4.75 (33.81)	0.00 (0.00)
Occupational therapist	4.12 (33.95)	3.09 (32.34)	0.00 (0.00)

TABLE 26 Costs for patients with complete data by treatment arm and cost category (£, 2017–18 prices) (continued)

Costs	Treatment arm, mean (£) (SE)		
	MUA	ACR	ESP
Total costs (NHS)			
Shoulder (a)	834.02 (752.66)	2271.09 (902.5)	599.06 (359.23)
^a Non-shoulder (b)	182.12 (228.98)	195.76 (304.22)	241.82 (366.23)
Productivity costs (c)	1995.29 (2999.85)	3735.61 (5031.35)	1308.70 (3165.177)
Private care costs (d)	31.23 (117.63)	21.40 (111.22)	40.00 (144.51)
Total broader costs (a + b + c + d)	3200.98 (3824.39)	5377.18 (4240.28)	1475.05 (2367.87)

a The cost of the intervention includes the costs of injections.

The total costs estimates shown here are unadjusted means and relate to complete cases; therefore, there is limited value in interpreting the differences between treatments. The mean differences for each surgical treatment compared with ESP, and the corresponding 95% CIs, adjusted for patient covariates and taking into consideration the correlation between costs and QALYs, are shown in *Cost-effectiveness analysis*.

Health-related quality of life and quality-adjusted life-years

A similar decrease in complete EQ-5D-5L questionnaires was seen throughout the trial follow-up across the three arms. Responses for the complete follow-up (i.e. baseline and 3, 6 and 12 months) were available for 369 (73%) participants: 156 (78%) in the MUA arm, 149 (73%) in the ACR arm and 64 (65%) in the ESP arm. The extent of incomplete EQ-5D-5L responses as a result of missing data strengthened the justification for using the imputed data sets as the base case (see *Appendix 16*).

The proportions of participants who reported the EQ-5D-5L levels (1–5) by dimension, group and time point are in *Appendix 16*. Comparing self-care levels between baseline and 12 months, we found that more people reported no problems across all the treatments, but the percentage was slightly higher in the ACR arm. When looking at usual activities, all groups had a similar increase over 12 months in the number of participants who classed themselves as level 1. A smaller percentage of participants were level 1 for pain/discomfort after 12 months in the ESP arm than in the MUA and ACR arms. When looking at anxiety and depression, we found that there was, again, a comparable increase in the percentage of participants in level 1 between baseline and 12 months across all treatments.

The overall distribution of the EQ-5D-5L scores (utilities) for the different follow-up assessments is illustrated in *Figure 12*. Patients allocated to MUA started from a higher utility value than patients allocated to ACR and ESP.

Table 27 summarises the mean EQ-5D-5L scores reported at each follow-up point for all available patients. Adjusted analysis shows that patients allocated to ACR and MUA had similar utility values at 12-month follow-up [ACR (mean 0.739) vs. MUA (mean 0.734)]. Similarly, patients allocated to the surgical arms had better utility values than those allocated to the ESP arm (mean 0.693). The QALY estimates at 1-year follow-up (adjusted for baseline utility) show that patients allocated to MUA accrued more QALYs than those in the other two arms: MUA (0.6765) >ESP (0.6492) >ACR (0.6475).

As for total costs, the HRQoL and QALYs estimates shown in this section are of limited value, as these estimates correspond exclusively to patients with complete EQ-5D-5L data (i.e. at baseline and at 3, 6 and 12 months). The mean differences in QALYs between the groups and corresponding 95% CIs, adjusted for all relevant covariates and taking into consideration the correlation between costs and QALYs, are shown in *Cost-effectiveness analysis*.

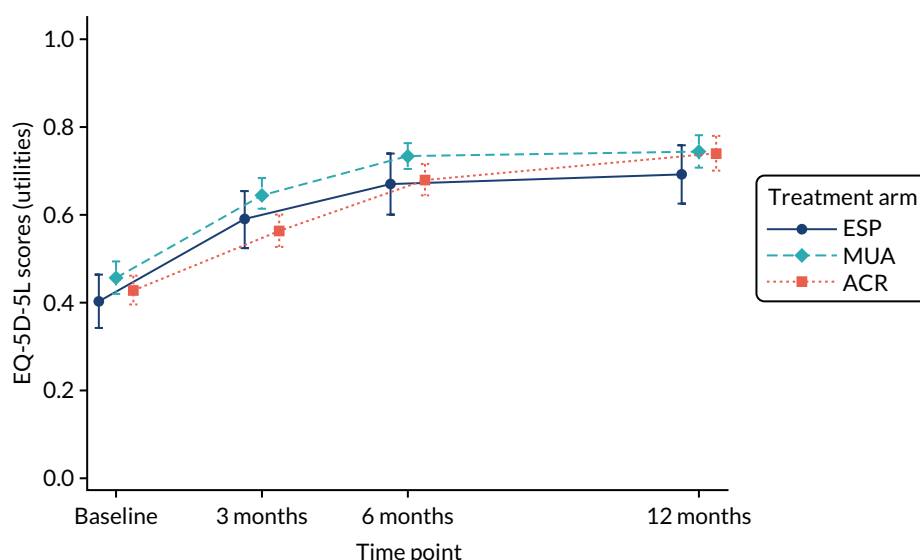


FIGURE 12 The EQ-5D-5L scores distribution at the different time points over the 12 months.

TABLE 27 Health-related quality of life: EQ-5D-5L summary scores (available cases) at each time point adjusted for baseline utility

Follow-up point	Treatment arm, mean (SE)		
	MUA	ACR	ESP
Baseline	0.456 (0.263)	0.428 (0.234)	0.402 (0.294)
3 months	0.632 (0.017)	0.567 (0.017)	0.606 (0.024)
6 months	0.729 (0.016)	0.677 (0.016)	0.680 (0.024)
12 months	0.734 (0.018)	0.739 (0.184)	0.693 (0.027)
QALYs (adjusted utility) (95% CI)	0.6765 (0.651 to 0.702)	0.6475 (0.621 to 0.674)	0.6492 (0.609 to 0.690)

QALY estimates (complete cases) adjusted for baseline utility.

Missing data

The UK FROST study collected data using the EQ-5D-5L at 3, 6 and 12 months. Health-care resource use was elicited from patients by postal questionnaire at 3, 6 and 12 months, and from health-care professionals by hospital forms 52 weeks after randomisation. A description of the economic variables in UK FROST can be found in *Appendix 17*.

Overall, the proportion of participants with complete economic data remained similar between treatment arms (see *Appendix 17*): 46.46% in the ESP arm, 58.21% in the MUA arm and 57.14% in the ACR arm. In all arms, more individuals are observed in month 12 than in month 6. Therefore, the missing data do not follow a monotonic pattern; in other words, there are participants with intermittent missing data (e.g. lost to follow-up at 6 months but remained subsequently). Hence, inverse probability weighting would be inappropriate under such a pattern. Similarly, complete-case analysis would be, as a minimum, inefficient because it would discard observed data from individuals with some missing outcomes.

Figures showing the pattern of missing data are in *Appendix 17*. As discussed above, missing data are shown to be non-monotonic, as individuals with missing data at one follow-up point may provide data subsequently.

Logistic regressions of indicators of missing cost and QALY data on treatment allocation and a selection of baseline variables showed that lower EQ-5D-5L scores at baseline are associated with missing cost and QALY data (see *Appendix 17*). Baseline age was also found to be a significant predictor of missing data on HRQoL. This suggests that the data are unlikely to be missing completely at random. The other baseline covariates (sex and diabetes) were associated with missingness but were not statistically significant at 5%. However, diabetes was significant predictor of costs and QALYs at 6 months and 1 year, which would support both covariate-dependent missing completely at random and MAR assumptions.

We also explored whether or not missingness is associated with previously observed outcomes by regressing indicators of missing costs or QALYs at each year on their previously observed values (e.g. regressing missing costs and QALYs at 1 year on costs and QALYs in previous months). Most regressions produced statistically insignificant results ($p > 0.05$), with two exceptions: missing QALYs at 1 year were significantly associated with QALYs at 3 months, and missing costs at 1 year were significantly associated with QALYs at 3 months and QALYs at 6 months. Although these regressions are likely to be affected by multicollinearity, they provide an indication that data are unlikely to be covariate-dependent missing completely at random.

Therefore, data were assumed to be MAR, and multiple imputation by chained equations was selected to handle missing data for this economic analysis. In the analysis, missingness is assumed to depend on baseline covariates (sex, diabetes, age, EQ-5D-5L at baseline and OSS score at baseline) and observed costs and QALYs but to be independent of unobservable costs at QALYs at 1 year. As it is impossible to know whether data are MNAR or MAR from the observed data, a mixed model is presented as per sensitivity analysis. Complete-case analysis, which is not valid when data are MAR, is presented for comparison only.

The multiple imputation model was validated by comparing the distribution of the observed UK FROST data with the imputed data (see *Appendix 17*).

Cost-effectiveness analysis

Base-case analysis

A bivariate regression, in the form of a seemingly related regression, conducted in the imputed data set was used to estimate the incremental costs and incremental health outcomes (i.e. QALYs) associated with the interventions (*Table 28*). Patients allocated to MUA showed a (non-significant) QALY gain compared with those allocated to ESP (mean difference 0.0396, 95% CI -0.0008 to 0.0800). Similarly, patients allocated to ACR showed a (non-significant) QALY gain compared with those allocated to ESP (mean difference 0.0103, 95% CI -0.0304 to 0.0510). Overall, those allocated to ACR had worse (non-significant) QALYs than those allocated to MUA at the 12-month follow-up (mean difference -0.0293, 95% CI -0.0616 to 0.0030).

The results of the fully incremental cost-effectiveness estimates and probability that each intervention is cost-effective at a threshold of £20,000 per QALY are also shown in *Table 28*. Compared with physiotherapy, MUA cost a mean of £276 more per patient (95% CI £65.67 to £487.35) and allowed patients to experience improved health outcomes at the end of the trial (on average 0.0396 more QALYs per participant than ESP, 95% CI -0.0008 to 0.0800). The resulting ICER for MUA was £6984 per additional QALY. ACR is significantly more costly than ESP [on average £1733.78 more expensive per participant (95% CI £1529.48 to £1938.06)], and, despite the QALY gain accrued by ACR participants (on average 0.0396 more QALYs per participant than ESP, 95% CI -0.0008 to 0.0800), this was not sufficient to prove that ACR is a cost-effective use of NHS resources when compared with ESP (i.e. ICER above recommended NICE threshold). Similarly, ACR is dominated by MUA, with higher mean costs and fewer QALYs.

TABLE 28 Adjusted mean differences in QALYs and costs between interventions (base case)

Treatment arm	Adjusted difference in means with SUR ^a (95% CI)			
Difference in costs (£)				
MUA vs. ESP	276.507 (65.67 to 487.35)			
ACR vs. ESP	1733.78 (1529.48 to 1938.06)			
ACR vs. MUA	1457.26 (1282.73 to 1631.79)			
Difference in QALYs				
MUA vs. ESP	0.0396 (-0.0008 to 0.0800)			
ACR vs. ESP	0.0103 (-0.0304 to 0.0510)			
ACR vs. MUA	-0.0293 (-0.0616 to 0.0030)			
	ICER (£ per QALY)	Probability that intervention is cost-effective at £13,000/QALY	Probability that intervention is cost-effective at £20,000/QALY	Probability that intervention is cost-effective at £30,000/QALY
MUA	6984	0.7942	0.8632	0.8978
ACR	> 100,000	0.0000	0.0002	0.002
ESP	-	0.2058	0.1366	0.1002

SUR, seemingly unrelated regression.

a Compared with ESP, as it is the alternative with lower costs and worse health outcomes.

The corresponding CEACs showing the probability that each treatment is cost-effective across a range of thresholds are shown in *Figure 13*. The probability that MUA surgery is cost-effective is 0.88 at a threshold of £20,000 per QALY gained. The CEAC indicates that, regardless of the value of the cost-effectiveness threshold, the probability that ACR is cost-effective does not exceed 0.002.

Sensitivity analyses and uncertainty

A number of scenario analyses were conducted to test the robustness to alternative assumptions, both related to costs and missing data. As already mentioned, we considered two sensitivity analysis around costs. *Table 29* shows the results of both scenarios that implied recalculating costs: scenario 1 [including non-shoulder costs (ITT)]; and scenario 2 (including productivity costs and private care costs). The ICER for MUA was £10,485 per QALY gained when including (primary care) non-shoulder resource use in the analysis, indicating that MUA would continue being a cost-effective use of NHS resources. By contrast, cost-effectiveness results were sensitive to a wider perspective scenario, suggesting that the ICER from a wider perspective was higher than the thresholds that NICE normally considers for reimbursement decisions. ACR continued to be dominated by MUA in both scenarios.

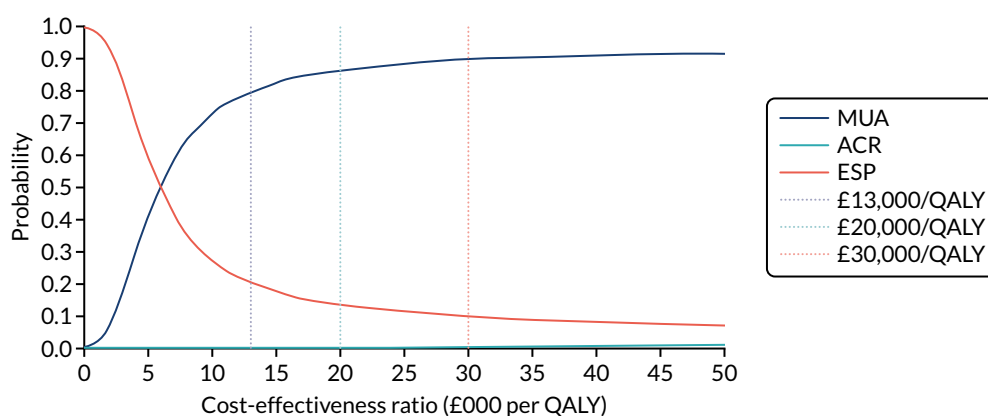


FIGURE 13 Base-case CEACs.

TABLE 29 Sensitivity analysis (scenarios 1 and 2): summary for incremental analysis (ITT), cost-effectiveness results and uncertainty under different cost scenarios

Treatment comparisons	Multiple imputation of costs (shoulder – NHS perspective) and QALYs analysis with SUR: base-case analysis	Multiple imputation of costs (shoulder and non-shoulder – NHS perspective) and QALYs analysis with SUR: scenario 1	Multiple imputation of costs (broader perspective ^a) and QALYs analysis with SUR: scenario 2
MUA vs. ESP			
Difference in costs (£)			
Mean	276.50	162.76	1031.86
SE	107.4462	112.83	595.33
95% CI	65.67 to 487.35	-58.39 to 383.91	-136.92 to 2200.65
Difference in QALYs			
Mean	0.039	0.0375	0.0375
SE	0.0206	0.0207	0.0207
95% CI	-0.001 to 0.080	-0.0032 to 0.0782	-0.0032 to 0.0781
ICER	6984	4336	27,522
Probability that MUA is cost-effective ^b	0.88	0.90	0.36
ACR vs. ESP			
Difference in costs (£)			
Mean	1733.78	1555.48	4109.96
SE	104.147	112.42	647.75
95% CI	1529.48 to 1938.06	1335.14 to 1775.82	2836.20 to 5383.73
Difference in QALYs			
Mean	0.0103	0.0080	0.0081
SE	0.0207555	0.0208	0.0208
95% CI	-0.0304 to 0.0510	-0.0328 to 0.0488	-0.0327 to 0.0488
ICER	168,613	194,895	507,707
Probability that ACR is cost-effective ^b	0.030	0.008	0.000
ACR vs. MUA			
Difference in costs (£)			
Mean	1457.26	1392.72	3078.10
SE	88.90998	91.41	548.27
95% CI	1282.73 to 1631.79	1213.56 to 1571.87	1999.07 to 4157.13
Difference in QALYs			
Mean	-0.0293	-0.0296	-0.0294
SE	0.0164678	0.0165	0.0165
95% CI	-0.0616 to 0.0030	-0.0619 to 0.0028	-0.0618 to 0.0030
ICER	ACR dominated by MUA	ACR dominated by MUA	ACR dominated by MUA
Probability that ACR is cost-effective ^b	0.00	0.00	0.00

a The broader perspective includes NHS costs for the shoulder, and non-shoulder, productivity and private costs.

b The probability of being cost-effective is estimated at a threshold of £20,000 per QALY.

Table 30 shows the results of the sensitivity analyses to test the impact of different methods for handling missing data in results. Given the results of the base-case analyses, sensitivity analyses around missing data were restricted to the comparison of MUA with ESP. The mean difference in costs and QALYs and the ICER changed according to the method. The differences in costs were £339 (95% CI £72 to £606) for complete-case analysis, £193 (95% CI -£14 to £399) for multiple imputation and £256 (95% CI £2 to £509) for the mixed model. The differences in QALYs adjusted for EQ-5D and baseline covariates were 0.016 (95% CI -0.034 to 0.066) for complete-case analysis, 0.036 (95% CI -0.004 to 0.076) for multiple imputation and 0.030 (95% CI -0.014 to 0.073) for the mixed model. The SEs are larger in the complete-case analysis, which reflects the smaller sample size. The mixed model has slightly larger SEs than multiple imputation in both the incremental costs and the QALYs, possibly because of the large number of parameters to estimate compared with the analysis model post multiple imputation. The average incremental costs in the complete-case analysis are higher than those estimated with multiple imputation and the mixed model, suggesting that a bias would be introduced if missing completely at random is assumed. However, both multiple imputation and the mixed model agree that MUA is the cost-effective alternative.

In this situation, sensitivity analyses to determine which departures from MAR can alter the conclusions are useful. Hence, the costs and QALYs were imputed under MAR and then shifted under different scenarios. These scenarios were judged of most interest after discussion with clinical experts. Hence, we considered a number of scenarios in which the costs of MUA and ESP were increased by 10% and 50% in both arms or by treatment arm; the same approach was followed for QALYs (see Appendix 17).

Increasing costs or decreasing QALYs in both patient groups makes little difference to the results. The probability changes considerably only when the QALYs of individuals with missing data who were allocated to MUA are decreased by 50%. Nevertheless, MUA remains the intervention most likely to be cost-effective even if its imputed QALYs are reduced by 10% or if its cost is increased by 50%. The results suggest, therefore, that the positive cost-effectiveness profile of MUA is robust to plausible departures from MAR.

TABLE 30 Sensitivity analyses (scenarios 3, 4 and 5): summary for incremental analysis (ITT), cost-effectiveness results and uncertainty under different missing data assumptions

MUA vs. ESP	Complete-case analysis with SUR	Multiple imputation of costs and utilities followed by SUR	Mixed model with adjustment for covariates
Difference in costs (£)			
Mean	339.3	192.68	255.7
SE	136.2	107.45	129.5
95% CI	72.2 to 606.3	-13.97 to 399.33	1.73 to 509.50
Difference in QALYs			
Mean	0.016	0.0357	0.030
SE	0.026	0.020	0.022
95% CI	-0.034 to 0.066	-0.004 to 0.076	-0.014 to 0.073
ICER (£)	21,443	5395.58	8562
Probability that MUA is cost-effective	0.48	0.89	0.76

Conclusion

This economic analysis has provided robust evidence on whether or not surgical management is cost-effective for the treatment of frozen shoulder. Over the trial period, the base-case analysis with the ITT approach showed that MUA was the intervention most likely to be cost-effective. The resulting ICER for MUA was £6984 per additional QALY when compared with ESP; over common threshold values of a QALY, the probability that MUA was cost-effective was high (> 85% from an NHS perspective). The finding indicates that ACR is dominated by MUA (higher mean costs and fewer QALYs), and ACR showed very low probability of being cost-effective (< 5% from an NHS perspective).

The positive cost-effectiveness profile of MUA is robust to plausible departures from MAR. Similarly, these results were robust to a number of sensitivity analyses, showing that MUA was the intervention most likely to be cost-effective at a threshold of £20,000 per QALY, with probabilities ranging across scenarios from 48% (complete-case analysis) to 99%. The only exemption was when we used the societal perspective to estimate the costs; ACR appeared to be dominated by MUA across all scenarios.

Discussion

The economic evaluation alongside the UK FROST trial was conducted following NICE methodological standards. We implemented a comprehensive strategy to handle missing data in accordance with methodological guidelines, and we used a number of analytical tools to address uncertainty, including sampling and methodological uncertainty. The results of the analyses suggest that MUA is a cost-effective option for the treatment of frozen shoulder in terms of QALYs gained calculated using the EQ-5D-5L. Compared with ESP, MUA cost a mean of £276 more per patient (95% CI £65.67 to £487.35) and allowed patients to experience improved health outcomes at the end of the trial [on average 0.0396 more QALYs per participant than ESP (95% CI -0.0008 to 0.0800)]. The ICER for the ITT approach in the imputed data set was £6984 per additional QALY. The probability that MUA is cost-effective is > 85%, whereas the probability that ESP is cost-effective does not exceed 20%. ACR is significantly more costly than ESP [on average £1733.78 more expensive per participant (95% CI £1529.48 to £1938.06)], and, despite the QALY gain accrued by ACR participants [on average 0.0396 more QALYs per participant than ESP (95% CI -0.0008 to 0.0800)], this was not sufficient to prove that ACR was a cost-effectiveness use of NHS resources when compared with ESP (i.e. the ICER was above the recommended NICE threshold). Despite the ACR arm having fewer additional interventions, ACR was dominated by MUA, with higher mean costs and fewer QALYs. Therefore, the worse outcomes observed in the ACR arm than in the other two arms, along with ACR's higher costs, make this treatment difficult to justify. The results of the base-case analysis remained robust to several sensitivity analyses that assessed the impact of areas of uncertainty around a number of study components.

There are two potential limitations to consider when interpreting these results. The first relates to the issue of missing data. Although the use of hospital forms reduced the number of incomplete data, the presence of missing data was unavoidable. We followed a comprehensive analysis to explore whether or not the MAR assumption is plausible given the actual missing data mechanism of the UK FROST data set. Our analysis showed that MAR is a plausible assumption fitting the UK FROST data set; therefore, multiple imputation was selected to handle missing data for the base-case analysis. Furthermore, the results were robust to alternative assumptions about the pattern of missing data, showing that the positive cost-effectiveness profile of MUA is robust to plausible departures from the MAR assumption. It is, therefore, highly unlikely that such assumptions regarding missing data will change the conclusions of our analysis. The second limitation relates to the duration of the study, which at 12 months might still be considered too short in terms of potential functioning. The clinical results showed that nearly 50% of the patients were only 3 points away from being in perfect health in the OSS, which, in turn, should influence positively the associated quality of life. Consequently, the clinical

trends observed during the trial would also suggest that it is unlikely that any important difference in QALYs would emerge beyond the trial follow-up. It is notable that fewer QALYs were observed in the ACR group than in the other two treatment arms at 3 months, which is consistent with the results of the OSS. Conversely, the ACR group had more QALYs and higher OSS scores at 12 months. Moreover, although MUA had marginally higher estimates of OSS scores than ESP and ACR for the average treatment effect over 12 months, this also applied to those in the MUA arm accruing more QALYs over the duration of the study. The extra additional cost for MUA to maximise QALYs would be considered good value for money to the UK NHS at the NICE threshold of willingness to pay. Regarding costs, we are confident that important costs, including costs of complications, have been captured during the trial, especially for MUA, as the most significant risk of this is fracture and it is very unlikely that this happened beyond follow-up. It should be noted that sensitivity analysis to explore whether the results were sensitive to under-reporting of complications did not change the positive cost-effectiveness results in favour of MUA.

Evidence presented in this analysis relates to interventions conducted in the UK. However, given the pragmatic design of the UK FROST trial, the results are generalisable to other health-care systems when patients are referred to secondary care with frozen shoulder, where the decision is whether to offer surgery relatively early or to continue to control symptoms with physiotherapy.

Chapter 5 Qualitative study

Introduction

Qualitative research is often conducted before, during or after a clinical trial to explore the personal perspectives of trial participants and/or health professionals on, commonly, the trial feasibility, participation, the data collection process and the effects of trial interventions.^{87,88} A comprehensive understanding of the subjective experiences to complement the quantitative evaluation of clinical and process outcomes in a trial will better inform patient-centred care and evidence-based practice.⁸⁸ The use of qualitative research methods in standalone studies or as part of clinical trials is gaining momentum in research on a wide range of musculoskeletal conditions.⁸⁹⁻⁹¹ However, there are very few published qualitative studies of people with frozen shoulder.^{18,92} Currently, no qualitative exploration of trial participants' and health professionals' experiences is available within frozen shoulder trials. Therefore, we conducted a qualitative study embedded within UK FROST²⁷ to provide trial participants' and health professionals' insights to guide clinical decision-making.

The objectives of the qualitative study were to explore (1) trial participants' experience and acceptability of the treatments and taking part in the trial and (2) health professionals' (surgeons and physiotherapists) experience of the treatments they delivered in the trial.

Methods

The UK FROST trial participants, surgeons and physiotherapists who had agreed to be contacted by the study team were invited to participate in the interviews. The trial participants were invited approximately 12 months after randomisation at the time of the primary end point of the trial.⁸ This allowed time for post-surgical recovery and for trial participants to reflect on their experience of the intervention received. Men and women with and without diabetes were included. Surgeons who delivered both surgical interventions and physiotherapists who delivered physiotherapy in all three arms of the trial were invited.

The study information sheet and consent forms were sent to the trial participants, surgeons and physiotherapists by post or e-mail. Non-respondents were sent reminders the second and fourth weeks after the invitation was sent. Once signed consent for participation and audio-recording of interviews had been received, a convenient date and time were arranged for a face-to-face or telephone interview. A physiotherapy researcher (CS) trained in qualitative research methods and not involved in the delivery of UK FROST treatments conducted the interviews. Interviews with the trial participants were semistructured, using open questions about their experience of living with frozen shoulder and the treatments in the trial. An interview schedule (see *Report Supplementary Material 21*) was used that was developed following a literature review and discussions with the research team, people with frozen shoulder, a physiotherapist and a surgeon with expertise in this area. The interview schedule for surgeons and physiotherapists (see *Report Supplementary Material 22*) covered the routine clinical management of frozen shoulder, their experience of treating participants in the trial, their personal treatment preferences, and the barriers to and enablers of positive treatment outcomes for frozen shoulder.

We planned to interview to the point of theoretical saturation⁹³ until no further useful categories emerged. We proposed to interview up to 45 trial participants and 15 health-care professionals.

The data were analysed in two ways:

1. The interviews were analysed using constant comparative methods.^{94,95} This involves comparing similarities and differences and developing themes with a shared essence. The data were coded and categorised into themes by CS. Another qualitative researcher (FT) reviewed these themes, and the two researchers discussed and reached an agreement. CS used NVivo 11 qualitative data software (QSR International, Warrington, UK) to organise the analysis.
2. The World Health Organization's *International Classification of Functioning, Disability and Health* (ICF)^{96,97} is a biopsychosocial framework used to conceptualise functioning and disability as a dynamic interaction between the following components: (1) body functions (denoted as 'b') and structures (denoted as 's'), (2) activities and participation (denoted as 'd'), and (3) contextual factors (environmental factors denoted as 'e'). Each component is arranged in hierarchal domains and has up to four levels of categories coded with the alphanumeric system. The first letter of the coding refers to the component followed by the first-level category (ICF chapter number designated for each component). For example, in d5, 'd' denotes the activities and participation component and '5' denotes the chapter on 'self-care'. A second-level category, d510, depicts a self-care problem, 'washing oneself'. Third- and fourth-level ICF categories are also available for some components. For example, b2801 denotes 'pain in body part' and b28013 denotes 'pain in the back'. A specific set of ICF categories that relate to common functional problems for different health conditions is available.

We aimed to map the problems reported by the UK FROST trial participants with a reference of second-level ICF categories (19 in body functions and structures, 34 in activities and participation, and 8 in environmental factors) identified in a previous study on chronic shoulder conditions, including frozen shoulder.⁹⁸

Results

Sixty interviews (trial participants, $n = 44$; surgeons, $n = 8$; physiotherapists, $n = 8$) were completed between August 2016 and January 2018. All interviews with the trial participants were conducted over the telephone. The majority (75%) of the interviews with surgeons and physiotherapists were telephone based and a few were face to face (surgeons, $n = 2$; physiotherapists, $n = 2$).

Interviews with trial participants

The flow diagram of interviews with the trial participants is presented in *Figure 14*. The characteristics of the trial participants are presented in *Table 31*. This includes participants who were allocated to ESP, MUA or ACR. All participants who were interviewed received their allocated treatment, except one participant allocated to MUA and another allocated to ACR, both of whom received ESP.

The five themes of the trial participants' experiences are described in the following sections. There was nothing to indicate that the UK FROST themes found for men and women, and those found for people with and people without diabetes, were different.

Living with frozen shoulder

Trial participants described how frozen shoulder had a major impact on all areas of their life. They perceived it as a combination of painful, restrictive and disabling. Many were not able to identify what had caused their shoulder problem. Some reported a previous shoulder injury and others attributed the problem to tasks such as gardening and lifting heavy weights:

How would I describe it? Well it is one hell of a pain. You are restricted of movement. You cannot move it like say, very far out from your body. You try and lift something and even when you are grabbing something, you can feel the pain in your shoulder. Like say if you were trying to lift up a cup of tea, you could not. I mean it is very awkward.

61 years, female (ESP)

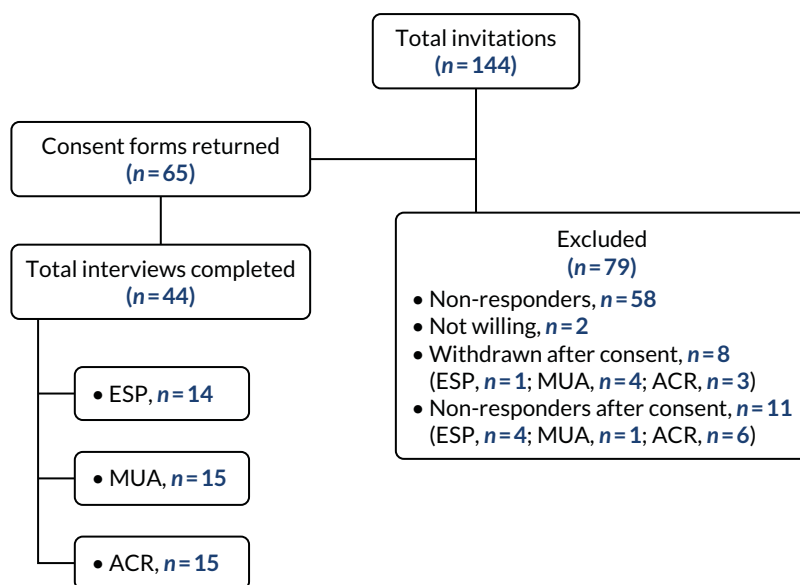


FIGURE 14 Flow diagram of the trial participant interviews.

TABLE 31 Characteristics of the trial participants interviewed

Treatment arm	Sex (n)		Age in years, median (IQR)	Diabetes status (n)	
	Male	Female		Diabetic	Non-diabetic
ESP	5	9	58 (51–63.5)	3	11
MUA	8	7	55 (53–57.5)	5	10
ACR	7	8	59 (53–69)	5	10

IQR, interquartile range.

It was a dramatic impact on my life. I felt I could hardly use my left arm at all, so it was restricting everything that I did.

53 years, female (ACR)

Trial participants had mixed experiences of the onset of symptoms. Some had a sudden sharp pain or a constant dull ache that gradually progressed to reduced movements and function. During the course of frozen shoulder, they experienced pain and movement impairments, sleep disturbances, limitations in day-to-day activities, and restrictions on participating in leisure, work and social activities. Tables 32–34 present the narratives of the trial participants mapped to 19 ICF categories (five in body functions and structures, 12 in activities and participation, and two in environmental factors) from a previous study.¹⁵

Participants described putting off making a GP appointment until their symptoms had worsened, and some described delays in receiving NHS care. Participants thought, in retrospect, that having a quicker NHS care pathway in terms of diagnosis and further specialist referrals was important. Participants also emphasised seeking early medical help and referrals from their GP:

I would say go straight to your doctor and get them to refer you. That would be the first thing because it doesn't just seem to go away of its own accord which I possibly thought initially and that's why I delayed in going to the doctor's in the first place.

64 years, female (ESP)

TABLE 32 Participant-reported problems mapped to ICF categories: body functions and structures

Trial participants' quotations	ICF category
<i>It was like a stabbing pain; it was very severe</i> 73 years, male (ACR)	b280 sensation of pain
<i>I couldn't move my shoulder at all. So it was stuck to my side</i> 59 years, female (ESP)	b710 mobility of joint functions
<i>Definitely lost my strength in my arm</i> 66 years, male (MUA)	b730 muscle power functions
<i>I didn't sleep at all</i> 48 years, female (ESP)	b134 sleep functions
<i>I was generally getting very depressed. I would have happily, towards the end, I would happily have them amputate the arm</i> 50 years, male (MUA)	b152 emotional functions

TABLE 33 Participant-reported problems mapped to ICF categories: activities and participation

Trial participants' quotations	ICF category
<i>Carrying was probably a bit of a problem because I couldn't move that arm so well</i> 64 years, female (ESP)	d430 lifting and carrying objects
<i>Reaching things from the tall shelves in the kitchen, reaching stuff out of the top of the wardrobe ... And it's things you take for granted really</i> 57 years, male (MUA)	d445 hand and arm use
<i>I couldn't drive my car even. I couldn't change gear</i> 55 years, male (ESP)	d475 driving
<i>I couldn't lift my hand up, my arm up, you know, to wash my hair or anything in the bath. And eventually I couldn't get into the bath properly</i> 76 years, male (ESP)	d510 washing oneself
<i>It was the pain and the stiffness, particularly the stiffness, and the inability to address and attend to myself at the toilet. It was becoming a personal hygiene issue for me</i> 53 years, male (ACR)	d530 toileting and d520 caring for body parts
<i>I couldn't get myself dressed. I couldn't get my tops on above my head. I had to wear slack things, so I could get my clothes on</i> 58 years, female (ACR)	d540 dressing
<i>I couldn't open anything, I couldn't use a tin opener, so I couldn't cook neither, I couldn't do anything</i> 55 years, female (ACR)	d630 preparing meals
<i>Doing housework was nigh impossible</i> 54 years, female (MUA)	d640 doing housework
<i>I'm a coachbuilder by trade ... I couldn't work overhead, I couldn't lift my arms up and I couldn't stretch my arms out, I just couldn't do it, so I wasn't, I didn't go to work, I was off for about 7 months</i> 55 years, male (MUA)	d850 employment
<i>Well I decorate cakes. I do novelty cakes, all kinds of decorating. That literally stopped</i> 64 years, female (ACR)	d920 recreation and leisure
<i>I like gardening and I was doing the decorating on my house, all sorts of jobs like that, sport, tennis, anything like that it sort of ruined everything really</i> 66 years, male (MUA)	d920 recreation and leisure
<i>I couldn't sleep on my left side any more and if I turned over in the night and tried to sleep on that side it instantly woke me up</i> 57 years, male (MUA)	d410 changing basic body position

TABLE 34 Participant-reported problems mapped to ICF categories: environmental factors

Trial participants' quotations	ICF category
<p><i>... my wife had to help me put my socks on, things like that, get in and out of the shower, she had to do all the gardening, shopping, things like that. It was just really, really sore</i></p> <p style="text-align: right;"><i>72 years, male (ESP)</i></p>	e310 and e320 immediate family and friends; facilitators
<p><i>I had a lot of support from my work, but the household work, I couldn't manage because my husband had to do that for me</i></p> <p style="text-align: right;"><i>53 years, female (MUA)</i></p>	e310 and e320 immediate family and friends; facilitators

Just speed it up. It was, I think I went to see the GP at the beginning of November, and it wasn't really till the following January before I got any kind of treatment other than pain relief, by which time I'd lost all movement.

50 years, male (MUA)

Some participants reported that they had had a range of treatments before the trial, such as painkillers, physiotherapy, acupuncture and steroid injections, whereas others had had no treatment at all. Participants felt that painkillers and injections used before the trial had not helped them. Similarly, participants found that physiotherapy treatments had not helped to increase their range of movement, partly because of the difficulty that they had exercising as a result of the pain:

I had a couple of sessions of physio before I was referred to [consultant's name]. And I will be honest with you; the physio basically said there was not much they could do for me at the time.

59 years, male (MUA)

Participants were concerned that they were stuck with the disability from frozen shoulder and were eager to get it sorted. This was their main motivation for taking part in UK FROST:

Just the fact that my life would seem to be on hold because I couldn't function properly, you know that was my main concern, I didn't want to be left like this permanently, I wanted something done about it, I didn't want to be continually taking painkillers, which I seem to be living on just to ease the pain, and I didn't want to be doing that. I thought, 'I need to get something done', that was my main concern.

59 years, female (ACR)

Improvements in outcomes and participant satisfaction following the trial

Trial participants considered pain relief and the return of shoulder movement and function to be important treatment outcomes:

Going back to normal ... When you had nil pain and full flexibility and movement within your shoulder. No sleepless nights and that.

62 years, female (ESP)

Trial participants said that they had experienced significant pain relief after their treatment. Participants in the ESP arm said that the steroid injections had reduced their pain and allowed them to start physiotherapy:

When I went to the surgeon I was injected into my shoulder and the pain down my arm that more or less went straight away. After that I went on that course for frozen shoulder, for therapy, and I went for 12 weeks running once a week. It seemed to go, and it's been fine since.

76 years, male (ESP)

So at the beginning I said the pain was 10 and now after all my physios, I'd say it was, I'd say it was about 2 now.

56 years, female (MUA)

I mean the pain in the beginning was just horrendous, it was really, really sore, really painful but after I'd had the physiotherapy, it was . . . I've got no pain at all now.

55 years, female (ACR)

Trial participants in all treatment arms reported increased shoulder movement:

Virtually full movement. My shoulder is fine as far as movement is concerned.

68 years, male (ESP)

Basically had all my full movement back.

58 years, male (MUA)

I got my life back again. I can walk my dogs. I can hold the dog leads. I can do my shopping. I can carry things again.

58 years, female (ACR)

Trial participants described how the physiotherapy sessions (ESP and PPP) had helped to improve their shoulder movement:

I could tell initially straightaway that my movement was starting to come, within a few days I could tell a difference of doing the exercises and as the weeks went on, it was just got better and better and by the time the 12 weeks was up, I virtually had full movements with no pain or anything, it was brilliant!

64 years, female (ESP)

After a few days I was doing my exercises and I was quite surprised already how much movement I had back and then it was regular physio appointments up at the hospital just to keep moving things around and that went really well . . . the physiotherapy was actually really, really beneficial.

45 years, female (MUA)

I felt that the physiotherapy I received was marvellous and improved the range of movements or showed me how to keep that range of movements much quicker than they did on the right-hand side, so I felt that everything went along fine, and I've got no complaints at all, none.

53 years, female (ACR)

Participants in the ACR group felt that their recovery in terms of pain and movement had been quicker than they had expected. Some had experienced improvements as early as after 1 or 2 weeks of physiotherapy after surgery:

It is almost like you have had a quick fix to fix your shoulder then you move on and I think personally for me because the surgery went very well and almost after a couple of weeks I was back to normal.

44 years, male (ACR)

Participants in all treatment arms said that their ability to do routine activities had improved:

I can lift my arm above my head now, you know? I can carry stuff, and I can lift it above my waist, and I can actually go swimming, you know? I can swim now.

54 years, male (ESP)

My little everyday things have come back; I have come back, yes.

52 years, female (MUA)

I can do everything – there's nothing that I can't do; I can wash my back, I can put my bra on, fasten it at the back, I can fasten my skirt at the side and the back now, there's nothing I can't do before I had the frozen shoulder everything I could do then I can now do again.

59 years, female (ACR)

In spite of achieving pain relief and improved function, participants experienced mild and occasional pain and restrictions during certain end-range activities:

I do get occasional pains in my arm, but it's very mild and yeah, I'm aware that I still don't have full movement in my shoulder, but it is much better than it was.

45 years, female (ESP)

There's still a wee bit of pain there, but it's nothing. You see, I'm not concerned about it.

55 years, male (MUA)

I still get twinges now and again but it's nothing, and that's only when I try to put my arm right around my back.

73 years, male (ACR)

Trial participants were satisfied with the UK FROST treatments they had received:

I'm absolutely delighted with the treatment that I was given. I feel as though it did everything that I wanted it to do and expected it to do.

64 years, female (ESP)

I would say I'm like, 100% happy with the treatment, and the study was, like, 100%, it's good, I didn't mind it.

56 years, female (MUA)

Very satisfied. I have no complaints at all.

61 years, male (ACR)

However, two participants were not satisfied with ESP. One had been treated by a private physiotherapist before the trial and did not improve after physiotherapy during the trial. The other was not pleased that the exercise sessions had been supervised by an unfamiliar physiotherapist:

Waste of time . . . I wouldn't be recommending it to a friend . . . Well because by the time I got to see the NHS physio, the private physio had already, if you like, done the hard work and they just had to, if you like, pick up the pieces and keep it OK, and they didn't.

50 years, female (ESP)

May I say 50%? That is mainly because when I saw the physio one to one, I was 100% happy and then when I went to the gym and the physiotherapist sort of left you to your own devices, they didn't really know who I was, why I was there, and they certainly didn't. It was basically just like going to any old gym and being supervised by someone who didn't know you from Adam. It was very, very disappointing.

59 years, female (ESP)

Trial participants' adherence to home exercises

Participants found that the exercises were difficult to begin with but eased off in subsequent sessions:

It [exercise] was difficult and painful. But I could tell week on week the pain was reducing, and my movement was increasing. So it was obviously working quite well.

57 years, male (MUA)

However, they were aware of the benefits of exercise and persevered with doing their home exercises regularly:

I did persevere, and I was doing what I was told, which was obviously you've got to have the pain to get back to normal.

48 years, female (ESP)

Most participants said that they had not continued their home exercises after the trial because they felt that they had regained their normal shoulder function. They shifted from doing the structured home exercise regime to daily functional activities to keep their shoulder mobile. A few participants did some shoulder stretches occasionally:

I'm working with my shoulder all the time so I'm not doing the exercises that the hospital gave me, because I'm working my, I'm swimming, I'm doing . . . I go on long walks, I take the dog out and what have you, so I'm using my arm.

54 years, male (ESP)

Trial participants' treatment preferences before and after the trial

Participants had various treatment preferences before the trial. Some of the participants allocated to MUA and ACR felt that physiotherapy would be ineffective because it had not worked for them previously or they felt that physiotherapy would be difficult with their painful shoulder. MUA was perceived as less invasive and ACR was perceived as an effective treatment. A few preferred physiotherapy so that they could avoid the risks of surgery. Some did not have a particular preference at all. Three participants with diabetes felt that it would take longer to recover after ACR because of their diabetes:

I'm not really too sure why I wouldn't choose physiotherapy. I just think surgery seems a more final option. Physiotherapy, it might work, it may help, it may not. But to me, if I was given surgery, the surgery would work. I had more faith in the surgery working than the physiotherapy itself.

53 years, male (ACR)

Well the first option would have been more intense physiotherapy which I didn't find would have been successful because I'd already had physiotherapy. The second option was to have been invasive surgery which means cutting open, an actual surgical procedure which I wasn't that keen on to be honest with you.

55 years, female (MUA)

Well I didn't want to go to surgery or anything like that, so I just had the needle in my shoulder. I don't think I'd have wanted to go surgery at my age.

76 years, male (ESP)

Just I didn't think, as I said, I didn't know if any was better than the other, I just went along with what was there.

55 years, male (MUA)

Despite preferences at the outset, at the end of the trial there was a sense from participants that they would choose the same treatment that they had been allocated, particularly those who had received ACR:

Oh no I would have the same treatment. As I say I was only there for 10 weeks and I mean, on my 11th week I was still hell of a lot better.

61 years, female (ESP)

Probably the same, because I had to get it moving to start with, it just felt as though it was never going to move. So again, the manipulation, it kind of kick-started it and got it moving, because I honestly thought it was never, ever going to move.

53 years, female (MUA)

I think if it happened to me again, I would be looking to be referred for keyhole surgery again. I think it was an excellent course of treatment and if it had to happen again, that's the treatment I would want.

53 years, male (ACR)

A few participants in the ESP and MUA arms wanted to choose ACR for a permanent and quicker solution for their shoulder problem:

If I had a recurrence of the frozen shoulder in the same joint, I would obviously look for alternative treatment for the simple reason because obviously that treatment, although it alleviated the symptoms, hasn't completely got rid of the symptoms then because if it recurs. So you would look for a permanent solution . . . I would think if it came back again, I would prefer to have the keyhole surgery, yes.

68 years, male (ESP)

Trial participants' experience of participating in the trial

Trial participants had altruistic and personal reasons for participating in the trial. They desired to help other people and contribute to research, and some expected to have their shoulder problem treated quickly:

Because I'm all in favour, if you can do something to help other people not go through the misery that you've been through, and gone through, then I would do it.

62 years, female (ESP)

That I would be seen to sooner than if I didn't do it. And I would have the opportunity to have any treatments much sooner than being on the waiting list. It's not a very nice reason for you to hear, but that's what I did it for.

53 years, female (ACR)

Trial participants found the trial questionnaires relevant and simple to complete. A few felt that the questions were lengthy and repetitive. Some had difficulties answering the questions about comparing their health status with that of the previous month:

It was quite easy, they were simple questions, and I just sort of flew through the questionnaire with no problem.

64 years, female (ESP)

Very repetitive [laughter]. Very repetitive. It went on and on and on. Just when you thought you'd finished there would be another page added on. They are a bit of a pain really.

57 years, male (MUA)

Trial participants said that the physiotherapists who delivered the physiotherapy sessions (ESP and PPP) were supportive and helpful:

My physiotherapist was very nice and he didn't push me to do anything that wasn't in my ability and yeah, we just took it at a nice, steady pace.

41 years, female (ESP)

The physiotherapist really knew what she was doing and straight away assessed exactly where I was at and what I needed to be doing, and that worked really well.

50 years, male (MUA)

They also liked seeing the same physiotherapist so that good connections and rapport could be developed throughout the programme:

I think seeing the same person is always helpful because otherwise it must be time saving as well because you haven't got to read up on the case every time and you get a good rapport like that.

53 years, male (MUA)

Participants in the MUA and ACR arms said that the surgical procedures were explained and went well:

The day in the surgery, I was told everything, how long I would be off work for, and I had targeted physiotherapy immediately after, I knew that was part of it. I attended on the day of my operation; everything went very, very smoothly.

58 years, female (MUA)

It was just a case of just sitting around, reading books, talking with other people, some people had been there two to three times and had various operations, everyone was just chatting and making everyone feel at ease with you know their own experiences . . . And then it was a case of get in, swabbed up and on the trolley through to the operating theatre, operation I believe obviously went fantastic, went really well. It was just a case of waking up, recovery.

44 years, male (ACR)

Trial participants felt that their treatment packages were well co-ordinated and did not require major modifications. Two MUA participants with diabetes suggested providing more information on the effects of pain block and steroid injections on blood sugar levels. A few participants in the ESP and ACR arms felt that the exercises were time-consuming:

Well personally I found it quite difficult to make sure I had them that many times in the day. I'm not too sure how somebody working or having a family could actually manage to fit it in because as I say, by the time going towards the middle to the end of the programme . . .

64 years, female (ESP)

Yeah that was quite difficult because it takes up quite a bit of time and it is quite tiring. I'm not the fittest person and I did find it quite tiring to do all the exercises required.

53 years, male (ACR)

Interviews with surgeons and physiotherapists

The flow diagram of the surgeons' and physiotherapists' interviews is presented in *Figure 15*. The characteristics of surgeons and physiotherapists are presented in *Table 35*.

The four themes from the interviews with surgeons and physiotherapists are described in the following sections.

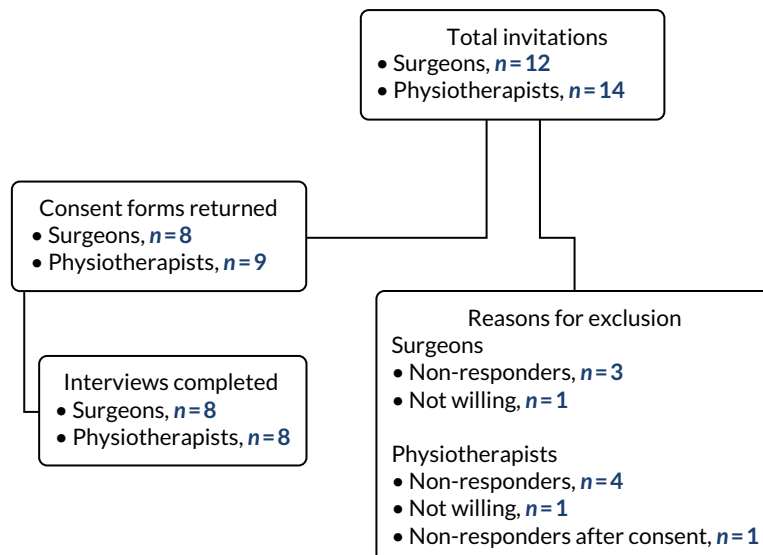


FIGURE 15 Flow diagram of surgeons' and physiotherapists' interviews.

TABLE 35 Characteristics of the surgeons and physiotherapists interviewed

Health professional	Sex		Median (IQR) years of experience in treating shoulder conditions
	Male	Female	
Surgeons	7	1	11 (7–16.25)
Physiotherapists	0	8	13.5 (12–15.75)

IQR, interquartile range.

A stage-based approach in routine treatment of frozen shoulder

Surgeons and physiotherapists described a stage-based treatment approach (from conservative to surgical interventions depending on the severity and duration of symptoms) in their routine practice. During the early painful phase, surgeons and physiotherapists thought that pain control with steroids would be the priority so that exercises could be initiated. Physiotherapists also provided patients with education as part of their treatment plan:

A humeral steroid injection is beneficial for pain relief. It allows them to do the physiotherapy. So, I do use that a lot.

Surgeon 6

I would always tend to go down the glenohumeral joint injection first to see if that settled things down and to give me a window where I could then push them with regards to their exercises.

Physiotherapist 1

And then it's education; tell them all about the condition, what we know about the condition, let them know that it is going to get better over time, and it may not resolve fully and so on. So we'll basically educate them, reassure them.

Physiotherapist 5

Surgeons felt that exercises during the early, painful, phase may aggravate the pain but that exercises after surgery are important for recovery:

Physiotherapy in isolation without an injection I've found most patients present having had really pain that's not manageable because they're trying to stretch and rehab a painful shoulder from frozen shoulder just increases their pain.

Surgeon 2

We do believe in physiotherapy afterwards. One is to have specialised physiotherapists who can allay their fears and talk about their fitness, talk about their recovery, talk about the time scales and the pain. That's one compliment of the physiotherapies is some health professionals who can talk to them on a regular basis.

Surgeon 3

During the stiffness-dominant and postoperative phases, physiotherapists focused on improving shoulder movements, strength and function. They prescribed an intensive exercise regimen that included vigorous shoulder stretches, joint manipulation/mobilisation and home or gym-based exercises:

I know that pain is not a problem then I can push them, I can do the more vigorous stretches, I can manipulate their joints or mobilise their joints and I put them on a more gym-based or exercise-based programme . . . I will go and more push them towards reaching their more functional goals and more towards achieving their return to work and that kind of thing task.

Physiotherapist 4

For people with severe or longstanding symptoms and those resistant to conservative treatments, both surgeons and physiotherapists considered surgery as the final treatment option. For most surgeons, ACR was the default procedure as it resulted in faster recovery and low risk of humeral fractures, and they were familiar with the procedure:

So, if a shoulder is extremely stiff, tight and quite painful I would probably avoid doing a manipulation on these patients and my preference would be ACR if the longevity of the stiffness and the symptoms is quite prolonged again my preference would be an ACR.

Surgeon 6

Generally, the people that don't respond to physio often will end up seeing a surgeon at some point. So the tricky more longer-term frozen shoulder patients are the ones with diabetes that tend to have more problems and end up requiring intervention.

Physiotherapist 2

Physiotherapists felt that the UK FROST physiotherapy programmes and the exercise booklet gave them flexibility to choose exercises that were compatible with their routine practice. A few suggestions were made, for example structuring the treatment components to suit the stages of frozen shoulder, spreading out the 12 weekly sessions over 6 months and having group sessions:

I think the interventions that were on the booklet were what I would use generally. There was always an option there for me to tick off what I would do so I was in agreement with the options that were there and in agreement with the options that they actually, didn't want you to use.

Physiotherapist 2

I think group sessions would be really useful because patients get a lot from each other, and having experience group sessions with other clients with different pathologies, you know, they find that really reassuring . . .

Physiotherapist 5

Physiotherapists also commented on the feasibility of the UK FROST physiotherapy programmes in the NHS, and there was a sense that it would be difficult to deliver the number of UK FROST physiotherapy sessions in their routine practice:

We don't normally get the luxury of being able to see patients as often as the FROST trial was letting us. I think it was 12 treatments we could have overall.

Physiotherapist 2

... they [trial participants] were seen with the 24 hours post surgery and they had 12 sessions which is a luxury because in our trust, that is never, not going to happen and that never used to happen.

Physiotherapist 4

In general practice, well if somebody is improving and they're self-managing, then we don't need to be seeing these patients every week and our service would not allow us to be able to see them every week.

Physiotherapist 7

Treatment expectations and preferences

Surgeons and physiotherapists had mixed treatment expectations. Although surgeons said that they maintained equipoise, they also said that ESP would not be as effective as surgery:

They usually come to me and say, 'What would you do doctor?' I just be honest with them that so far, I have done this, and my results are reasonable, but I wouldn't say that this is the only answer. There are other answers that are equally valid and have to be tested ... It is an ethical thing to do because all three of the treatments are valid and accepted treatment for frozen shoulder.

Surgeon 3

My expectation is that physio won't work. My expectation is that the other two are probably equivocal.

Surgeon 7

Some surgeons and physiotherapists expected similar outcomes across treatment arms. A few physiotherapists felt that the surgical arms, especially the ACR arm, would perform better than the ESP arm. Although some considered MUA an outdated intervention with a risk of injury, some felt that it would be comparable with ACR. Some physiotherapists mentioned that post-surgical soreness is common with ACR:

I expected my patients to get better on whichever arm they chose.

Surgeon 3

I'm expecting the MUAs to be surprisingly better than I would expect. I think the arthrolysis do great anyway and the physio is an unfair one, because if we're seeing them over such a long time the natural history of the frozen shoulder is it will get better. Going against my own profession here; is it the physio that made it better or is it just time.

Physiotherapist 6

I've had experience of patients doing well in them in them all and equally patients not doing well in them all. I thought they were; I don't think they were randomised equally. I thought there were more randomised to surgery. It was more for surgery than physio. So no, I didn't have any predisposed feelings about it at all.

Physiotherapist 8

Surgeons and physiotherapists would have preferred that hydrodilatation had been one of the UK FROST trial arms. Hydrodilatation was described as an easy-to-use, less invasive and inexpensive procedure and an alternative for reducing NHS waiting lists for surgery:

I'd definitely have a hydrodilatation group because part of your trial is trying to work out if the cheaper operation is better than the more expensive operation and hydrodilations probably gained quite popularity since we started the trial design and it reflects current practice.

Surgeon 5

It needs hydrodilatation in it. I personally think it gets really good results on a big bulk of patients and it's a wasted opportunity to have done this study and not have that as one of the arms.

Physiotherapist 6

Factors that influence treatment outcomes

Similar to trial participants, surgeons and physiotherapists perceived that pain relief and improved movements and function were important outcomes. Pain relief was the priority outcome for physiotherapists:

Better in a sentence? That their pain is resolved, and their movements improve, and they return to full function within their daily living, both at work and for recreation and home. That's what I expected to achieve in any of the arm of the trial.

Surgeon 3

So yeah, pain relief first and foremost and then everything else after that; increasing their range of movement, function, return to activities, whether it be sport or work or hobbies and so on. But yeah, definitely pain is the number one.

Physiotherapist 5

I think pain is always the predominant thing with these patients. They just want someone to do something to help with the pain.

Physiotherapist 7

Surgeons and physiotherapists felt that diabetes negatively affects treatment outcomes:

Well my experience with diabetics they have been very bad for a hell of a long time, OK, and that's again maybe just my experience, I haven't measured that experience, but I've seen plenty of them of male diabetics that say, 'Well, I've been stiff for three years', or two-and-a-half years, that's not uncommon.

Surgeon 8

I think my understanding is diabetic patients are slightly more prone to developing a frozen shoulder. It seems to be potentially more complications, slightly more resistant to treatment. And the chances of them maybe developing it again are slightly higher, I think.

Physiotherapist 2

They also felt that participants engaging with treatments and having positive expectations leads to better outcomes:

With any treatment they do they have to engage, and that is something that I emphasise to my patients. So I spent a lot of time in the initial consultations giving them the knowledge. So what makes them better, it's the patient themselves and their knowledge.

Surgeon 3

I think the expectations and belief is probably the most noticeable factor that affects people's outcome; if they believe something is the right thing, the best thing for them, they seem to do well.

Physiotherapist 5

Perceptions about trial participants' experience

Surgeons and physiotherapists felt that the trial participants were happy to be involved in UK FROST. At the same time, participants came to the trial with fixed ideas about the treatment they wanted:

Once they had consented to be part of the study, and they had no problems because they were equally welcome on what treatment they would get . . . all those who once we enrolled them on to the study, were OK with that.

Surgeon 6

So, we saw a lot of frozen shoulders coming in, but a lot of them had fixed ideas of what treatment they wanted. They didn't want surgery yet, or they didn't want to take time off work was the other one, but less so. The standout one was they didn't want an operation, or they wanted to try physiotherapy and injection and then they would opt for surgery. They wanted it to be continuum like that, not a one or the other.

Physiotherapist 8

Surgeons and physiotherapists said that some people declined to take part in the trial because their previous physiotherapy had not worked and so they did not want to be randomised to receive ESP:

Many of them they say, 'Look, I would love to contribute to the greater good and be involved in clinical trials, but I've come to the point that I will not consent for physiotherapy if I was randomised to that'.

Surgeon 8

Like I say, if patients have already had physiotherapy, some of those patients have not wanted to be recruited at risk of repeating what's already not worked. So that was quite awkward to do.

Physiotherapist 6

Physiotherapists said that the UK FROST interventions were well received. Surgeons and physiotherapists said that participants were surprised with the number of PPP sessions they received during the trial. Surgeons and physiotherapists described how a few participants randomised to ESP felt that they were not improving:

They seem to be quite happy with the intervention [ESP] generally. If they had had an injection they were happier because their pain level was better and they were able to tolerate the exercises a bit better.

Physiotherapist 2

I think most of them were with my experience, the patients were really, really very happy, the ones who went to the manipulation as well as arthroscopy capsular release.

Physiotherapist 4

Discussion

Principal findings

An embedded qualitative interview study was conducted within UK FROST to explore the experiences of the trial participants and health professionals (surgeons and physiotherapists) of the trial interventions. The key findings were as follows.

Trial participants described that frozen shoulder had a major impact on all aspects of their life. They were keen to get their shoulder problem sorted, which motivated them to participate in the trial. They also insisted on seeking early medical help and a quicker NHS care pathway. In general, trial participants were satisfied with the UK FROST interventions and found them acceptable. They reported improvements in pain, movement and function. Participants who had received ACR described quicker recovery than they had expected.

Surgeons and physiotherapists followed a stage-based treatment approach in their routine practice. Both felt that people with diabetes tend to have poorer outcomes. They suggested that hydrodilatation could have been a treatment arm of the trial. Both described that some people who had received previously ineffective physiotherapy did not want to take part in the trial.

The common perceptions among trial participants, surgeons and physiotherapists were that (1) trial participants were happy to be part of UK FROST; (2) pain relief and regaining shoulder movements and function are important outcomes; (3) steroids help pain relief and to initiate shoulder exercises; (4) a progressive physiotherapy programme would improve shoulder movements and function; (5) adherence to prescribed exercises is important for better outcomes; and (6) all had their personal preferences among UK FROST treatments.

Frozen shoulder has a negative impact on all areas of life

Although frozen shoulder has a self-resolving natural history, our findings indicate that it is a painful and debilitating condition causing a considerable level of disability and reduced quality of life. This resonates with the results of previously published studies on this topic.^{18,99,100} The problems due to frozen shoulder, as described in our participants' interviews, were mapped to the ICF biopsychosocial framework of disability.^{96,97} This is the first time, to our knowledge, that the ICF has been used specifically to describe functioning and disability due to frozen shoulder. Our findings support the range of problems reported in previous studies by people living with chronic shoulder pain.^{18,98-101} Our results on participants' concerns in seeking early diagnosis and referrals are comparable with those of a previous qualitative study in people with frozen shoulder.¹⁸ However, in the context of variable prognosis (from self-resolving to resistant/chronic cases) of frozen shoulder, a screening tool to identify the subgroup of patients who might benefit from early referral would be helpful. Factors such as chronicity, severity, diabetes, inability to cope with functional restriction and pain tolerance could be incorporated to predict the need for further treatment.

Pain relief and regaining shoulder movements and function are important treatment outcomes

Trial participants and health professionals described pain relief and improvements in function and range of motion as the main outcomes to be achieved from the treatment of frozen shoulder. Their priorities resonate with similar results from a previous survey of 225 health-care professionals¹⁰² and from other studies.^{13,18,103,104}

Steroids help with pain relief and with initiating exercises

Our interviews with trial participants and health professionals support the existing evidence on the use of corticosteroid injections for pain relief. Pain relief is important as it enables physiotherapy exercises to be undertaken to maintain the range of movement and to avoid long-term symptoms. There is moderate evidence to support the efficacy of steroids for pain, function and disability when compared with placebo^{21,105} and additional benefit with shoulder exercises.¹⁰⁶ Steroids are also reported to be a potentially cost-effective option.¹³

Commitment to adhering to prescribed exercises is important for better outcomes

Participants and health professionals had similar views that continued patient engagement with the prescribed exercise is important for better outcomes.¹⁰⁷⁻¹⁰⁹ During the early recovery phase, UK FROST participants were motivated by treatment benefits¹⁰⁹ and had self-determination¹¹⁰ to cope with the

pain associated with exercise. However, during the recovery phase, participants prioritised daily functional activities and did not seem to mind about the minor residual deficits they still had. They aimed for pain relief and enough movement to allow adequate daily function. These findings are in line with a previous study which conceptualised participants' views on 'ideal' (no symptoms at all) and 'adequate' (return to function with residual deficits) recovery from musculoskeletal complaints.¹¹¹

Trial participants, surgeons, and physiotherapists had their personal preferences among UK FROST treatments

Our interview findings suggest that treatment choices did exist among trial participants^{112,113} and health professionals.¹¹⁴⁻¹¹⁶ The preferences of surgeons and physiotherapists were mainly based on their clinical experience in routine practice.¹¹⁴⁻¹¹⁶ It would be highly unlikely for experts not to acquire personal preferences, especially when treatment decisions are expertise based because of a lack of strong evidence. The trial participants also had a range of preferences before participating in the trial. This is evident from both the main trial data and the interviews. In spite of having personal treatment preferences before the trial, the trial participant interviews indicated that all UK FROST interventions were well received and accepted. This also supports the main trial findings, which indicated that patient preferences did not influence the treatment outcomes.

Frozen shoulder and diabetes

Frozen shoulder is a common complaint in people with diabetes, with an incidence ranging between 10% and 36%.¹¹⁷ Our interview findings indicate that the presence or absence of diabetes did not influence trial participants' experiences of the trial interventions. These support the main trial findings, which indicate no significant between-group differences in the mean OSS between diabetic and non-diabetic participants across the treatment arms. The perceptions of surgeons and physiotherapists were that people with diabetes tend to have poorer outcomes (prolonged/severe symptoms, or resistant to conservative management). These are supported by the existing literature,^{118,119} and findings from the main trial found diabetes as a significant predictor of outcome in people with frozen shoulder.

Hydrodilatation as one of the UK FROST treatments

Hydrodilatation involves stretching the capsule of the shoulder joint and reducing the inflammation within it by injecting a mixture of sterile saline, local anaesthetic and steroid. The UK FROST surgeons and physiotherapists suggested that hydrodilatation should have been one of the treatment arms of the trial. They perceived it as an easy-to-administer, less invasive and cost-effective alternative to combat the NHS waiting lists for surgery. However, the available evidence on the effects of hydrodilatation is inconclusive.^{25,120,121} A meta-analysis¹²⁰ of seven small RCTs concluded that hydrodilatation combined with corticosteroid has no significant clinical effect on pain, disability and shoulder movements compared with corticosteroid alone. A further RCT²⁵ in 50 participants with severe frozen shoulder found ACR compared with hydrodilatation improved OSS at 6 months. Despite the lack of sufficient evidence, hydrodilatation appears to be growing in popularity and is being increasingly used by shoulder surgeons.¹²² This resonates with the views of the UK FROST surgeons. More large-scale and high-quality RCTs evaluating the clinical effectiveness and safety of hydrodilatation compared with those of other treatments are essential to make recommendations and to guide evidence-based practice.¹²³

Surgery preferred for prolonged or resistant frozen shoulder cases

Following conservative management, people with frozen shoulder might continue to have persistent pain and poorer outcomes.^{123,124} Often, people with prolonged symptoms and those resistant to conservative treatments of at least 6 months are seen in secondary care and recommended for surgery as the final treatment option.^{26,124,125} Of the two surgical procedures used in UK FROST, evidence shows that surgeons commonly perform ACR¹²³ for the controlled procedure of capsular release¹²⁶ and improved clinical outcomes.¹²⁷ This reflects the views of surgeons and physiotherapists who indicated that people with prolonged frozen shoulder symptoms or not improving with conservative treatment might need ACR. This aligns with the main trial findings, which confirmed that participants who received ACR were least likely to require further treatment.

Study limitations

First, the interviews were conducted with participants who took part in UK FROST and therefore may not be relevant outside this context. Second, the qualitative study only included two trial participants who did not receive their allocated treatments, which could have influenced the predominantly positive experiences towards the trial interventions. Third, given the geographical spread of trial participants and health professionals interviewed, 93% of interviews were conducted via the telephone. Therefore, we are uncertain if participants would have expressed their views differently if face-to-face interviews were conducted. Last, the response rate to participate was low across all arms. Those who did not participate might have reported different experiences.

Strengths

To our knowledge, UK FROST is the first clinical trial to explore the perspectives of both trial participants and health professionals involved in the trial. Interviews were conducted by a researcher not involved in the trial and by using open-ended questions that allowed trial participants and health professionals to express their opinions freely. The interview codes and themes were reviewed by another qualitative researcher to ensure rigour of analysis and interpretation of data.

Implications for clinical practice

Our findings indicate the following implications for clinical practice:

1. Frozen shoulder has a major impact on all aspects of an individual's life. A better understanding of patients' problems and identifying ways to address their concerns during clinical assessments would optimise holistic and patient-centred care.
2. Trial participants had their own treatment preferences. Some preferred surgery as a quick solution to their shoulder problem and some perceived physiotherapy to be a low-risk intervention. These personal treatment preferences should be well understood by health professionals and patients should be provided with opportunities to address their preferences during shared decision-making.
3. Health professionals should also consider their own preferences for treatment and how these affect their treatment decisions. They should carefully consider the evidence available for the treatments they provide. All UK FROST treatments were perceived as acceptable, beneficial and satisfactory. Steroids play an important role in reducing pain and helping people begin their physiotherapy exercises. The evidence on the benefits and anticipated risks of these treatments must be considered in treatment decision-making and clearly communicated to participants.

Conclusion

This qualitative study has provided a fuller understanding of the perspectives of UK FROST trial participants and health professionals and has complemented some of the key findings of the main trial. Our findings indicate that although the content of the physiotherapy interventions was acceptable to trial participants and health professionals, they also highlight concerns about delivering this intensity of treatment within the constraints of the NHS. Future trial designs would usefully include qualitative research as part of intervention development to ensure the feasibility of the interventions in the NHS. More primary qualitative studies on people with frozen shoulder are needed to integrate patient perspectives into informing patient-centred care and shared decision-making.

Chapter 6 Discussion and conclusion

To our knowledge, UK FROST is the largest RCT to date that evaluates three commonly used options to treat frozen shoulder. The trial was sufficiently powered to allow strong conclusions to be drawn about the effectiveness of these treatments. Crucially, all arms of UK FROST involved physiotherapy protocols that were designed to provide pathways to reduce variations in usual NHS care and to optimise clinical practice. It is therefore important to emphasise that, although physiotherapy is a common treatment in NHS practice, the ESP intervention was a specifically designed, standardised, new physiotherapy pathway for UK FROST that was based on the best available evidence and on expert consensus. The pragmatic, multicentre design focused on delivering good standards of practice for all treatment options. Importantly, unlike previous RCTs, a thorough and detailed economic evaluation was undertaken to assess the relative cost-effectiveness of the three treatments during the trial follow-up period. The primary analysis perspective is the NHS and this will have direct applicability to informing future policy and commissioning decisions in the UK. In this chapter, we begin by summarising the main results, before going on to explore the potential risks of bias that might challenge the trial's validity and applicability. We conclude by discussing the application of the trial's findings to clinical practice and our recommendations for future research.

Principal findings of clinical effectiveness

Primary outcome

At the 12-month primary end point, participants randomised to ACR had, on average, a statistically significantly higher (better) OSS than MUA (2.01 points, 95% CI 0.10 to 3.91 points) and ESP (3.06 points, 95% CI 0.71 to 5.41 points), based on ITT analysis. Although statistically significant, mean estimates were short of the minimal clinically important effect size of 4–5 OSS points (the trial was powered for differences of 4 points for comparing MUA with ACR, and 5 points for comparisons with ESP). Differences of clinically important magnitude, however, were included in the 95% CIs for the benefit of MUA and ESP compared with ACR at 3 months, and of ACR compared with ESP at 12 months. Clinically meaningful group differences may, therefore, exist for these comparisons in the wider population.

Additionally collected OSS scores to assess the impact of waiting times revealed little change between baseline and the start of any of the treatments. Six months following treatment, scores improved more in the surgical arms than in the ESP arm and were similar to final follow-up scores across all arms by 8 months. Analyses of the data incorporating all available time points for each participant (day of treatment, 6 months post treatment, and 3, 6 and 12 months post randomisation) found that, compared with the primary analysis, group differences at the different follow-up points tended to be of smaller magnitude, except between ACR and ESP at 12 months (3.26 points in favour of ACR, 95% CI 1.18 to 5.35 points). The 95% CI still included the minimal clinically important difference for this comparison of 5 OSS points.

There was no statistically significant effect of treatment group for interactions with participants' diabetes status, receipt of previous physiotherapy, baseline treatment preference or duration of frozen shoulder symptoms at baseline.

Secondary outcomes

Of the secondary outcomes, QuickDASH and shoulder pain followed a similar pattern to the OSS, in that significantly poorer outcomes were observed for ACR patients at 3 months but better outcomes were observed at 12 months post randomisation than for MUA or ESP patients. There were no statistically significant differences between the treatment arms for reduction in frozen shoulder symptoms as measured by the extent of recovery. In terms of pain or stiffness at the end of physiotherapy, participants in the ESP arm had relatively lower levels of predominant pain by the end of physiotherapy, whereas participants in the ACR arm had relatively lower levels of predominant stiffness, than those in the other arms.

Fidelity of treatment

Of the participants randomised to their allocated treatment, 82% completed MUA, 80% completed ACR and 81% completed ESP. Only 16 participants (3%) crossed over to a different trial treatment, and 17 (3%) received an alternative, non-trial, treatment. As part of the surgical treatments, optimal release was reported as achieved in 92% of MUA procedures and 98% of ACR procedures. Steroid injection was delivered for all completed MUAs and 28% of ACRs. Steroid injection was also given to 80% of patients randomised to ESP. Participants who completed the ESP intervention attended a median of 9 sessions, whereas PPP following surgical procedures had slightly fewer sessions (median of 7 for MUA and 8 for ACR).

Further treatment

Following completion of their randomised treatment, a number of participants received further treatment. There were no specific criteria to inform this decision, which was at the discretion of the treating surgeon. Most commonly, this was ACR for participants allocated to MUA (seven participants); and further physiotherapy (six participants) or ACR (four participants) for participants allocated to ESP. Participants in the ACR arm received fewest further treatments.

Safety

In total, only 10 SAEs were reported for nine participants, of whom eight were randomised to ACR and two were randomised to MUA. The events mainly related to serious medical complications such as chest infection or stroke, some of which may be related to comorbidities or surgery in general, rather than being specifically related to the trial procedures. As an example, a stroke was diagnosed 3 months after ACR. Furthermore, of the eight SAEs in participants randomised to ACR there were two participants who did not have ACR (one had MUA and the other had a non-trial physiotherapy treatment). Only one of the two participants allocated to MUA who experienced a SAE actually received MUA, and the other participant had no treatment for their frozen shoulder. There was, therefore, only a marginal difference in the safety profile between MUA and ESP for which in the latter group there were half of the participants. There were 33 non-serious AEs reported for 31 participants, with comparable rates in the three arms.

Systematic review update of the currently available evidence

To place the trial findings in the context of current evidence, the HTA systematic review about management of frozen shoulder was updated.¹³ The updated review focused only on evidence from RCTs and the interventions and outcomes collected in UK FROST. Hydrodilatation, however, was also included as its popularity has increased since a survey was undertaken to inform the design of UK FROST.²² Moreover, during the qualitative interviews with health-care professionals in the nested study, some surgeons and physiotherapists commented that this could have been a treatment option in the trial.

Nine trials were identified, including UK FROST. The number of participants in the other trials ranged from 26 to 136; therefore, UK FROST was substantially larger. All trials, including UK FROST, were rated as being at high risk of bias in terms of blinding of participants and clinicians and outcome assessment.^{23-25,128-132} Three trials were rated as being at high risk of bias for incomplete outcome reporting,^{24,128,130} two trials were rated as being at high risk of bias for selective reporting^{129,130} and two trials were rated as being at high risk of bias for 'other' biases.^{23,128} Owing to the considerable heterogeneity of the interventions and the generally limited evidence for many of the comparisons, only two trials, UK FROST and one other trial,¹²⁹ were pooled in a meta-analysis, which compared long-term shoulder functioning between ACR and physiotherapy plus steroid injection. The pooled effect favoured ACR, but was smaller in magnitude than the clinical threshold of the standard effect size used in UK FROST. The second trial provided little additional weighted evidence. Overall, most of the comparisons between treatments were informed by single trials, based in single centres, with considerable variation in the interventions used and timing of outcome assessments. UK FROST provides the strongest evidence with broad generalisability of the three treatments it evaluated.

Although UK FROST did not include hydrodilatation, evidence of hydrodilatation's effectiveness from four trials was inconclusive.^{23,25,130,132}

Cost-effectiveness

The base-case economic analysis showed that at 12 months MUA was, on average, £276 more costly per participant (95% CI £65.67 to £487.35) than ESP. MUA was slightly more beneficial in terms of utilities than ESP (on average 0.0396 more QALYs per participant than ESP, 95% CI -0.0008 to 0.0800). The ICER for the ITT approach in the imputed data set between MUA and ESP was £6984 per additional QALY. ACR was more costly than ESP (on average £1733.78 more expensive per participant, 95% CI 1529.48 to 1938.06). Despite the QALY gain accrued by ACR participants (on average 0.0103 more QALYs per participant than ESP, 95% CI -0.0304 to 0.0510), the ICER was over £100,000 per additional QALY. ACR was more expensive than MUA and resulted in slightly fewer QALYs. Therefore, given the limited differences in outcomes observed in the ACR arm compared with the other two treatment options, along with the much higher costs, it is difficult to justify ACR as a first-line treatment option on evidence of cost-effectiveness. MUA was the intervention most likely to be cost-effective at a £20,000 per QALY threshold (MUA 86% > ESP 14% > ACR 0%).

The results of the base-case analysis remained robust to several sensitivity analyses that assessed the impact of areas of uncertainty around a number of study components. This included our analyses being robust to missing data and the assumptions around missing data. However, the cost-effectiveness of MUA compared with ESP was sensitive to the addition of non-shoulder costs and the broader perspective that included private treatment costs and days off work. A key cost driver in these analyses was days off work, at £113.80 per day. During the 12-month follow-up, participants allocated to the ESP arm had a median of no days off work, participants allocated to the MUA arm had a median of 6 days off work, and participants allocated to the ACR arm had a median of 2 weeks off work. This potentially could be related to quicker access to treatment for ESP participants, and may be important in patient decision-making. The analysis was also limited to a 12-month follow-up. However, as the results of the OSS at 12 months show that 50% of the participants were only 5 points away from regaining full function, this suggests it is unlikely that an important difference in QALYs would emerge during longer-term follow-up. Regarding costs, the important costs of treatment, and complications, were expected to have been captured during the 12-month follow-up.

Qualitative study findings

Trial participants described how frozen shoulder had a major impact on all aspects of their life. They were keen to get their shoulder problem resolved, which motivated them to participate in the trial. They thought that seeking early medical help and having a quicker NHS care pathway were important. In general, trial participants were satisfied with the UK FROST interventions and found them acceptable. They reported improvements in pain, shoulder movements and function. Participants who received ACR described recovering more quickly than they had expected. Surgeons and physiotherapists followed a stage-based treatment approach in their routine practice. Both groups of professionals felt that people with diabetes tend to have poorer outcomes. They suggested that hydrodilatation could have been a treatment arm of the trial. Both groups of professionals commented that some people who had received previously ineffective physiotherapy had not wanted to take part in the trial.

Trial validity and minimising bias

Various measures were taken to ensure trial validity and minimise bias, or to explore the potential for bias, of which some are discussed here.

The secure randomisation method helped to ensure comparability in the characteristics of the participants in the three treatment arms. A greater number of participants in the MUA arm were currently in paid work, and there was some arm imbalance in having had a similar shoulder problem on the opposite side from the reference shoulder. A sensitivity analysis of the primary outcome, which included employment status as an additional covariate, found that the results were similar to those observed in the primary analysis. The use of unequal random allocation reflected differential treatment effect expectations. The greater number of participants allocated to the surgery arms than allocated to physiotherapy allowed for a larger effect size to justify the higher costs of and potential risks associated with surgery.²⁷

To help ensure a good standard of care, surgeons were advised to use techniques with which they were familiar, which also helped to avoid learning curve problems. Most operations were carried out by consultant surgeons for both surgical procedures, and most operating surgeons routinely performed both procedures up to once per month. Both ESP and PPP were delivered by qualified physiotherapists who were predominantly band 6 and treated two or three frozen shoulder patients per month. It is unlikely that not including students or assistants in delivering physiotherapy introduced bias, as this was applied consistently across all treatment arms. The number of physiotherapy sessions across the three trial arms was similar. All participants were provided with standardised, written physiotherapy advice detailing the home exercises they needed to perform.

There were low levels of attrition in the completion of the primary outcome and there was no evidence of differential dropout in any of the treatment arms. There were no systematic differences in baseline characteristics compared between those included in the primary analyses and all randomised participants. The use of a mixed-effect, repeated measures analysis model that included data from any participants with at least one valid follow-up meant that only 6% of participants were not included in the primary analysis. This also increased the statistical power of the analyses compared with the single time point comparison used for the sample size calculation.²⁷ There was a ceiling effect at 12-month follow-up, in that 24% of participants had regained full function (top OSS score). Although it is encouraging that participants across all three treatment arms were recovering well, it could be argued that this limited the potential to find clinically meaningful differences at the primary end point.

Given the nature of the trial treatments, blinding participants and clinicians to treatment allocation was not possible or desirable in this pragmatic trial. The statistician and health economist were blinded to group allocation until after the data had been hard locked and no further changes could be made. The lack of any subgroup effect of participant baseline preferences on treatment outcome (using the OSS) may in part mitigate concerns about introducing bias from a lack of blinding in the participant self-reported primary outcome.

It could be argued that a potential bias of the primary analyses concerned the different waiting times for treatment delivery, with ESP starting at around 14 days, and MUA and ACR starting at around 57 days and 72 days post randomisation, respectively. This could have benefited those in the ESP arm at the 3-month follow-up when compared with participants who had not yet received a trial intervention or who were recovering from a surgical procedure. To account for different waiting times, participants also completed the OSS on the day of treatment and 6 months later. Reassuringly, the OSS appeared to stay stable between baseline and the start of any of the treatments. Analyses incorporating all data were largely consistent with the primary analysis findings. This analysis is limited, however, as it reflects treatment effects at pragmatic follow-up times accounting for the different outcome trajectories, rather than observing what would have happened had all three trial arms been delivered at similar times.

A further potential threat to study validity is non-compliance because the treatments were not delivered as planned to all participants. This could dilute the treatment effect observed in the ITT primary analysis. Only 16 participants (3%) crossed over to a different trial treatment, and 17 (3%) received an alternative treatment that was not a trial intervention (e.g. steroid injection only). However, around 20% of

participants across all three trial arms did not complete their treatment, according to our defined criteria. This was expected, as the natural history is for frozen shoulder to resolve,¹³³⁻¹³⁵ particularly for participants awaiting MUA or ACR, who did not receive their allocated treatment because the waiting time was 57 and 72 days, respectively. For ESP patients, although up to 12 sessions of physiotherapy were encouraged, based on existing evidence,^{38,39} we used strict criteria to define 'compliers' as having to complete eight or more sessions, or fewer if the participant and/or physiotherapist were satisfied with their progress. To explore the effect of non-compliance on the OSS at the primary end point of 12 months, an instrumental variable regression was undertaken comparing ESP compliers and those who would have complied in the two surgery groups. ESP outcomes were lower at 12 months, as in the primary analysis, but this was neither statistically significant nor clinically important, with a difference of < 2 points on the OSS. Interestingly, unadjusted OSS at 12 months found that participants who complied with ESP scored, on average, 5 points higher on the OSS than those who did not, which is potentially clinically important. Finally, a steroid injection was delivered for all who received completed MUA and for 80% of patients randomised to ESP, compared with 28% of those who received ACR, who had a steroid injection at the surgeon's discretion. This could be argued to be a bias against ACR, but it is consistent with our finding from a survey of 53 surgeons, carried out when developing the trial protocol, that only 30% routinely provide a steroid injection alongside ACR.²⁷ Therefore, this result reflects clinical practice.

Applicability of results

Characteristics of the trial population

Among the 914 patients screened who met the inclusion criteria, the application of the eligibility criteria meant that only 95 patients were excluded for genuine clinical reasons, the frequency of which was similar across the eligibility criteria. A further 21 patients were excluded for other reasons, 295 eligible patients did not consent and the recruitment target was met with 503 participants randomised into the trial. Review of the baseline characteristics confirmed the inclusion of appropriate trial participants who were in their sixth decade of life and slightly more women.^{133,135} Characteristics were comparable between patients who did and patients who did not consent to take part.

The consent rate among eligible patients was 63%. Nearly one-third of patients did not consent because they 'wanted surgery' or 'did not want physiotherapy'. There were 41% who did not take part because they preferred 'keyhole surgery' (ACR), where over half thought it would be a 'fairly' or 'very' effective treatment. Among patients who did take part, around half had no treatment preference, but the majority of the remaining patients preferred surgery. This preference for surgery could be explained by the fact that trial participants had already had symptoms for around 8-9 months at the time of enrolment. Moreover, nearly two-thirds of trial participants had previously received physiotherapy for their affected shoulder. Although it is recognised that the trial was not powered to detect statistically significant effects between treatment allocation and subgroups, none was found when exploring the effect on treatment outcome of whether or not participants had previously had physiotherapy, their treatment preferences, or the duration of their symptoms.

Finally, 30% of trial participants had diabetes, a common complaint in people with frozen shoulder, ranging between 10% and 36%.¹¹⁷ Participants with diabetes tended to have poorer outcomes at all time points, which is why we stratified for this at randomisation.⁶² The subgroup analyses, however, showed that whether or not participants were diabetic did not have a statistically significant effect on treatment comparisons.

Applicability of the trial findings

The pragmatic design and setting of UK FROST helps to ensure that the findings have immediate applicability to the NHS. The criteria used to enrol participants were minimised, and exclusions were kept to a minimum. Nor were there stringent criteria as to which surgeons could operate on

participants. Most of those surgeons who did operate were consultants, as would be expected. Although trial physiotherapy had to be delivered by qualified physiotherapists (i.e. not by students or assistants), in routine clinical practice students or assistants would be supervised by qualified physiotherapists. The provision of standardised, written physiotherapy advice detailing the home exercises participants needed to perform may not have been entirely reflective of all NHS practice, but it ensured that a good standard of care was applied across all groups.

The trial protocol stipulated that the surgical procedures should be performed within 18 weeks of randomisation, in keeping with NHS waiting list targets at the time, and MUA and ACR were delivered, on average, at around 57 days and 72 days post randomisation. Encouragement was given to deliver both ESP and PPP as soon as possible, particularly PPP within 24 hours of surgery.

It is important to emphasise that, although physiotherapy is a common treatment in NHS practice, the ESP intervention was a specifically designed, standardised and new physiotherapy pathway to test the optimal delivery of physiotherapy in the NHS. Both of the physiotherapy groups were developed using evidence from various sources^{6,13,18,33,34} and consensus from an expert Delphi study,³⁵ which encouraged the delivery of up to 12 treatment sessions. In the case of ESP a steroid injection was to be offered at the first opportunity, whereas with PPP it was not anticipated that a steroid injection would normally be given. Current NHS pressures and waiting times, however, may compromise early access to physiotherapy and timely access to the surgical procedures. The seven or eight sessions of physiotherapy delivered across the three trial arms, along with 80% of participants allocated to ESP receiving a steroid injection, could also have been more than what is routinely provided in the NHS. For example, at baseline, randomised participants reported that they had received only five sessions of physiotherapy and 53% had received a steroid injection. Physiotherapy services may also vary substantially across the UK.²⁸ In the context of the trial, standardised and structured physiotherapy protocols were applied to ensure the rigorous and optimal delivery across all three treatment groups.

During the design of UK FROST, a national survey of health-care professionals found that only 5% used hydrodilatation to treat a frozen shoulder.²² Therefore, UK FROST focused on the more urgent comparisons of ESP with the more costly, invasive surgical interventions. Since then, hydrodilatation appears to have increased in popularity. When the trial team undertook an informal survey with surgeons and physiotherapists who have collaborated with us on UK FROST and on another upper limb orthopaedic surgical trial (ProFHER-2), we found that 52 out of 78 respondents used hydrodilatation to treat a frozen shoulder in a hospital setting. The qualitative interviews, from our nested study, found that some physiotherapists and surgeons thought that hydrodilatation was a treatment option for consideration. The systematic review we have undertaken presents inconclusive evidence of the effectiveness of hydrodilatation from two trials, both small and rated as at high risk of bias, that compared MUA with hydrodilatation,^{23,130} and two further trials that compared hydrodilatation with physiotherapy and steroid injection¹³² and ACR.²⁵ Although the applicability of the UK FROST findings needs to be considered in the context of hydrodilatation's increasing popularity, there is a paucity of rigorous evidence to support its use.

Most trials in the systematic review that compared the treatments included in UK FROST appeared to involve a single centre. By contrast, in UK FROST participants were recruited across a range of urban and rural areas that included 28 hospitals in England, six hospitals in Scotland and one hospital in Wales. The large number of participating hospitals and health-care professionals improves generalisability of findings. There could be concerns about the influence on patient outcome of the small number of participants to whom surgeons and physiotherapists delivered treatment. This effect was statistically controlled for by including the adjustment of hospital site in the primary model.

Application of the trial results to clinical practice

The characteristics of the trial participants and duration of their symptoms were as expected. Therefore, at the primary end point of 12 months for the primary outcome, it was encouraging to find that participants in all three treatment arms had improved considerably since they had been enrolled into the trial. Although participants in the ACR arm did a little better, the mean differences between the treatment options were not of the magnitude of the minimally clinically important difference that we sought. By contrast, at the earlier time point of 3-month follow-up, participants in the MUA arm did a little better than participants in the ACR arm, and a clinically important difference was approached between participants in the ESP and ACR arms in favour of the former. The timing of delivery of these interventions could explain these findings. However, the analyses that attempted to account for variation in waiting times illustrated that the differences between treatment options were smaller than those from the primary analyses, except for a further benefit in favour of ACR compared with ESP, and smaller still than the minimally clinically important difference. The findings for the secondary outcomes showed a similar pattern. Therefore, there is evidence that ESP has potential early benefits compared with ACR. Although it could be argued that these benefits are confounded by waiting times and the surgical procedures being performed in more carefully selected participants whose frozen shoulder had not resolved naturally, pragmatically, this reflects the quicker access to this intervention in clinical practice, with waiting times for surgery likely to be longer than those during the trial. Importantly, participants in the nested qualitative study commented on the need for their frozen shoulder to be resolved and so ESP offers quick access to an effective treatment. Otherwise, the evidence is inconclusive on whether any of the three treatment options is superior in terms of primary and secondary outcomes. For these findings to be replicated in clinical practice, ESP with a steroid injection would need to be delivered as rigorously as it was in UK FROST. Although this might be potentially challenging in routine care in the NHS, the effective delivery of ESP could prevent the 'opportunity cost' of using theatre resources for MUA or ACR, and avoid the need for PPP.

These findings also apply to people with diabetes, as the presence or absence of diabetes in participants did not have a statistically significant effect on treatment comparisons.

All three treatments had similar completion rates, at around 80%. There was an optimal release in 92% and 98% of participants who had MUA or ACR, respectively. Overall, among those allocated to ACR, this treatment was more definitive, with further treatment required for 4% of participants. Nearly twice as many participants in the MUA arm (7%) required further treatment and even more in the ESP arm (15%) required further treatment. Serious complications were rare, although the ACR arm was relatively less safe (4%). Only two participants allocated to MUA experienced a serious complication (1%). One of the participants in the ACR arm who was diagnosed with a deep-vein thrombosis actually received non-trial physiotherapy. Therefore, only a marginal difference was seen in the safety profiles of MUA and ESP, with the latter arm having half the participants. The systematic review that was undertaken in an attempt to further underpin UK FROST findings found that most of the comparisons between treatments were limited by the availability of single trials, often in a single centre with small sample sizes, and the considerable presence of bias and the heterogeneity of treatment interventions. None of the included trials helped to produce conclusive findings about the effectiveness of the interventions evaluated in UK FROST. Although hydrodilatation has increased in popularity in clinical practice, and although research in this area has also increased, evidence of its effectiveness was inconclusive in the systematic review. In this context, hydrodilatation will need to be considered carefully as a treatment option for patients with a frozen shoulder.

Finally, ESP was the least expensive intervention, as most participants in this arm did not require a surgical procedure. MUA was the second most expensive treatment option, with participants spending one-third of the time in theatre compared with ACR, for which the latter was by far the most expensive option. MUA, however, resulted in more QALYs over the duration of the study than either ESP or ACR. This meant that MUA had an 86% probability of being a cost-effective intervention

at the threshold of £20,000 per QALY threshold if commissioners of services would be willing to pay £6984 per additional QALY.

Conclusion

UK FROST has provided robust clinically relevant evidence that none of the three treatments was clearly superior in improving patient-reported shoulder pain and functioning at 12 months. Our specifically designed ESP pathway can be accessed quickly in the NHS and has lower costs. However, the likelihood of further treatment being required is higher with ESP than with the other two interventions. MUA produced the most QALYs overall. At a modest additional cost, MUA is the most cost-effective option to the NHS, with an ICER of £6984 per additional QALY. Patients who receive ACR are the least likely to need further treatment, but ACR is associated with relatively higher risks and costs. These findings should help inform treatment decisions by patients, providers and commissioners of care.

The conclusions should be interpreted with some caution given the potential confounding effect of waiting times to surgery, which have also lengthened since the trial. This may have meant that participants with a more resistant frozen shoulder were those who received an operation. It also could be challenging to implement the ESP pathway in clinical practice with the same optimal timing of access and to the same standard of delivery as in UK FROST.

Recommendations for research

To address the increasing popularity of hydrodilatation, and the paucity of rigorous evidence for hydrodilatation's effectiveness and cost-effectiveness, we recommend the inclusion of hydrodilatation in a high-quality RCT with an economic evaluation. Trial participants had their own treatment preferences in the nested qualitative study; some perceived a surgical procedure to be a quick solution to their shoulder problem, whereas physiotherapy was perceived as a low-risk alternative. Given patient preferences for different treatment options and the trial findings, we propose the RCT be a three-arm trial that compares hydrodilatation versus ESP with steroid injection versus MUA with a steroid injection followed with PPP, as MUA was the more cost-effective of the two surgical interventions. When designing this RCT, including an outcome measure that is not limited by a ceiling effect should be considered and rigorously assessed for stiffness. Finally, in clinical practice it could be complex for patients and surgeons to discuss the risks and benefits of the three treatment options evaluated in UK FROST, along with the inconclusive evidence for hydrodilatation. Therefore, it could be of value to undertake research on how to integrate patient and clinician perspectives on the evidence to inform patient-centred care and shared decision-making.

Acknowledgements

Foremost, we thank the patients participating in this trial, without whom this trial would not have been possible. A general thanks also to all those people whose commitment and efforts in the design and conduct of UK FROST allowed us to successfully complete this study and this final report.

A specific thanks to the following, who made a valuable contribution to the study but who are not named as authors: Dr Sarwat Shah, who was a trial co-ordinator of the study at the University of York, Mr Joe Millar who was the research and development manager at South Tees Hospitals NHS Foundation Trust, and Mrs Amanda Goodman, who provided comprehensive support to the chief investigator and trial team. We also thank the team in York, who contributed to the design of data collection forms, information systems and management of data. A special thanks to Mr Daniel Powell and Mrs Linda Murray, who were our patient representatives in the TMG, and also to the shoulder research users group at James Cook University Hospital.

We thank the British Elbow and Shoulder Society, which funded the transport costs of the tissue and blood samples in the nested study.

This project was funded by the National Institute for Health Research HTA programme (project number 13/26/01).

Collaborators: principal investigators

Mr Philip Ahrens (Royal Free Hospital NHS Foundation Trust).

Miss Alison Armstrong (University Hospitals of Leicester NHS Trust).

Miss Cheryl Baldwick (Northern Devon Healthcare NHS Trust).

Mr Amit Bidwai (Sherwood Forest Hospitals NHS Foundation Trust).

Mr Andrew Brooksbank (Glasgow Royal Infirmary, NHS Greater Glasgow and Clyde).

Mr Muhammad Butt (The Dudley Group NHS Foundation Trust).

Mr Jamie Candel-Couto (Northumbria Healthcare NHS Foundation Trust).

Mr Charalambos Charalambous (Blackpool Teaching Hospitals NHS Foundation Trust).

Mr Mark Crowther (North Bristol NHS Trust).

Mr Steve Drew (University Hospitals Coventry and Warwickshire NHS Trust).

Mr Sunil Garg (James Paget University Hospitals NHS Foundation Trust).

Mr Richard Hawken (Torbay and South Devon NHS Foundation Trust).

Mr Cormac Kelly (The Robert Jones and Agnes Hunt Orthopaedic Hospital NHS Foundation Trust).

Mr Matthew Kent (Royal Liverpool and Broadgreen University Hospitals NHS Trust).

ACKNOWLEDGEMENTS

Mr Kapil Kumar (Woodend Hospital, Aberdeen, NHS Grampian).

Mr Tom Lawrence (University Hospitals Coventry and Warwickshire NHS Trust).

Mr Christopher Little (Oxford University Hospitals NHS Foundation Trust).

Mr Iain Macleod (Hampshire Hospitals NHS Foundation Trust).

Mr Jodi George Malal (Bedfordshire Hospitals NHS Foundation Trust).

Mr Tim Matthews (Cardiff & Vale University Health Board, University Hospital of Wales).

Mr Damian McClelland (University Hospitals of North Midlands NHS Trust).

Mr Neal Millar (West Glasgow Ambulatory Care Hospital, NHS Greater Glasgow and Clyde).

Mr Prabhakar Motkur (United Lincolnshire Hospitals NHS Trust).

Mr Rajesh Nanda (North Tees and Hartlepool Hospitals NHS Foundation Trust).

Mr Chris Peach (Manchester University NHS Foundation Trust).

Mr Tim Peckham (Basildon and Thurrock University Hospitals NHS Foundation Trust).

Mrs Jayanti Rai (East Kent Hospitals University NHS Foundation Trust).

Professor Amar Rangan (South Tees Hospitals NHS Foundation Trust).

Mr Ravi Ray (Basildon and Thurrock University Hospitals NHS Foundation Trust).

Mr Douglas Robinson (Forth Valley Royal Hospital, NHS Forth Valley).

Lt Con Philip Rosell (Frimley Health NHS Foundation Trust).

Mr Adam Ruman (East and North Hertfordshire NHS Trust).

Mr Adnan Saithna (Southport and Ormskirk Hospital NHS Trust).

Mr Colin Senior (Dorset County Hospital NHS Foundation Trust).

Mr Harish Shanker (Royal Alexandra Hospital, NHS Greater Glasgow and Clyde).

Mr Barnaby Sheridan (Taunton and Somerset NHS Foundation Trust).

Mr Kanthan Theivendran (Sandwell and West Birmingham Hospitals NHS Trust).

Mr Simon Thomas (Perth Royal Infirmary, NHS Tayside).

Mr Balachandran Venateswaran (The Mid Yorkshire Hospitals NHS Trust).

Research associates/nurses at collaborating sites

Lynne Allsop, Pam Ashdown, Manjit Attwell, Piet Bakker, Jenna Bardsley, Dawn Barker, Steven Barnfield, Geraldine Belcher, Sue Bell, Helen Black, Lesley Boulton, Rebecca Boulton, Charlotte Boustead, Joanne Bradley-Potts, Sally Breakspear, Tracy Brear, Emma Brennan, Sarah Brown, Sarah Buckley, Ellis Burton, Julie Butler, Melanie Caswell, Diana Cedron, Carla Christie, Louise Clarkson, Rachel Clarkson, Cheryl Cleary, Katherine Coates, Julie Colley, Elise Cooke, Carina Cruz, Win Culley, Carol Dalton, Rachael Daw, Angie Dempster, Christine Dobb, Joanne Donnachie, Alison Dube, Anne Elliot, Barbara Finson, Joanna Forbes, Kelly Goffin, Francesca Gowing, Natalie Grocott, Christian Hacon, Ruth Halliday, Anne Hardwick, Zena Haslam, Melissa Hawkes-Blackburn, Steven Henderson, Christine Herriott, Dennise Hill, Martin Howard, Maggie Hughes, Bianca Hulance, Jane Hunt, Alison Innes, Caron Innes, Juliet James, Joan Joyce, Cherry Kilbride, Tracy Lee, Jane Martin, Richard McCormick, Kerri MCGowen, Pauline Mercer, Alanna Milne, Norma Murray, Dominic Nash, Kimberley Netherton, Anne Nicholson, Claire Nott, Helen Nutt, Joshua O'Donnell, Deborah Page-Browne, Pravin Pastil, Ann Quirk, Angelo Ramos, Alex Ramshaw, Joanne Riches, Cheryl Ritchie, Tessa Rowland, Terri-Ann Sewell, Simon Sharpe, Emma Sharpe, Sarah Sheldon, Emma Shinn, Helen Shirley, Amanda Skinner, Carol Snipe, Jan Stephen, Emma Stoddard, Yogita Stokes, Natalie Taylor, Mairead Tennent, Kerry Third, Susan Thomson, Anne Todd, Dawn Trodd, Sylvia Turner, Mark Vertue, Michael Villaruel, Joanna Wales, Bridget Watkins, Thanuja Weerasinghe, Lori Wells, Kim Wheway, Alison Whitcher, Jessica Whiteman, Matthew Williams, Georgina Williamson, Caroline Wrey Brown and Andrea Young.

Surgeons at collaborating sites

A Armstrong, A Baan, A Bidwai, A Brooksbank, A Desai, A Dunkley, A Leonidou, A Modi, A Putti, A Rajput, A Rangan, A Rumian, A Saithna, B Charalambous, B Sheridan, B Venkateswaran, Banclore, C Baldwick, C Jones, C Kelly, C Modi, C Peach, C Senior, D Cairns, D Robinson, F Ashton, H Colaco, H Ingoe, H Shanker, H Singh, Hilley, I Macleod, J Candal-Couto, J Kurian, J Malal, J Relwani, K Kumar, K Mangat, K Theivendran, M Butt, M Crowther, M Kent, M Smith, Morrison, N Bayliss, N Millar, N Modi, P Ahrens, P Holland, P Jenkins, P Moses, P Motkur, P O'Farrell, Qadir, R Aveledo, R Bansal, R Evans, R Hawken, R Jeavons, R Nanda, R Ray, R Sofat, R Taranu, S Barker, S Chan, S Cloke, S Drew, S Garg, S Sawalha, S Srinivasan, S Thomas, S Venkatachalam, T Lawrence, T Matthews, T Peckham, T Yarashi, U Fazzi, Velu and W Tabi.

Physiotherapists at collaborating sites

A Quirk, C Smith, L MacPherson, M Robertson, N Carr, S Barklie, C Powell, G Hill, S Leahy, C McClymont, D Page-Browne, L Steed, L Wells, S Pak, T Andrews, C Blundell, A Daniel, A Smillie, C Argles, C Holmes, C Hubner, D Coyle, D Keighley, E Haynes, E Newby, H Moulton, H O'Brien, J Bardsley, J James, J Phillips, J Winton, K Clay, K Moses, L Bellis, M Dickinson, N Munnelly, P Lewis, R Daw, S Rogers, D Edmonds, K Williams, V Walker, A Doodson, C Clahson, D Buckle, E Palmer, J Neal, N Howarth, N Maher, C Brookes, M Peak, A Deutsch, A Glover, E Phipps, I Carpenter, L Brookes, M Bailey, S Wang, V Hawke, G Whitting, J Rai, K Flexton, R Camp, S Meadows, A Coutts, C Grady, C Meadley, G Watson, K Williamson, L Dawson, M Taylor, B Danks, D Norris, H Alsafar, L Mercer, N Goodfellow, V Forgie, A Kunda, C Gallagher, G Cranfield, H Sell, L Clark, M London, S Garrett, M Bawden, S Flaherty, S Ingram, C Bourne, R Witham, A Butler, A Fairley, A Haddick, A Meddes, A Rafferty, A Warsama, A Wood, B Bull, B Woods, C Birch, C Day, C McCarthy, C Robinson, C Tarn, D Dernie, D Irvine, D Ryan, G Baker, H Harper, J Watson, J Wilson, N Wild, R Conway, R Cross, R Whittle, S Henderson, S Mtopo, T Docherty, A Mclean, D Holmes, J Darling, L Roberts, R Holman, R Jenks, S Treble, B Davison, C Clark, C Rowell, C Smith, D Robson, J Preston, J Pringle, M Bourke, N Dinsdale, N Watt, P Moss, R Haugh, R Mather, S Chester, T Calvert, T McKay, T Sanderson, J Moser, C Bennett, J Arnot, L Watson, N Bryson, P Jaggard, J Lee, L Scott, M Tsuru, P Patel, R Littleworth,

S Brayshaw, T Richardson, C Andrews, J Lloyd Evans, L Fox, I Gray, Y Edmiston, H Du Toit, J Padkin, L Beaman, N Depala, P Canning, R Jones, R Morgan, S Baynes, S Orbell, D Henderson, M Durell, B Mathew, J Paulraj, K Barak, K Baraks, Y Stokes, A Arnold, A Emmott, A Southgate, D Anthony, D Smith, D Wharton, J Blackmall, J Fletcher, J Grainger, J Moroney, J Rohun, K Budge, K Revitt, L Bardgett, M Korimbocus, M West, R Heyer, S Brown, S Deaton, T Mullally, V Bullock, V Georgopoulos, W Stone, C Gallen, D Mishra, J Curley, M Franklin, M Langley, S Titherley, A Christer, A Donnelly, B Cole, C Usher, J Calvert, J Goodwin, J Hatfield, K Langman, K Thornton, L Goodchild, L Readyhough, L Thomson, L Wildgoose, M Heaton, N Goodenough, R Sheridan, S Boucher, S Cook, S Willingham, C Whiteman, L Hills, M Dagless, C Watts, A Pritchard, C Hegarty, C Khanna, D Selvin, F Deery, H Bush, K Hollway, K Stevenson, L Abbey, L Burton, L Fort, L Nicholls, T Rhodes-Jones, A Din, A Loizou, D Egerton, E Wallis, G Bennett, G Brear, G Cartner, J Farrar, J Taylor, K Madden, R Brooke, R Carter, R Gymer, R Holmes, T Grimson, C Duffield, C Handy, C Mutiziva, F Cunningham, G Betts, G Wright, H Giga, H Tunnicliffe, J Davis, J Leigh, K Aghada, M Hughes, N Cook, N Horsley, P Frostick, P Lewis, P Stanton, R Bove, R Dixey, R Shillcock, S Gaurovski, S Joshi, S Latif, S Mistry, T Swinley, C Gillies, H Hamilton, K Shields, K Taylor and M Kaan.

Trial Steering Committee independent members

Professor Chris Rogers (Chairperson; Professor of Medical Statistics and Clinical Trials, University of Bristol).

Mr Marcus Bateman (Upper Limb Consultant Physiotherapist, Derby Hospitals NHS Foundation Trust).

Mr David Stanley (Consultant Orthopaedic Surgeon, Sheffield Teaching Hospitals NHS Foundation Trust).

Mrs Denise Denton (Independent Member; Hull Ambassador for NHS Hull Clinical Commissioning Group).

Data Monitoring and Ethics Committee independent members

Professor John Norrie [Chairperson; Chairperson of Medical Statistics and Trial Methodology, University of Edinburgh (he is also a member of the National Institute for Health Research HTA and EME Editorial Board)].

Dr Chris Littlewood (Physiotherapist/Senior Lecturer in Clinical Trials, Keele University).

Mr Roger Hackney (Consultant Orthopaedic Surgeon, Leeds Teaching Hospitals NHS Trust).

Contributions of authors

Dr Stephen Brealey (<https://orcid.org/0000-0001-9749-7014>) (Research Fellow, Trial Manager) was a co-applicant and the trial manager. He contributed to the design and conduct of the trial throughout the study, and contributed to and commented on all drafts of the report.

Dr Matthew Northgraves (<https://orcid.org/0000-0001-9260-8643>) (Research Fellow, Trial Co-ordinator) contributed to the day-to-day running of the trial and the ongoing development of the study protocol, and contributed to and commented on all drafts of the report.

Dr Lucksy Kottam (<https://orcid.org/0000-0002-7976-2416>) (Trauma and Orthopaedic Surgery Trial Manager and Trial Co-ordinator) was a co-applicant and trial co-ordinator who contributed to the design and conduct of the trial, and contributed to and commented on all drafts of the report.

Miss Ada Keding (<https://orcid.org/0000-0002-1182-887X>) (Lecturer in Statistics, Statistician) was a co-applicant and statistician on the trial who contributed to the design and monitoring of the trial throughout the study, conducted the statistical analysis, and contributed to and commented on drafts of the report.

Mrs Belen Corbacho (<https://orcid.org/0000-0002-2359-0379>) (Research Fellow, Health Economist) led the write-up of the chapter on the economic evaluation and commented on drafts of the report.

Mrs Lorna Goodchild (<https://orcid.org/0000-0001-6301-7684>) (Clinical Physiotherapist, Co-leader of Physiotherapy Arm) was involved in the design, conduct and delivery of the trial throughout the duration and commented on drafts of the reports.

Dr Cynthia Srikesavan (<https://orcid.org/0000-0002-3540-8052>) (Postdoctoral Research Assistant in Physiotherapy, Qualitative Researcher) contributed to the data acquisition, analysis and interpretation, and preparation of the HTA chapter on the embedded qualitative study of UK FROST.

Mrs Saleema Rex (<https://orcid.org/0000-0002-0318-6125>) (Medical Statistician) contributed to the analysis and reporting of the systematic review that combines UK FROST trial results with other relevant RCTs to treat idiopathic frozen shoulder.

Mr Charalambos P Charalambous (<https://orcid.org/0000-0003-0680-9970>) (Consultant in Trauma and Orthopaedics, Honorary Professor) was a co-applicant who contributed to the design and conduct of the trial, and contributed to and commented on drafts of the report.

Dr Nigel Hanchard (<https://orcid.org/0000-0002-6570-4346>) (Reader in Orthopaedics, Co-Leader of Physiotherapy Arm) was a co-applicant who contributed to the design and conduct of the trial, and commented on drafts of the report.

Mrs Alison Armstrong (<https://orcid.org/0000-0002-6254-0411>) (Consultant Orthopaedic Surgeon) was a co-applicant and was the principal investigator at the Leicester site. She contributed to the design and conduct of the trial, and contributed to and commented on drafts of the report.

Mr Andrew Brooksbank (<https://orcid.org/0000-0002-9271-4259>) (Consultant Orthopaedic Surgeon) was a principal investigator and co-applicant on the original grant. He contributed to study design, trial set up/recruitment and co-ordination of the Scottish centres, and commented on drafts of the report.

Professor Andrew Carr (<https://orcid.org/0000-0001-5940-1464>) (Nuffield Professor of Orthopaedics) was a co-applicant who contributed to the design of the trial, and contributed to and commented on all drafts of the report.

Mrs Cushla Cooper (<https://orcid.org/0000-0002-3790-3428>) (Operational and Trials Development Lead) was a trial co-ordinator who contributed to the conduct of the trial and management of participating sites and the development and implementation of the tissue substudy, and contributed to and commented on all drafts of the report.

Professor Joseph Dias (<https://orcid.org/0000-0001-5360-4543>) (Professor of Orthopaedic and Hand Surgery) was a co-applicant who contributed to the design of the trial, and contributed to and commented on all drafts of the report.

Mrs Iona Donnelly (<https://orcid.org/0000-0002-9794-5347>) (Trial Co-ordinator) made a substantial contribution to the acquisition of data as well as commenting on drafts of the report.

ACKNOWLEDGEMENTS

Professor Catherine Hewitt (<https://orcid.org/0000-0002-0415-3536>) (Professor of Trials and Statistics/Deputy Director York Trials Unit) was a co-applicant and had overall responsibility for the statistical analyses of the trial data. She contributed to the design and analysis of the study, and commented on all drafts of the report.

Professor Sarah E Lamb (<https://orcid.org/0000-0003-4349-7195>) (Professor and Director of Oxford Clinical Trials Research Unit) was a co-applicant and co-supervised the qualitative elements and the development and reporting of the intervention. She has contributed to the design and conduct of the study, and contributed to and commented on drafts of the report.

Dr Catriona McDaid (<https://orcid.org/0000-0002-3751-7260>) (Reader in Trials) was a co-applicant and contributed to the design and conduct of the trial and systematic review, and commented on all drafts of the report.

Professor Gerry Richardson (<https://orcid.org/0000-0002-2360-4566>) (Professor of Health Economics) was a co-applicant who contributed to the design and conduct of the economic evaluation alongside the trial, and contributed to and commented on all drafts of the report.

Mrs Sara Rodgers (<https://orcid.org/0000-0003-1757-541X>) (Research Fellow, Trial Co-ordinator) was a trial co-ordinator who contributed to the conduct of the trial and commented on drafts of the final report.

Mrs Emma Sharp (<https://orcid.org/0000-0001-7988-9508>) (Orthopaedic Research Manager) was a trial co-ordinator who contributed to the study implementation at multiple sites throughout Scotland.

Professor Sally Spencer (<https://orcid.org/0000-0001-6330-7200>) (Director of Postgraduate Medical Institute, Lead for Patient-reported Outcome Measures) was a co-applicant who contributed to the design of the trial, and contributed to and commented on all drafts of the report.

Professor David Torgerson (<https://orcid.org/0000-0002-1667-4275>) (Director of York Trials Unit) was a co-applicant and helped write the study protocol and is a trial methodologist responsible for the overall methods of the trial.

Dr Francine Toye (<https://orcid.org/0000-0002-8144-6519>) (Qualitative Researcher) was a co-applicant who contributed to the design and conduct of the trial of the qualitative research embedded in this trial throughout the duration of the study, and contributed to and commented on all drafts of the report.

Professor Amar Rangan (<https://orcid.org/0000-0002-5452-8578>) (Professor of Orthopaedic Surgery) was the lead applicant and chief investigator. He contributed to the design and conduct of the trial throughout the duration of the study, and contributed to and commented on all drafts of the report.

Data-sharing statement

All data requests should be submitted to the corresponding author for consideration. Access to anonymised data may be granted following review with the chief investigator and the trial team.

References

1. Boyle-Walker KL, Gabard DL, Bietsch E, Masek-VanArsdale DM, Robinson BL. A profile of patients with adhesive capsulitis. *J Hand Ther* 1997;**10**:222–8. [https://doi.org/10.1016/s0894-1130\(97\)80025-7](https://doi.org/10.1016/s0894-1130(97)80025-7)
2. Hand GC, Athanasou NA, Matthews T, Carr AJ. The pathology of frozen shoulder. *J Bone Joint Surg Br* 2007;**89**:928–32. <https://doi.org/10.1302/0301-620X.89B7.19097>
3. Bunker TD, Anthony PP. The pathology of frozen shoulder. A Dupuytren-like disease. *J Bone Joint Surg Br* 1995;**77**:677–83. <https://doi.org/10.1302/0301-620X.77B5.7559688>
4. Hand C, Clipsham K, Rees JL, Carr AJ. Long-term outcome of frozen shoulder. *J Shoulder Elbow Surg* 2008;**17**:231–6. <https://doi.org/10.1016/j.jse.2007.05.009>
5. Buchbinder R, Green S. Effect of arthrographic shoulder joint distension with saline and corticosteroid for adhesive capsulitis. *Br J Sports Med* 2004;**38**:384–5. <https://doi.org/10.1136/bjism.2004.013532>
6. Hanchard NC, Goodchild L, Thompson J, O'Brien T, Davison D, Richardson C. Evidence-based clinical guidelines for the diagnosis, assessment and physiotherapy management of contracted (frozen) shoulder: quick reference summary. *Physiotherapy* 2012;**98**:117–20. <https://doi.org/10.1016/j.physio.2012.01.001>
7. van der Windt DA, Koes BW, de Jong BA, Bouter LM. Shoulder disorders in general practice: incidence, patient characteristics, and management. *Ann Rheum Dis* 1995;**54**:959–64. <https://doi.org/10.1136/ard.54.12.959>
8. Walker-Bone K, Palmer KT, Reading I, Coggon D, Cooper C. Prevalence and impact of musculoskeletal disorders of the upper limb in the general population. *Arthritis Rheum* 2004;**51**:642–51. <https://doi.org/10.1002/art.20535>
9. Bunker TD. Time for a new name for 'frozen shoulder'. *Br Med J* 1985;**290**:1233–4. <https://doi.org/10.1136/bmj.290.6477.1233>
10. Park SW, Lee HS, Kim JH. The effectiveness of intensive mobilization techniques combined with capsular distension for adhesive capsulitis of the shoulder. *J Phys Ther Sci* 2014;**26**:1767–70. <https://doi.org/10.1589/jpts.26.1767>
11. Rundquist PJ, Ludewig PM. Patterns of motion loss in subjects with idiopathic loss of shoulder range of motion. *Clin Biomech* 2004;**19**:810–18. <https://doi.org/10.1016/j.clinbiomech.2004.05.006>
12. Codman E. *Rupture of the Supraspinatus Tendon and Other Lesions in or about the Subacromial Bursa*. Boston, MA: Privately published; 1934.
13. Maund E, Craig D, Suekarran S, Neilson A, Wright K, Brealey S, et al. Management of frozen shoulder: a systematic review and cost-effectiveness analysis. *Health Technol Assess* 2012;**16**(11). <https://doi.org/10.3310/hta16110>
14. Walmsley S, Osmotherly PG, Rivett DA. Movement and pain patterns in early stage primary/idiopathic adhesive capsulitis: a factor analysis. *Physiotherapy* 2014;**100**:336–43. <https://doi.org/10.1016/j.physio.2014.02.001>
15. Rundquist PJ, Anderson DD, Guanache CA, Ludewig PM. Shoulder kinematics in subjects with frozen shoulder. *Arch Phys Med Rehabil* 2003;**84**:1473–9. [https://doi.org/10.1016/S0003-9993\(03\)00359-9](https://doi.org/10.1016/S0003-9993(03)00359-9)

16. Wolf EM, Cox WK. The external rotation test in the diagnosis of adhesive capsulitis. *Orthopedics* 2010;**33**:303. <https://doi.org/10.3928/01477447-20100329-11>
17. Noboa E, López-Graña G, Barco R, Antuña S. Distension test in passive external rotation: validation of a new clinical test for the early diagnosis of shoulder adhesive capsulitis. *Rev Esp Cir Ortop Traumatol* 2015;**59**:354–9. <https://doi.org/10.1016/j.recot.2014.10.005>
18. Jones S, Hanchard N, Hamilton S, Rangan A. A qualitative study of patients' perceptions and priorities when living with primary frozen shoulder. *BMJ Open* 2013;**3**:e003452. <https://doi.org/10.1136/bmjopen-2013-003452>
19. Linsell L, Dawson J, Zondervan K, Rose P, Randall T, Fitzpatrick R, Carr A. Prevalence and incidence of adults consulting for shoulder conditions in UK primary care; patterns of diagnosis and referral. *Rheumatology* 2006;**45**:215–21. <https://doi.org/10.1093/rheumatology/kei139>
20. Wofford JL, Mansfield RJ, Watkins RS. Patient characteristics and clinical management of patients with shoulder pain in U.S. primary care settings: secondary data analysis of the National Ambulatory Medical Care Survey. *BMC Musculoskelet Disord* 2005;**6**:4. <https://doi.org/10.1186/1471-2474-6-4>
21. Rangan A, Hanchard N, McDaid C. What is the most effective treatment for frozen shoulder? *BMJ* 2016;**354**:i4162. <https://doi.org/10.1136/bmj.i4162>
22. Dennis L, Brealey S, Rangan A, Rookmonee M, Watson J. Managing idiopathic frozen shoulder: a survey of health professionals' current practice and research priorities. *Shoulder Elbow* 2010;**2**:294–300. <https://doi.org/10.1111/j.1758-5740.2010.00073.x>
23. Quraishi NA, Johnston P, Bayer J, Crowe M, Chakrabarti AJ. Thawing the frozen shoulder. A randomised trial comparing manipulation under anaesthesia with hydrodilatation. *J Bone Joint Surg Br* 2007;**89**:1197–200. <https://doi.org/10.1302/0301-620X.89B9.18863>
24. Kivimäki J, Pohjolainen T, Malmivaara A, Kannisto M, Guillaume J, Seitsalo S, Nissinen M. Manipulation under anesthesia with home exercises versus home exercises alone in the treatment of frozen shoulder: a randomized, controlled trial with 125 patients. *J Shoulder Elbow Surg* 2007;**16**:722–6. <https://doi.org/10.1016/j.jse.2007.02.125>
25. Gallacher S, Beazley JC, Evans J, Anaspure R, Silver D, Redfern A, *et al.* A randomized controlled trial of arthroscopic capsular release versus hydrodilatation in the treatment of primary frozen shoulder. *J Shoulder Elbow Surg* 2018;**27**:1401–6. <https://doi.org/10.1016/j.jse.2018.04.002>
26. Grant JA, Schroeder N, Miller BS, Carpenter JE. Comparison of manipulation and arthroscopic capsular release for adhesive capsulitis: a systematic review. *J Shoulder Elbow Surg* 2013;**22**:1135–45. <https://doi.org/10.1016/j.jse.2013.01.010>
27. Brealey S, Armstrong AL, Brooksbank A, Carr AJ, Charalambous CP, Cooper C, *et al.* United Kingdom Frozen Shoulder Trial (UK FROST), multi-centre, randomised, 12 month, parallel group, superiority study to compare the clinical and cost-effectiveness of Early Structured Physiotherapy versus manipulation under anaesthesia versus arthroscopic capsular release for patients referred to secondary care with a primary frozen shoulder: study protocol for a randomised controlled trial. *Trials* 2017;**18**:614. <https://doi.org/10.1186/s13063-017-2352-2>
28. Rangan A, Goodchild L, Gibson J, Brownson P, Thomas M, Rees J, *et al.* BESS/BOA patient care pathways – frozen shoulder. *Shoulder Elbow* 2015;**7**:299–307. <https://doi.org/10.1177/1758573215601779>
29. New Zealand Guidelines Group. *The Diagnosis and Management of Soft Tissue Shoulder Injuries and Related Disorders*. Wellington: New Zealand Guidelines Group; 2004.

30. Hanchard NC, Howe TE, Gilbert MM. Diagnosis of shoulder pain by history and selective tissue tension: agreement between assessors. *J Orthop Sports Phys Ther* 2005;**35**:147–53. <https://doi.org/10.2519/jospt.2005.35.3.147>
31. Carr AJ, Cooper CD, Campbell MK, Rees JL, Moser J, Beard DJ, *et al.* Clinical effectiveness and cost-effectiveness of open and arthroscopic rotator cuff repair [the UK Rotator Cuff Surgery (UKUFF) randomised trial.] *Health Technol Assess* 2015;**19**(80). <https://doi.org/10.3310/hta19800>
32. Rangan A, Handoll H, Brealey S, Jefferson L, Keding A, Corbacho Martin B, *et al.* Surgical vs nonsurgical treatment of adults with displaced fractures of the proximal humerus: the PROFHER randomised clinical trial. *JAMA* 2015;**313**:1037–47. <https://doi.org/10.1001/jama.2015.1629>
33. Hanchard N, Goodchild L, Thompson J, O'Brien T, Watson H. Evaluation of clinical guidelines for contracted (frozen) shoulder 12–18 months after publication. *Int J Ther Rehabil* 2013;**20**:543–9. <https://doi.org/10.12968/ijtr.2013.20.11.543>
34. Hanchard NC, Goodchild L, Thompson J, O'Brien T, Davison D, Richardson C. A questionnaire survey of UK physiotherapists on the diagnosis and management of contracted (frozen) shoulder. *Physiotherapy* 2011;**97**:115–25. <https://doi.org/10.1016/j.physio.2010.08.012>
35. Hanchard NCA, Goodchild L, Brealey SD, Lamb SE, Rangan A. Physiotherapy for primary frozen shoulder in secondary care: developing and implementing stand-alone and post operative protocols for UK FROST and inferences for wider practice. *Physiotherapy* 2020;**107**:150–60. <https://doi.org/10.1016/j.physio.2019.07.004>
36. Babcock HM, Matava MJ, Fraser V. Postarthroscopy surgical site infections: review of the literature. *Clin Infect Dis* 2002;**34**:65–71. <https://doi.org/10.1086/324627>
37. Dakin SG, Rangan A, Martinez F, Brealey S, Northgraves M, Kottam L, *et al.* Tissue inflammation signatures point towards resolution in adhesive capsulitis. *Rheumatology* 2019;**58**:1109–11. <https://doi.org/10.1093/rheumatology/kez007>
38. Carette S, Moffet H, Tardif J, Bessette L, Morin F, Frémont P, *et al.* Intraarticular corticosteroids, supervised physiotherapy, or a combination of the two in the treatment of adhesive capsulitis of the shoulder: a placebo-controlled trial. *Arthritis Rheum* 2003;**48**:829–38. <https://doi.org/10.1002/art.10954>
39. Ryans I, Montgomery A, Galway R, Kernohan WG, McKane R. A randomized controlled trial of intra-articular triamcinolone and/or physiotherapy in shoulder capsulitis. *Rheumatology* 2005;**44**:529–35. <https://doi.org/10.1093/rheumatology/keh535>
40. Bloom JE, Rischin A, Johnston RV, Buchbinder R. Image-guided versus blind glucocorticoid injection for shoulder pain. *Cochrane Database Syst Rev* 2012;**8**:CD009147. <https://doi.org/10.1002/14651858.CD009147.pub2>
41. Dawson J, Fitzpatrick R, Carr A. Questionnaire on the perceptions of patients about shoulder surgery. *J Bone Joint Surg Br* 1996;**78**:593–600. <https://doi.org/10.1302/0301-620X.78B4.0780593>
42. Dawson J, Rogers K, Fitzpatrick R, Carr AJ. The Oxford Shoulder Score revisited. *Arch Orthop Trauma Surg* 2009;**129**:119–23. <https://doi.org/10.1007/s00402-007-0549-7>
43. Constant CR, Murley AH. A clinical method of functional assessment of the shoulder. *Clin Orthop Relat Res* 1987;**214**:160–4.
44. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-Item Health Survey 1.0. *Health Econ* 1993;**2**:217–27. <https://doi.org/10.1002/hec.4730020305>

45. Dawson J, Hill G, Fitzpatrick R, Carr AJ. The benefits of using patient-based methods of assessment. *J Bone Joint Surg* 2001;**83**:877–82. <https://doi.org/10.1302/0301-620X.83B6.0830877>
46. Kennedy CA, Beaton DE, Solway S, McConnell S, Bombardier C. *The DASH Outcome Measure User's Manual*. Toronto, ON: Institute for Work and Health; 2011.
47. Beaton DE, Wright JG, Katz JN, Upper Extremity Collaborative Group. Development of the QuickDASH: comparison of three item-reduction approaches. *J Bone Joint Surg Am* 2005;**87**:1038–46. <https://doi.org/10.2106/JBJS.D.02060>
48. Schmitt JS, Di Fabio RP. Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. *J Clin Epidemiol* 2004;**57**:1008–18. <https://doi.org/10.1016/j.jclinepi.2004.02.007>
49. Staples MP, Forbes A, Green S, Buchbinder R. Shoulder-specific disability measures showed acceptable construct validity and responsiveness. *J Clin Epidemiol* 2010;**63**:163–70. <https://doi.org/10.1016/j.jclinepi.2009.03.023>
50. Brooks R. EuroQol: the current state of play. *Health Policy* 1996;**37**:53–72. [https://doi.org/10.1016/0168-8510\(96\)00822-6](https://doi.org/10.1016/0168-8510(96)00822-6)
51. Group EQ. EuroQol – a new facility for the measurement of health-related quality of life. *Health Policy* 1990;**16**:119–208. [https://doi.org/10.1016/0168-8510\(90\)90421-9](https://doi.org/10.1016/0168-8510(90)90421-9)
52. Bharmal M, Thomas J. Comparing the EQ-5D and the SF-6D descriptive systems to assess their ceiling effects in the US general population. *Value Health* 2006;**9**:262–71. <https://doi.org/10.1111/j.1524-4733.2006.00108.x>
53. 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>
54. Angst F, Schwyzer H-K, Aeschlimann A, Simmen BR, Goldhahn J. Measures of adult shoulder function. *Arthritis Care Res* 2011;**63**:S174–88. <https://doi.org/10.1002/acr.20630>
55. Holmgren T, Björnsson Hallgren H, Öberg B, Adolfsson L, Johansson K. Effect of specific exercise strategy on need for surgery in patients with subacromial impingement syndrome: randomised controlled study. *BMJ* 2012;**344**:e787. <https://doi.org/10.1136/bmj.e787>
56. Farrar JT, Young JP, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain* 2001;**94**:149–58. [https://doi.org/10.1016/S0304-3959\(01\)00349-9](https://doi.org/10.1016/S0304-3959(01)00349-9)
57. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SFMPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care Res* 2011;**63**:S240–52. <https://doi.org/10.1002/acr.20543>
58. National Institute for Health and Care Excellence. *Surgical Site Infection: Prevention and Treatment of Surgical Site Infection*. Clinical guideline 74. London: NICE; 2008.
59. Baker P, Nanda R, Goodchild L, Finn P, Rangan A. A comparison of the Constant and Oxford shoulder scores in patients with conservatively treated proximal humeral fractures. *J Shoulder Elbow Surg* 2008;**17**:37–41. <https://doi.org/10.1016/j.jse.2007.04.019>
60. Schulz KF, Grimes DA. Multiplicity in randomised trials I: endpoints and treatments. *Lancet* 2005;**365**:1591–5. [https://doi.org/10.1016/S0140-6736\(05\)66461-6](https://doi.org/10.1016/S0140-6736(05)66461-6)

61. Hewitt CE, Torgerson DJ, Miles JN. Individual allocation had an advantage over cluster randomization in statistical efficiency in some circumstances. *J Clin Epidemiol* 2008;**61**:1004–8. <https://doi.org/10.1016/j.jclinepi.2007.12.002>
62. Griggs SM, Ahn A, Green A. Idiopathic adhesive capsulitis. A prospective functional outcome study of nonoperative treatment. *J Bone Joint Surg Am* 2000;**82**:1398–407. <https://doi.org/10.2106/00004623-200010000-00005>
63. Dawson J, Harris KK, Doll H, Fitzpatrick R, Carr A. A comparison of the Oxford shoulder score and shoulder pain and disability index: factor structure in the context of a large randomized controlled trial. *Patient Relat Outcome Meas* 2016;**7**:195–203. <https://doi.org/10.2147/PROM.S115488>
64. Ashby R, Turner G, Cross B, Mitchell N, Torgerson D. A randomized trial of electronic reminders showed a reduction in the time to respond to postal questionnaires. *J Clin Epidemiol* 2011;**64**:208–12. <https://doi.org/10.1016/j.jclinepi.2010.01.020>
65. Mitchell N, Hewitt CE, Lenaghan E, Platt E, Shepstone L, Torgerson DJ, SCOOP study team. Prior notification of trial participants by newsletter increased response rates: a randomized controlled trial. *J Clin Epidemiol* 2012;**65**:1348–52. <https://doi.org/10.1016/j.jclinepi.2012.05.008>
66. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, *et al*. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health* 2012;**15**:708–15. <https://doi.org/10.1016/j.jval.2012.02.008>
67. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal*. London: NICE; 2013.
68. Manca A, Palmer S. Handling missing data in patient-level cost-effectiveness analysis alongside randomised clinical trials. *Appl Health Econ Health Policy* 2005;**4**:65–75. <https://doi.org/10.2165/00148365-200504020-00001>
69. Curtis L, Burns A. *Unit Costs of Health and Social Care 2018*. Canterbury: Personal Social Services Research Unit, University of Kent; 2018.
70. NHS England. *NHS Reference Costs 2017 to 2018*. URL: www.england.nhs.uk/national-cost-collection/ (accessed 10 September 2020).
71. Joint Formulary Committee. *British National Formulary* (online) London: BMJ Group and Pharmaceutical Press. URL: www.medicinescomplete.com (accessed 10 September 2020).
72. Office for National Statistics. *EARN01: Average Weekly Earnings*. London: Office for National Statistics; 2019.
73. Devlin N, Shah K, Feng Y, Mulhern B, Van Hout B. *An EQ-5D-5L Value Set for England*. London: Office of Health Economics Research Paper; 2015.
74. NICE. *Position Statement on Use of the EQ-5D-5L Value Set for England (Updated October 2019)*. URL: www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisal-guidance/eq5d5l_nice_position_statement.pdf (accessed 1 November 2020).
75. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ* 2005;**14**:487–96. <https://doi.org/10.1002/hec.944>
76. Billingham LJ, Abrams KR, Jones DR. Methods for the analysis of quality-of-life and survival data in health technology assessment. *Health Technol Assess* 1999;**3**(10). <https://doi.org/10.3310/hta3100>

REFERENCES

77. 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>
78. Faria R, Gomes M, Epstein D, White IR. A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. *Pharmacoeconomics* 2014;**32**:1157–70. <https://doi.org/10.1007/s40273-014-0193-3>
79. Leurent B, Gomes M, Faria R, Morris S, Grieve R, Carpenter JR. Sensitivity analysis for not-at-random missing data in trial-based cost-effectiveness analysis: a tutorial. *Pharmacoeconomics* 2018;**36**:889–901. <https://doi.org/10.1007/s40273-018-0650-5>
80. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011;**30**:377–99. <https://doi.org/10.1002/sim.4067>
81. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. Hoboken, NJ: John Wiley & Sons, Inc.; 2004.
82. Johannesson M, Weinstein MC. On the decision rules of cost-effectiveness analysis. *J Health Econ* 1993;**12**:459–67. [https://doi.org/10.1016/0167-6296\(93\)90005-Y](https://doi.org/10.1016/0167-6296(93)90005-Y)
83. Manca A, Lambert PC, Sculpher M, Rice N. Cost-effectiveness analysis using data from multinational trials: the use of bivariate hierarchical modeling. *Med Decis Making* 2007;**27**:471–90. <https://doi.org/10.1177/0272989X07302132>
84. Gomes M, Grieve R, Nixon R, Ng ES, Carpenter J, Thompson SG. Methods for covariate adjustment in cost-effectiveness analysis that use cluster randomised trials. *Health Econ* 2012;**21**:1101–18. <https://doi.org/10.1002/hec.2812>
85. Claxton K, Martin S, Soares M, Rice N, Spackman E, Hinde S, *et al.* *Methods for the Estimation of the NICE Cost Effectiveness Threshold*. CHE Research Paper 81. York: University of York; 2013.
86. Lomas J, Claxton K, Martin S, Soares M. Resolving the ‘cost-effective but unaffordable’ paradox: estimating the health opportunity costs of nonmarginal budget impacts. *Value Health* 2018;**21**:266–75. <https://doi.org/10.1016/j.jval.2017.10.006>
87. Clement C, Edwards SL, Rapport F, Russell IT, Hutchings HA. Exploring qualitative methods reported in registered trials and their yields (EQUITY): systematic review. *Trials* 2018;**19**:589. <https://doi.org/10.1186/s13063-018-2983-y>
88. O’Cathain A, Thomas KJ, Drabble SJ, Rudolph A, Hewison J. What can qualitative research do for randomised controlled trials? A systematic mapping review. *BMJ Open* 2013;**3**:e002889. <https://doi.org/10.1136/bmjopen-2013-002889>
89. Holmberg C, Karner JJ, Rappenecker J, Witt CM. Clinical trial participants’ experiences of completing questionnaires: a qualitative study. *BMJ Open* 2014;**4**:e004363. <https://doi.org/10.1136/bmjopen-2013-004363>
90. Minns Lowe CJ, Moser J, Barker KL. Why participants in the United Kingdom Rotator Cuff Tear (UKUFF) trial did not remain in their allocated treatment arm: a qualitative study. *Physiotherapy* 2018;**104**:224–31. <https://doi.org/10.1016/j.physio.2017.09.002>
91. Prothero L, Sturt J, de Souza S, Lempp H, TITRATE Programme Investigators. Intensive management for moderate rheumatoid arthritis: a qualitative study of patients’ and practitioners’ views. *BMC Rheumatol* 2019;**3**:12. <https://doi.org/10.1186/s41927-019-0057-8>
92. Carter B. Clients’ experiences of frozen shoulder and its treatment with Bowen technique. *Complement Ther Nurs Midwifery* 2002;**8**:204–10. <https://doi.org/10.1054/ctnm.2002.0645>

93. Baker SE, Edwards R. *How Many Qualitative Interviews is Enough? Expert Voices and Early Career Reflections on Sampling and Cases in Qualitative Research*. Southampton: National Centre for Research Methods; 2012.
94. Charmaz K. *Constructing Grounded Theory – A Practical Guide Through Qualitative Analysis*. London: SAGE Publications Ltd; 2006.
95. Rindflesch AB. A grounded-theory investigation of patient education in physical therapy practice. *Physiother Theory Pract* 2009;**25**:193–202. <https://doi.org/10.1080/09593980902776613>
96. World Health Organization (WHO). *Towards a Common Language for Functioning, Disability, and Health: ICF. The International Classification of Functioning, Disability and Health*. Geneva: WHO; 2002.
97. Stucki G. International Classification of Functioning, Disability, and Health (ICF): a promising framework and classification for rehabilitation medicine. *Am J Phys Med Rehabil* 2005;**84**:733–40. <https://doi.org/10.1097/01.phm.0000179521.70639.83>
98. Roe Y, Bautz-Holter E, Juel NG, Soberg HL. Identification of relevant International Classification of Functioning, Disability and Health categories in patients with shoulder pain: a cross-sectional study. *J Rehabil Med* 2013;**45**:662–9. <https://doi.org/10.2340/16501977-1159>
99. Bagheri F, Ebrahimzadeh MH, Moradi A, Bidgoli HF. Factors associated with pain, disability and quality of life in patients suffering from frozen shoulder. *Arch Bone Jt Surg* 2016;**4**:243–7.
100. Toprak M, Erden M. Sleep quality, pain, anxiety, depression and quality of life in patients with frozen shoulder. *J Back Musculoskelet Rehabil* 2019;**32**:287–91. <https://doi.org/10.3233/BMR-171010>
101. Page MJ, O'Connor DA, Malek M, Haas R, Beaton D, Huang H, *et al*. Patients' experience of shoulder disorders: a systematic review of qualitative studies for the OMERACT Shoulder Core Domain Set. *Rheumatology* 2019;**58**:1410–21. <https://doi.org/10.1093/rheumatology/kez046>
102. Rodgers S, Brealey S, Jefferson L, McDaid C, Maund E, Hanchard N, *et al*. Exploring the outcomes in studies of primary frozen shoulder: is there a need for a core outcome set? *Qual Life Res* 2014;**23**:2495–504. <https://doi.org/10.1007/s11136-014-0708-6>
103. Page MJ, Huang H, Verhagen AP, Gagnier JJ, Buchbinder R. Outcome reporting in randomized trials for shoulder disorders: literature review to inform the development of a core outcome set. *Arthritis Care Res* 2018;**70**:252–9. <https://doi.org/10.1002/acr.23254>
104. Page MJ, McKenzie JE, Green SE, Beaton DE, Jain NB, Lenza M, *et al*. Core domain and outcome measurement sets for shoulder pain trials are needed: systematic review of physical therapy trials. *J Clin Epidemiol* 2015;**68**:1270–81. <https://doi.org/10.1016/j.jclinepi.2015.06.006>
105. Sun Y, Lu S, Zhang P, Wang Z, Chen J. Steroid injection versus physiotherapy for patients with adhesive capsulitis of the shoulder: a PRIMSA systematic review and meta-analysis of randomized controlled trials. *Medicine* 2016;**95**:e3469. <https://doi.org/10.1097/MD.0000000000003469>
106. Page MJ, Green S, Kramer S, Johnston RV, McBain B, Chau M, Buchbinder R. Manual therapy and exercise for adhesive capsulitis (frozen shoulder). *Cochrane Database Syst Rev* 2014;**8**:CD011275. <https://doi.org/10.1002/14651858.CD011275>
107. Bailey DL, Holden MA, Foster NE, Quicke JG, Haywood KL, Bishop A. Defining adherence to therapeutic exercise for musculoskeletal pain: a systematic review. *Br J Sports Med* 2020;**54**:326–331. <https://doi.org/10.1136/bjsports-2017-098742>

REFERENCES

108. Chester R, Jerosch-Herold C, Lewis J, Shepstone L. Psychological factors are associated with the outcome of physiotherapy for people with shoulder pain: a multicentre longitudinal cohort study. *Br J Sports Med* 2018;**52**:269–75. <https://doi.org/10.1136/bjsports-2016-096084>
109. Picorelli AM, Pereira LS, Pereira DS, Felício D, Sherrington C. Adherence to exercise programs for older people is influenced by program characteristics and personal factors: a systematic review. *J Physiother* 2014;**60**:151–6. <https://doi.org/10.1016/j.jphys.2014.06.012>
110. Teixeira PJ, Carraça EV, Markland D, Silva MN, Ryan RM. Exercise, physical activity, and self-determination theory: a systematic review. *Int J Behav Nutr Phys Act* 2012;**9**:78. <https://doi.org/10.1186/1479-5868-9-78>
111. Carroll LJ, Lis A, Weiser S, Torti J. How well do you expect to recover, and what does recovery mean, anyway? Qualitative study of expectations after a musculoskeletal injury. *Phys Ther* 2016;**96**:797–807. <https://doi.org/10.2522/ptj.20150229>
112. Patients' preferences within randomised trials: systematic review and patient level meta-analysis. *BMJ* 2008;**337**:a1864. <https://doi.org/10.1136/bmj.a1864>
113. Walter SD, Turner RM, Macaskill P, McCaffery KJ, Irwig L. Estimation of treatment preference effects in clinical trials when some participants are indifferent to treatment choice. *BMC Med Res Methodol* 2017;**17**:29. <https://doi.org/10.1186/s12874-017-0304-x>
114. Cook C, Sheets C. Clinical equipoise and personal equipoise: two necessary ingredients for reducing bias in manual therapy trials. *J Man Manip Ther* 2011;**19**:55–7. <https://doi.org/10.1179/106698111X12899036752014>
115. Ghogawala Z, Barker F, Carter BS. Clinical equipoise and the surgical randomized controlled trial. *Neurosurgery* 2008;**62**:N9–10. <https://doi.org/10.1227/01.neu.0000333288.21621.a9>
116. Kwaees TA, Charalambous CP. Surgical and non-surgical treatment of frozen shoulder. Survey on surgeons treatment preferences. *Muscles Ligaments Tendons J* 2014;**4**:420–4. <https://doi.org/10.32098/mltj.04.2014.05>
117. Noorani A, Goldring M, Jaggi A, Gibson J, Rees J, Bateman M, et al. BESS/BOA patient care pathways: atraumatic shoulder instability. *Shoulder Elbow* 2019;**11**:60–70. <https://doi.org/10.1177/1758573218815002>
118. Barbosa F, Swamy G, Salem H, Creswell T, Espag M, Tambe A, Clark D. Chronic adhesive capsulitis (frozen shoulder): comparative outcomes of treatment in patients with diabetes and obesity. *J Clin Orthop Trauma* 2019;**10**:265–8. <https://doi.org/10.1016/j.jcot.2018.02.015>
119. Hsu CL, Sheu WH. Diabetes and shoulder disorders. *J Diabetes Investig* 2016;**7**:649–51. <https://doi.org/10.1111/jdi.12491>
120. Saltychev M, Laimi K, Virolainen P, Fredericson M. Effectiveness of hydrodilatation in adhesive capsulitis of shoulder: a systematic review and meta-analysis. *Scand J Surg* 2018;**107**:285–93. <https://doi.org/10.1177/1457496918772367>
121. Sinha R, Patel P, Rose N, Tuckett J, Banerjee AN, Williams J, et al. Analysis of hydrodilatation as part of a combined service for stiff shoulder. *Shoulder Elbow* 2017;**9**:169–77. <https://doi.org/10.1177/1758573216687273>
122. Heire P, Braham R, Mubashar M, Bhatti W. *Ultrasound-guided Hydrodilatation for Adhesive Capsulitis – A Step-by-Step Guide*. Paper presented at European Congress of Radiology, 4–8 March 2015, Vienna, Austria.
123. Guyver PM, Bruce DJ, Rees JL. Frozen shoulder – a stiff problem that requires a flexible approach. *Maturitas* 2014;**78**:11–16. <https://doi.org/10.1016/j.maturitas.2014.02.009>

124. Arce G. Primary frozen shoulder syndrome: arthroscopic capsular release. *Arthrosc Tech* 2015;4:e717–20. <https://doi.org/10.1016/j.eats.2015.06.004>
125. Le HV, Lee SJ, Nazarian A, Rodriguez EK. Adhesive capsulitis of the shoulder: review of pathophysiology and current clinical treatments. *Shoulder Elbow* 2017;9:75–84. <https://doi.org/10.1177/1758573216676786>
126. Mubark IM, Ragab AH, Nagi AA, Motawea BA. Evaluation of the results of management of frozen shoulder using the arthroscopic capsular release. *Ortop Traumatol Rehabil* 2015;17:21–8. <https://doi.org/10.5604/15093492.1143530>
127. Kanbe K. Clinical outcome of arthroscopic capsular release for frozen shoulder: essential technical points in 255 patients. *J Orthop Surg Res* 2018;13:56. <https://doi.org/10.1186/s13018-018-0758-5>
128. Smitherman JA, Struk AM, Cricchio M, McFadden G, Dell RB, Horodyski M, Wright TW. Arthroscopy and manipulation versus home therapy program in treatment of adhesive capsulitis of the shoulder: a prospective randomized study. *J Surg Orthop Adv* 2015;24:69–74. <https://doi.org/10.3113/JSOA.2015.0069>
129. De Carli A, Vadalà A, Perugia D, Frate L, Iorio C, Fabbri M, Ferretti A. Shoulder adhesive capsulitis: manipulation and arthroscopic arthrolysis or intra-articular steroid injections? *Int Orthop* 2012;36:101–6. <https://doi.org/10.1007/s00264-011-1330-7>
130. Jacobs LG, Smith MG, Khan SA, Smith K, Joshi M. Manipulation or intra-articular steroids in the management of adhesive capsulitis of the shoulder? A prospective randomized trial. *J Shoulder Elbow Surg* 2009;18:348–53. <https://doi.org/10.1016/j.jse.2009.02.002>
131. Mukherjee RN, Pandey RM, Nag HL, Mittal R. Frozen shoulder – a prospective randomized clinical trial. *World J Orthop* 2017;8:394–9. <https://doi.org/10.5312/wjo.v8.i5.394>
132. Mun SW, Baek CH. Clinical efficacy of hydrodistention with joint manipulation under interscalene block compared with intra-articular corticosteroid injection for frozen shoulder: a prospective randomized controlled study. *J Shoulder Elbow Surg* 2016;25:1937–43. <https://doi.org/10.1016/j.jse.2016.09.021>
133. Vastamäki H, Kettunen J, Vastamäki M. The natural history of idiopathic frozen shoulder: a 2- to 27-year followup study. *Clin Orthop Relat Res* 2012;470:1133–43. <https://doi.org/10.1007/s11999-011-2176-4>
134. Diercks RL, Stevens M. Gentle thawing of the frozen shoulder: a prospective study of supervised neglect versus intensive physical therapy in seventy-seven patients with frozen shoulder syndrome followed up for two years. *J Shoulder Elbow Surg* 2004;13:499–502. <https://doi.org/10.1016/j.jse.2004.03.002>
135. Reeves B. The natural history of the frozen shoulder syndrome. *Scand J Rheum* 1975;4:193–6. <https://doi.org/10.3109/03009747509165255>

Appendix 1 Participating trusts

- Basildon and Thurrock University Hospitals NHS Foundation Trust.
- Bedfordshire Hospital NHS Foundation Trust.
- Blackpool Teaching Hospitals NHS Foundation Trust.
- Cardiff & Vale University Health Board (University Hospital of Wales).
- Dorset County Hospital NHS Foundation Trust.
- East and North Hertfordshire NHS Trust.
- East Kent Hospitals University NHS Foundation Trust.
- Frimley Health NHS Foundation Trust.
- Hampshire Hospitals NHS Foundation Trust.
- James Paget University Hospitals NHS Foundation Trust.
- Manchester University NHS Foundation Trust.
- NHS Forth Valley (Forth Valley Royal Hospital).
- NHS Grampian (Aberdeen Woodend Hospital).
- NHS Greater Glasgow and Clyde (Glasgow Royal Infirmary, Royal Alexandra Hospital, West Glasgow Ambulatory Care Hospital).
- NHS Tayside (Perth Royal Infirmary).
- North Bristol NHS Trust.
- North Tees and Hartlepool NHS Foundation Trust.
- Northern Devon Healthcare NHS Trust.
- Northumbria Healthcare NHS Foundation Trust.
- Oxford University Hospitals NHS Foundation Trust.
- Royal Free London NHS Foundation Trust.
- Royal Liverpool and Broadgreen University Hospitals NHS Trust.
- Sandwell and West Birmingham Hospitals NHS Trust.
- Sherwood Forest Hospitals NHS Foundation Trust.
- South Tees Hospitals NHS Foundation Trust.
- Southport and Ormskirk Hospital NHS Trust.
- Taunton and Somerset NHS Foundation Trust.
- The Dudley Group NHS Foundation Trust.
- The Mid Yorkshire Hospitals NHS Trust.
- The Robert Jones and Agnes Hunt Orthopaedic Hospital NHS Foundation Trust.
- Torbay and South Devon NHS Foundation Trust.
- United Lincolnshire Hospitals NHS Trust.
- University Hospitals Coventry and Warwickshire NHS Trust.
- University Hospitals of Leicester NHS Trust.
- University Hospitals of North Midlands NHS Trust.

Appendix 2 Table of amendments

Type (non-substantial or substantial)	Approved date	Documents amended	Brief description of amendment
Substantial amendment 1	12 January 2015	Update to trial protocol (V2.0_12/01/15) Update to trial participant information sheet (V2.0_12/01/15) Addition of shoulder home exercise leaflet (V1.0_01/12/14)	<ol style="list-style-type: none"> 1. Clarification in the trial participant information sheet about the possible need for further treatment after all three treatments. Also clarification that the £5 unconditional payment at 12 months is for all treatment arms 2. Addition of home exercise leaflet 3. Change to trial protocol to be explicit the analysis that will adjust for age, sex and diabetes and that an exploratory subgroup analysis will be done for the presence of diabetes
Substantial amendment 2	24 May 2016	Update to trial protocol (V3.0_20/04/16) Update to pre-treatment form (V2.0_20/04/2016) Clinic staff poster (V1.0_04/12/2014)	<ol style="list-style-type: none"> 1. Change to allow trial poster to be available in a public part of the hospital 2. Permission for hospitals to publicise the trial through initiatives such as the 'OK to ask' campaign 3. Amended the protocol with additional subgroup analysis as proposed by DMEC 4. Amended the protocol to update sites on what treatment the participant will have while awaiting surgery 5. Amended protocol for pre-treatment form in the ESP arm to complete either on first day of physiotherapy or before steroid injection, whichever is first 6. Inclusion of text messaging SWAT at 3-month time point 7. Updated protocol with amended protocol regarding feedback comments at the funder's request 8. Added a list of amendment changes to the protocol since original REC approval at the funder's request
Non-substantial amendment 1	17 August 2016	N/A	Addition of new participating sites
Substantial amendment 3	21 December 2016	UK FROST tissue and blood approach letter (V1.0_15/11/2016) UK FROST tissue and blood consent form (V1.0_15/11/2016) UK FROST tissue and blood PIL (V1.0_15/11/2016) UK FROST trial protocol (V4.0_15/11/2016)	The protocol was updated, and accompanying materials were provided, to allow us to undertake a nested shoulder capsular tissue and blood study within the host trial
Non-substantial amendment 2	13 October 2017	N/A	Change in principal investigator at Forth Valley and Basildon sites

N/A, not applicable; REC, Research Ethics Committee; SWAT, study within a trial.

Appendix 3 Recruitment

TABLE 36 Recruitment by site

Site	Screened (n)	Randomised, n (% of screened)	Withdrawn, n (% of randomised)
1	58	12 (21)	0 (0)
2	9	6 (67)	0 (0)
3	12	7 (58)	0 (0)
4	8	6 (75)	0 (0)
5	45	11 (24)	0 (0)
6	8	5 (63)	0 (0)
7	49	17 (35)	1 (6)
8	8	4 (50)	0 (0)
9	11	9 (82)	2 (22)
10	12	12 (100)	1 (8)
11	20	18 (90)	1 (6)
12	4	0 (0)	N/A
13	79	34 (43)	1 (3)
14	17	11 (65)	4 (36)
15	7	7 (100)	0 (0)
16	15	1 (7)	0 (0)
17	48	45 (94)	1 (2)
18	15	7 (47)	0 (0)
19	48	22 (46)	0 (0)
20	2	1 (50)	0 (0)
21	16	14 (88)	0 (0)
22	18	11 (61)	0 (0)
23	10	3 (30)	0 (0)
24	58	26 (45)	0 (0)
25	12	5 (42)	0 (0)
26	26	18 (69)	0 (0)
27	3	0 (0)	N/A
28	10	9 (90)	0 (0)
29	69	49 (71)	2 (4)
30	13	5 (38)	0 (0)
31	52	27 (52)	2 (7)
32	35	16 (46)	1 (6)
33	11	5 (45)	0 (0)

continued

TABLE 36 Recruitment by site (continued)

Site	Screened (n)	Randomised, n (% of screened)	Withdrawn, n (% of randomised)
34	32	23 (72)	1 (4)
35	34	32 (94)	0 (0)
36	32	22 (69)	2 (9)
37	8	3 (38)	0 (0)
Total	914	503 (55)	19 (4)

N/A, not applicable.

Appendix 4 Practitioner characteristics

TABLE 37 Practitioner characteristics

Characteristic	MUA	ACR	ESP	Total
Surgeons	N = 58	N = 65	-	N = 90
Operating surgeon grade, n (%)				
Consultant	36 (62)	42 (65)	-	49 (54)
Registrar	5 (9)	3 (5)	-	7 (8)
Unknown	17 (29)	20 (31)	-	34 (38)
Number of operations of this type performed per month by operating surgeon, n (%)				
0–1	28 (48)	30 (46)	-	38 (42)
2–3	8 (14)	9 (14)	-	11 (12)
≥ 4	3 (5)	4 (6)	-	4 (4)
Missing	19 (33)	22 (34)	-	37 (41)
Physiotherapists	N = 148	N = 175	N = 78	N = 285
Physiotherapist band, n (%)				
Band 5	18 (12)	28 (16)	9 (12)	47 (16)
Band 6	71 (48)	87 (50)	36 (46)	139 (49)
Band 7	43 (29)	47 (27)	23 (29)	73 (26)
Band ≥ 8	15 (10)	12 (7)	10 (13)	24 (8)
Missing	1 (< 1)	1 (1)	-	2 (1)
Physiotherapist experience, n (%)				
Treating 0–1 frozen shoulders per month	44 (30)	58 (33)	18 (23)	94 (33)
Treating 2–3 frozen shoulders per month	65 (44)	75 (43)	34 (44)	127 (45)
Treating ≥ 4 frozen shoulders per month	35 (24)	38 (22)	24 (31)	59 (21)
Missing	4 (3)	4 (2)	2 (3)	5 (2)

Appendix 5 Elements of physiotherapy

TABLE 38 Physiotherapy elements received (common treatments)

Treatment given	ESP (N = 80; average number of sessions, n = 8.7)		MUA (N = 158; average number of sessions, n = 7.9)		ACR (N = 156; average number of sessions, n = 8.3)	
	Predominant pain ^a (average sessions, n = 4.9)	Predominant stiffness (average sessions, n = 3.8)	Predominant pain (average sessions, n = 4.0)	Predominant stiffness (average sessions, n = 3.8)	Predominant pain (average sessions, n = 5.1)	Predominant stiffness (average sessions, n = 3.2)
Patients recording problem at least once	72	63	140	134	146	120
Advice and education						
Patients, n (%)	71 (99)	63 (100)	139 (99)	134 (100)	146 (100)	118 (98)
Number of sessions, mean (SD)	5.3 (3.4)	4.6 (2.4)	4.5 (3.5)	4.2 (3.3)	5.3 (3.5)	4.0 (2.8)
Median (minimum, maximum)	5 (1, 12)	4 (1, 12)	4 (1, 17)	3 (1, 17)	4 (1, 15)	3 (1, 12)
% of total sessions ^b	96 (378/395)	96 (288/300)	97 (621/639)	94 (568/602)	97 (778/802)	95 (468/492)
Home exercises (instruction/review)						
Patients, n (%)	64 (89)	55 (87)	139 (99)	134 (100)	146 (100)	119 (99)
Number of sessions, mean (SD)	4.6 (3.4)	4.3 (2.8)	4.4 (3.5)	4.3 (3.3)	5.3 (3.5)	4.0 (2.7)
Median (minimum, maximum)	4 (1, 12)	4 (1, 12)	3 (1, 17)	3 (1, 17)	4 (1, 15)	4 (1, 12)
% of total sessions ^b	75 (295/395)	80 (239/300)	97 (617/639)	95 (572/602)	96 (772/802)	96 (473/492)
Supervised exercises (gentle active/self-assisted)						
Patients, n (%)	71 (99)	61 (97)	132 (94)	123 (92)	142 (97)	108 (90)
Number of sessions, mean (SD)	5.3 (3.4)	4.6 (2.5)	4.2 (3.4)	3.8 (3.1)	5.0 (3.6)	3.8 (2.6)
Median (minimum, maximum)	5 (1, 12)	4 (1, 12)	3 (1, 17)	3 (1, 17)	4 (1, 15)	3 (1, 12)
% of total sessions ^b	95 (374/395)	93 (280/300)	86 (552/639)	78 (471/602)	89 (711/802)	83 (406/492)

Treatment given	ESP (N = 80; average number of sessions, n = 8.7)		MUA (N = 158; average number of sessions, n = 7.9)		ACR (N = 156; average number of sessions, n = 8.3)	
	Predominant pain ^a (average sessions, n = 4.9)	Predominant stiffness (average sessions, n = 3.8)	Predominant pain (average sessions, n = 4.0)	Predominant stiffness (average sessions, n = 3.8)	Predominant pain (average sessions, n = 5.1)	Predominant stiffness (average sessions, n = 3.2)
Supervised exercises (function based)						
Patients, n (%)	12 (17)	60 (95)	64 (46)	109 (81)	77 (53)	99 (83)
Number of sessions, mean (SD)	3.2 (2.8)	3.5 (2.3)	3.3 (3.2)	3.5 (3.1)	3.3 (2.8)	3.3 (2.6)
Median (minimum, maximum)	2 (1, 9)	3 (1, 11)	2 (1, 13)	2 (1, 17)	2 (1, 13)	3 (1, 12)
% of total sessions ^b	10 (38/395)	70 (211/300)	33 (214/639)	64 (386/602)	31 (251/802)	66 (326/492)
Manual shoulder mobilisation						
Patients, n (%)	8 (11)	17 (27)	82 (59)	82 (61)	90 (62)	71 (59)
Number of sessions, mean (SD)	1.9 (1.8)	2.1 (1.1)	3.2 (2.1)	3.4 (2.9)	3.3 (2.6)	2.9 (2.2)
Median (minimum, maximum)	1 (1, 6)	2 (1, 4)	3 (1, 10)	2 (1, 15)	2 (1, 11)	2 (1, 11)
% of total sessions ^b	4 (15/395)	12 (36/300)	41 (264/639)	47 (280/602)	37 (293/802)	41 (204/492)
Posture correction						
Patients, n (%)	32 (44)	23 (27)	57 (41)	47 (35)	68 (47)	51 (43)
Number of sessions, mean (SD)	3.2 (2.5)	2.3 (1.5)	2.6 (2.1)	2.9 (2.6)	3.1 (2.4)	2.5 (2.2)
Median (minimum, maximum)	2 (1, 10)	2 (1, 6)	2 (1, 9)	2 (1, 11)	2 (1, 13)	2 (1, 11)
% of total sessions ^b	26 (101/395)	18 (54/300)	23 (149/639)	22 (135/602)	26 (208/802)	26 (128/492)
Other						
Patients, n (%)	27 (38)	22 (35)	61 (44)	34 (25)	50 (34)	29 (24)
Number of sessions, mean (SD)	2.6 (2.7)	2.9 (2.5)	2.6 (2.5)	2.6 (2.2)	3.2 (3.4)	2.8 (2.6)
Median (minimum, maximum)	1 (1, 12)	2 (1, 10)	2 (1, 12)	2 (1, 9)	1.5 (1, 15)	2 (1, 11)
% of total sessions ^b	17 (69/395)	21 (64/300)	25 (158/639)	15 (89/602)	20 (158/802)	16 (81/492)
<p>^a Columns relating to predominant pain also include sessions during which patients indicated pain and stiffness equally, as the applicable treatments were the same.</p> <p>^b Percentage out of the total sessions for predominant pain (or stiffness), including sessions during which the particular treatment was not given.</p>						

TABLE 39 Physiotherapy elements received (pain-specific treatments)

Treatment given	ESP	MUA	ACR
	Predominant pain ^a (average sessions, <i>n</i> = 4.9)	Predominant pain (average sessions, <i>n</i> = 4.0)	Predominant pain (average sessions, <i>n</i> = 5.1)
Patients recording problem at least once	72	140	146
Hydrotherapy			
Patients, <i>n</i> (%)	5 (7)	8 (6)	8 (5)
Number of sessions, mean (SD)	3.6 (2.6)	4.0 (3.1)	4.8 (2.1)
Median (minimum, maximum)	4 (1, 7)	3.5 (1, 8)	5 (1, 7)
% of total sessions ^b	5 (18/395)	5 (32/639)	5 (38/802)
Relaxation techniques			
Patients, <i>n</i> (%)	32 (44)	17 (12)	25 (17)
Number of sessions, mean (SD)	2.3 (1.8)	2.1 (1)	1.8 (1.4)
Median (minimum, maximum)	1 (1, 8)	2 (1, 4)	1 (1, 6)
% of total sessions	18 (72/395)	5 (35/639)	6 (45/802)
Superficial cold			
Patients, <i>n</i> (%)	15 (21)	7 (5)	6 (4)
Mean (SD) sessions	2.3 (1.4)	1.6 (1.1)	2.8 (2.6)
Median (minimum, maximum)	2 (1, 6)	1 (1, 4)	1.5 (1, 7)
% of total sessions	9 (35/395)	2 (11/639)	2 (17/802)
TENS			
Patients, <i>n</i> (%)	9 (13)	4 (3)	2 (1)
Number of sessions, mean (SD)	2 (1)	2.5 (1.7)	1.5 (0.7)
Median (minimum, maximum)	2 (1, 4)	2 (1, 5)	1.5 (1, 2)
% of total sessions	5 (18/395)	2 (10/639)	0.4 (3/802)
Trigger-point therapy			
Patients, <i>n</i> (%)	21 (29)	16 (11)	19 (13)
Number of sessions, mean (SD)	3.2 (2.5)	2.1 (1.5)	1.7 (1.1)
Median (minimum, maximum)	2 (1, 10)	2 (1, 6)	1 (1, 4)
% of total sessions	17 (67/395)	5 (34/639)	4 (33/802)

TENS, transcutaneous electrical nerve stimulation.

a Columns relating to predominant pain also include sessions during which patients indicated pain and stiffness equally, as the applicable treatments were the same.

b Percentage out of the total sessions for predominant pain, including sessions during which the particular treatment was not given.

TABLE 40 Physiotherapy elements received (stiffness-specific treatments)

Treatment given	ESP	MUA	ACR
	Predominant stiffness (average sessions, <i>n</i> = 3.8)	Predominant stiffness (average sessions, <i>n</i> = 3.8)	Predominant stiffness (average sessions, <i>n</i> = 3.2)
Patients recording problem at least once	63	134	120
Supervised exercises (stretching)			
Patients, <i>n</i> (%)	1 (2%)	75 (56%)	67 (56%)
Number of sessions, mean (SD)	4.0 (-)	3.0 (2.7)	2.7 (2.4)
Median (minimum, maximum)	4 (4, 4)	2 (1, 15)	2 (1, 11)
% of total sessions ^a	1 (4/300)	37 (223/602)	36 (178/492)
Supervised exercises (strengthening)			
Patients, <i>n</i> (%)	20 (32%)	59 (44%)	52 (43%)
Number of sessions, mean (SD)	2.0 (1.3)	3.1 (2.6)	2.4 (2)
Median (minimum, maximum)	1.5 (1, 5)	2 (1, 11)	2 (1, 8)
% of total sessions ^a	13 (40/300)	30 (182/602)	26 (126/492)
Soft-tissue techniques			
Patients, <i>n</i> (%)	9 (14%)	36 (27%)	34 (28%)
Number of sessions, mean (SD)	2.6 (1.4)	2.6 (2.1)	2.3 (1.8)
Median (minimum, maximum)	3 (1, 5)	2 (1, 8)	2 (1, 8)
% of total sessions ^a	8 (23/300)	15 (92/602)	16 (77/492)
PNF			
Patients, <i>n</i> (%)	22 (35%)	16 (12%)	22 (18%)
Number of sessions, mean (SD)	2.3 (2)	2.6 (1.8)	2.4 (1.9)
Median (minimum, maximum)	1 (1, 8)	2 (1, 6)	2 (1, 9)
% of total sessions ^a	17 (50/300)	7 (42/602)	11 (53/492)
Spinal/scapulothoracic manual therapy			
Patients, <i>n</i> (%)	22 (35%)	21 (16%)	16 (13%)
Number of sessions, mean (SD)	2.0 (1.3)	2.8 (1.8)	2.1 (1.7)
Median (minimum, maximum)	2 (1, 5)	2 (1, 7)	1 (1, 7)
% of total sessions ^a	15 (45/300)	10 (59/602)	7 (33/492)

PNF, proprioceptive neuromuscular facilitation.

^a Percentage out of the total sessions for predominant stiffness, including sessions during which the particular treatment was not given.

Appendix 6 Oxford Shoulder Score subdomains

TABLE 41 Unadjusted OSS pain subdomain by treatment arm

Time point	Treatment arm			Total
	MUA	ACR	ESP	
Baseline				
<i>n</i>	200	202	99	501
Mean (SD)	3.7 (2.59)	3.3 (2.11)	3.6 (2.38)	3.5 (2.37)
Median	3	3	3	3
Minimum, maximum	0, 16	0, 12	0, 11	0, 16
3 months				
<i>n</i>	178	179	90	447
Mean (SD)	7.8 (3.68)	6.7 (3.57)	8.8 (4.00)	7.6 (3.78)
Median	8	7	9	8
Minimum, maximum	0, 16	0, 16	0, 16	0, 16
6 months				
<i>n</i>	177	170	83	430
Mean (SD)	10.7 (3.97)	10.2 (3.75)	10.5 (4.36)	10.5 (3.96)
Median	11	11	11	11
Minimum, maximum	0, 16	1, 16	0, 16	0, 16
12 months				
<i>n</i>	183	175	88	446
Mean (SD)	11.3 (4.22)	12.2 (3.84)	11.3 (4.20)	11.7 (4.09)
Median	12	13	12	12
Minimum, maximum	1, 16	1, 16	1, 16	1, 16

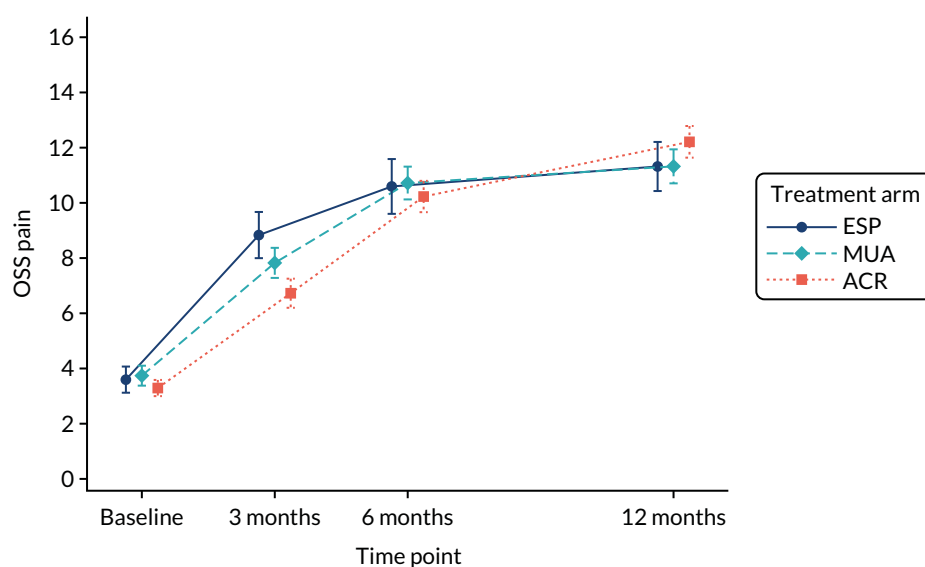


FIGURE 16 Unadjusted mean OSS pain items and 95% CIs by treatment arm.

TABLE 42 Unadjusted OSS function subdomain by treatment arm

Time point	Treatment arm			Total
	MUA	ACR	ESP	
Baseline				
<i>n</i>	200	202	99	501
Mean (SD)	16.7 (6.92)	15.9 (6.21)	16.7 (6.41)	16.4 (6.54)
Median	17	16	17	16
Minimum, maximum	2, 32	1, 29	2, 31	1, 32
3 months				
<i>n</i>	178	179	90	447
Mean (SD)	23.8 (7.24)	20.7 (7.99)	23.9 (7.41)	22.6 (7.72)
Median	26	22	26	25
Minimum, maximum	4, 32	2, 32	1, 32	1, 32
6 months				
<i>n</i>	177	170	83	430
Mean (SD)	27.9 (6.25)	26.2 (6.73)	26.0 (7.07)	26.9 (6.64)
Median	30	29	28	29
Minimum, maximum	2, 32	4, 32	4, 32	2, 32
12 months				
<i>n</i>	183	175	88	446
Mean (SD)	28.0 (6.19)	28.5 (6.50)	27.5 (6.68)	28.1 (6.41)
Median	31	32	30	31
Minimum, maximum	3, 32	1, 32	3, 32	1, 32

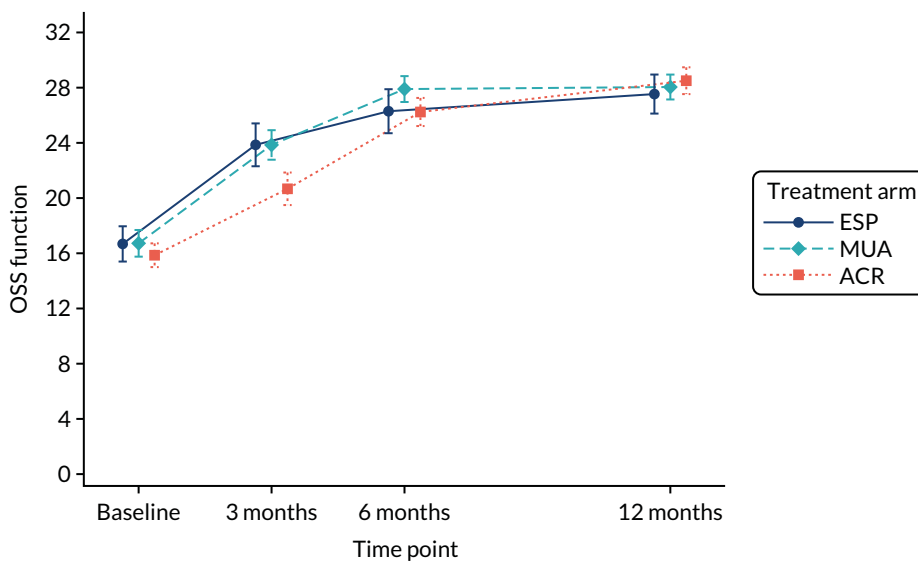


FIGURE 17 Unadjusted mean OSS function items and 95% CIs by treatment arm.

Appendix 7 Illustration of treatment effects

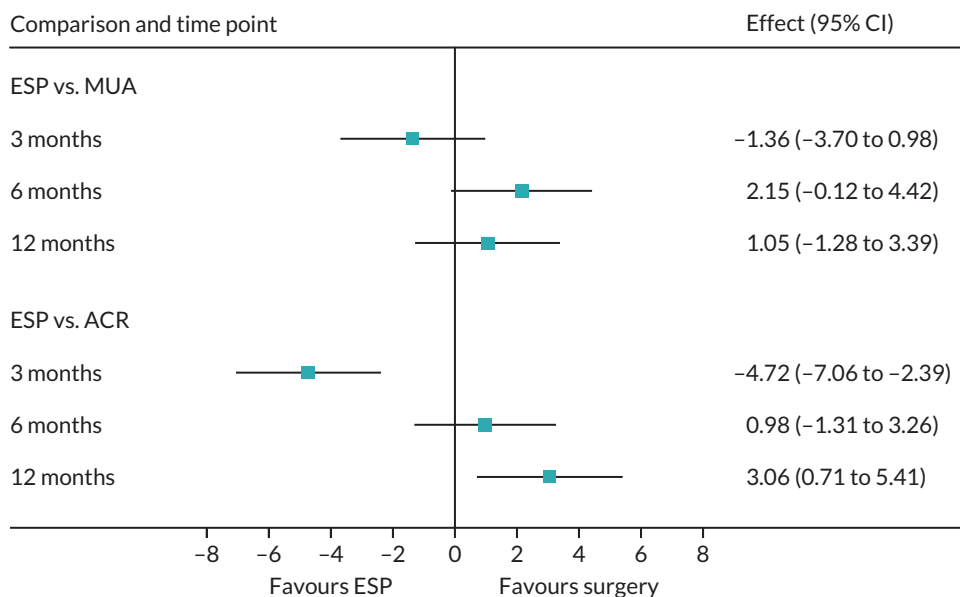


FIGURE 18 Estimated mean OSS differences from primary analysis model by treatment arm (ESP vs. surgery). Grey lines indicate sought minimal clinically important difference.

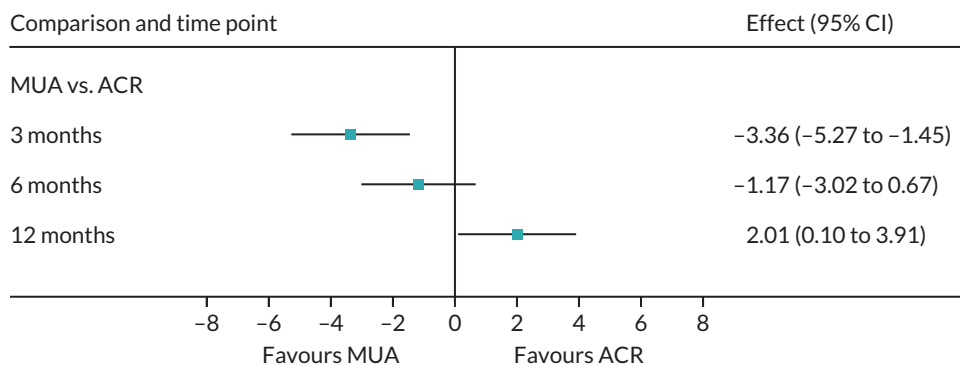


FIGURE 19 Estimated mean OSS differences from primary analysis model by treatment arm (MUA vs. ACR). Grey lines indicate sought minimal clinically important difference.

Appendix 8 Oxford Shoulder Score by treatment completion at start and end of trial

TABLE 43 Unadjusted OSS scores at 12 months by treatment completion

Characteristic	MUA		ACR		ESP	
	Completed treatment	Did not complete treatment	Completed treatment	Did not complete treatment	Completed treatment	Did not complete treatment
Baseline						
<i>n</i>	163	37	161	41	80	19
Mean (SD)	20.4 (8.9)	20.8 (8.9)	19.0 (7.6)	19.9 (8.4)	20.7 (7.8)	18.3 (8.4)
Median (minimum, maximum)	20 (2, 48)	20 (3, 36)	19 (1, 37)	19 (4, 35)	20 (2, 42)	18 (4, 34)
12 months						
<i>n</i>	157	26	147	28	77	11
Mean (SD)	39.8 (9.3)	36.5 (12.4)	41.1 (9.5)	38.8 (12.3)	39.5 (10.2)	34.2 (11.8)
Median (minimum, maximum)	43 (4, 48)	39 (7, 48)	45 (2, 48)	44.5 (7, 48)	43 (4, 48)	39 (10, 46)

Appendix 9 Other secondary analyses

Analysis excluding questionnaire responses received more than 6 weeks beyond the intended follow-up

As > 5% of responses were received beyond the intended follow-up (6% at 3 months, 10% at 6 months, 9% at 12 months), these data were excluded from the primary analysis model in a secondary analysis. Overall, the results remained similar to those observed in the primary analysis (Table 44). The magnitude of differences between MUA and ESP was slightly reduced at all time points, and treatment differences were shown to be less in favour of ACR at all time points when compared with ESP and MUA.

Analysis adjusting for baseline imbalances

As employment status was found to be slightly imbalanced between treatment arms and associated with OSS scores (participants in paid work having better outcomes), it was included as an additional covariate in the analysis model as a sensitivity analysis. The results were similar to those observed in the primary analysis (Table 45).

TABLE 44 Estimated mean OSS differences by treatment arm (estimates from analysis, excluding data received after 6 weeks)

Time point	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)	p-value
	MUA	ESP	Difference	
3 months	30.6 (29.2 to 32.0)	31.8 (29.8 to 33.9)	-1.24 (-3.65 to 1.18)	0.32
6 months	37.3 (35.9 to 38.6)	35.5 (33.5 to 37.5)	1.74 (-0.60 to 4.09)	0.15
12 months	38.5 (37.1 to 40.0)	37.5 (35.5 to 39.6)	0.98 (-1.45 to 3.40)	0.43
	ACR	ESP	Difference	
3 months	26.7 (25.3 to 28.1)	31.8 (29.8 to 33.9)	-5.11 (-7.53 to -2.68)	< 0.01
6 months	35.8 (34.4 to 37.2)	35.5 (33.5 to 37.5)	0.28 (-2.07 to 2.64)	0.81
12 months	40.0 (38.6 to 41.5)	37.5 (35.5 to 39.6)	2.50 (0.05 to 4.94)	0.05
	ACR	MUA	Difference	
3 months	26.7 (25.3 to 28.1)	30.6 (29.2 to 32.0)	-3.87 (-5.80 to -1.95)	< 0.01
6 months	35.8 (34.4 to 37.2)	37.3 (35.9 to 38.6)	-1.46 (-3.32 to 0.40)	0.12
12 months	40.0 (38.6 to 41.5)	38.5 (37.1 to 40.0)	1.52 (-0.44 to 3.47)	0.13

TABLE 45 Estimated mean OSS differences by treatment arm (estimates from analysis, adjusted for employment status)

Time point	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)	p-value
	MUA	ESP	Difference	
3 months	29.9 (28.5 to 31.3)	31.5 (29.6 to 33.5)	-1.63 (-3.97 to 0.71)	0.17
6 months	36.7 (35.3 to 38.1)	34.9 (33.0 to 36.8)	1.82 (-0.46 to 4.11)	0.12
12 months	37.9 (36.5 to 39.3)	37.2 (35.2 to 39.1)	0.78 (-1.56 to 3.11)	0.51
	ACR	ESP	Difference	
3 months	26.7 (25.3 to 28.1)	31.5 (29.6 to 33.5)	-4.84 (-7.17 to -2.50)	< 0.01
6 months	35.6 (34.3 to 37.0)	34.9 (33.0 to 36.8)	0.77 (-1.51 to 3.06)	0.51
12 months	40.1 (38.7 to 41.5)	37.2 (35.2 to 39.1)	2.89 (0.55 to 5.24)	0.02
	ACR	MUA	Difference	
3 months	26.7 (25.3 to 28.1)	29.9 (28.5 to 31.3)	-3.21 (-5.13 to -1.29)	< 0.01
6 months	35.6 (34.3 to 37.0)	36.7 (35.3 to 38.1)	-1.05 (-2.91 to 0.81)	0.27
12 months	40.1 (38.7 to 41.5)	37.9 (36.5 to 39.3)	2.12 (0.21 to 4.03)	0.03

Appendix 10 Subgroup descriptive statistics (Oxford Shoulder Score)

TABLE 46 Unadjusted OSS by treatment arm and diabetes status

Time point	Diabetic (N = 150, 30%)			Non-diabetic (N = 353, 70%)		
	MUA	ACR	ESP	MUA	ACR	ESP
Baseline						
<i>n</i>	141	142	69	12	12	5
Mean (SD)	20.1 (8.43)	19.3 (7.27)	20.8 (8.48)	24.8 (8.55)	20.5 (9.26)	20.8 (8.04)
Median	20	19.5	20	23	22.5	21
Minimum, maximum	2, 40	4, 37	2, 42	15, 48	6, 35	9, 30
3 months						
<i>n</i>	127	128	64	12	11	5
Mean (SD)	32.6 (10.05)	28.9 (10.39)	34.1 (9.75)	33.9 (7.77)	28.3 (11.46)	35 (14.32)
Median	34	31	36	35	32	40
Minimum, maximum	7, 48	5, 48	7, 48	20, 46	9, 43	14, 48
6 months						
<i>n</i>	125	123	60	12	9	3
Mean (SD)	39.9 (8.24)	37.7 (9.26)	38.2 (9.37)	36.6 (9.23)	34.3 (11.74)	33.3 (20.43)
Median	42	40	40	38.5	36	42
Minimum, maximum	5, 48	9, 48	6, 48	15, 48	7, 46	10, 48
12 months						
<i>n</i>	126	122	62	12	10	5
Mean (SD)	40.4 (8.94)	42.0 (8.43)	40.8 (7.79)	39.5 (6.67)	39.2 (13.65)	34.6 (11.78)
Median	43	45	43	38.5	43	35
Minimum, maximum	7, 48	8, 48	10, 48	27, 48	3, 48	20, 48

TABLE 47 Unadjusted OSS by treatment arm and previous physiotherapy

Time point	Had previous physiotherapy for affected shoulder (N = 308, 61%)			Did not have previous physiotherapy for affected shoulder (N = 192, 38%)		
	MUA	ACR	ESP	MUA	ACR	ESP
Baseline						
<i>n</i>	125	123	59	75	77	39
Mean (SD)	20.9 (8.25)	19.3 (7.64)	20.3 (8.54)	19.8 (9.85)	19.1 (7.78)	20 (7.11)
Median	21	19	19	20	20	20
Minimum, maximum	2, 40	2, 37	2, 42	2, 48	1, 37	6, 39
3 months						
<i>n</i>	111	112	58	67	66	32
Mean (SD)	31.4 (10.13)	27.1 (11.40)	32.3 (11.41)	32.1 (10.93)	28.2 (10.50)	33.3 (10.20)
Median	34	27	34.5	34	31	35.5
Minimum, maximum	7, 48	2, 47	4, 46	5, 48	8, 48	8, 48
6 months						
<i>n</i>	108	105	53	69	64	30
Mean (SD)	39.4 (9.01)	36.4 (10.41)	35.4 (11.95)	37.4 (10.65)	36.9 (9.12)	38.5 (9.19)
Median	42	39	39	40	39.5	40
Minimum, maximum	5, 48	7, 48	6, 48	3, 48	10, 48	7, 48
12 months						
<i>n</i>	115	109	54	68	65	34
Mean (SD)	40.4 (8.67)	40.4 (10.18)	38.8 (10.53)	37.7 (11.49)	41.51 (9.48)	39 (10.59)
Median	43	44	42.5	42	45	42.5
Minimum, maximum	4, 48	2, 48	4, 48	5, 48	6, 48	10, 48

TABLE 48 Unadjusted OSS by treatment arm and patient preference

Time point	Randomised to preferred treatment (N = 131, 26%)			Randomised to non-preferred treatment (N = 105, 21%)			Had no preference at baseline (N = 263, 52%)		
	MUA	ACR	ESP	MUA	ACR	ESP	MUA	ACR	ESP
Baseline									
n	56	64	11	39	27	38	103	110	49
Mean (SD)	18.8 (9.52)	16.6 (7.30)	25 (5.76)	21.8 (8.73)	20.2 (8.04)	16.6 (8.08)	20.9 (8.33)	20.4 (7.62)	21.8 (7.22)
Median	18	15	25	22	20	16.5	21	21	21
Minimum, maximum	2, 38	1, 31	16, 39	2, 48	4, 37	2, 42	2, 40	2, 37	6, 37
3 months									
n	47	54	9	36	24	37	94	100	43
Mean (SD)	31.7 (11.31)	26.0 (11.60)	37.9 (5.49)	29.6 (9.18)	26.3 (10.79)	28.6 (11.48)	32.4 (10.44)	28.3 (10.96)	35.3 (10.35)
Median	35	25.5	37	30	27	32	35	30	38
Minimum, maximum	5, 48	3, 47	31, 45	7, 47	8, 44	6, 45	8, 48	2, 48	4, 48
6 months									
n	47	53	9	34	24	32	95	92	42
Mean (SD)	37.6 (11.29)	35.4 (9.82)	41.6 (5.36)	38.68 (6.83)	36.3 (9.72)	33.19 (11.87)	39.0 (9.82)	37.1 (10.21)	37.9 (10.81)
Median	40	38	43	40	40	37	42	40	41
Minimum, maximum	4, 48	10, 48	32, 48	15, 48	14, 48	6, 48	3, 48	7, 48	6, 48
12 months									
n	50	53	10	36	24	34	95	97	43
Mean (SD)	39.2 (9.47)	39.4 (9.84)	42.7 (4.27)	38.5 (8.57)	41.0 (8.92)	38.2 (10.19)	40.0 (10.10)	41.4 (10.39)	38.8 (11.60)
Median	42.5	42	44	40.5	45.5	40.5	44	45	43
Minimum, maximum	4, 48	6, 48	35, 48	10, 48	21, 48	10, 48	5, 48	2, 48	4, 48

TABLE 49 Unadjusted OSS by treatment arm and duration of symptoms at baseline

Time point	Duration of symptoms: < 9 months (N = 249, 61%)			Duration of symptoms: ≥ 9 months (N = 245, 49%)		
	MUA	ACR	ESP	MUA	ACR	ESP
Baseline						
<i>n</i>	103	95	51	93	105	47
Mean (SD)	18.3 (8.33)	18.8 (7.53)	19.6 (7.53)	22.7 (9.04)	19.6 (7.90)	21.0 (8.51)
Median	18	19	18	24	19	21
Minimum, maximum	2, 36	2, 34	2, 39	1, 48	1, 37	4, 42
3 months						
<i>n</i>	88	83	46	86	95	43
Mean (SD)	31.4 (10.21)	28.2 (11.61)	30.7 (10.89)	32.0 (10.76)	26.9 (10.60)	34.5 (10.72)
Median	34	31	32	34	27	36
Minimum, maximum	8, 46	2, 48	6, 48	5, 48	3, 47	4, 48
6 months						
<i>n</i>	89	81	42	84	88	40
Mean (SD)	38.2 (10.42)	37.2 (9.63)	37.6 (8.35)	39.0 (9.07)	36.0 (10.05)	35.1 (13.3)
Median	40	41	40	42	39	40
Minimum, maximum	3, 48	7, 48	12, 48	5, 48	10, 48	6, 48
12 months						
<i>n</i>	94	84	44	86	90	43
Mean (SD)	39.0 (11.35)	40.9 (10.45)	39.1 (10.02)	40.0 (8.00)	40.7 (9.60)	38.5 (8.00)
Median	43	45	43.5	43	44	43
Minimum, maximum	4, 48	2, 48	10, 48	10, 48	6, 48	10, 48

TABLE 50 Unadjusted OSS for patients who completed treatment by receipt of steroid injection

Time point	MUA		ACR		ESP	
	Received steroid injection	Did not receive steroid injection	Received steroid injection	Did not receive steroid injection	Received steroid injection	Did not receive steroid injection
Baseline						
<i>n</i>	163	-	45	-	64	10
Mean (SD)	20.4 (8.89)		18.1 (7.51)		20.7 (6.83)	18.1 (10.04)
Median	20		18		20	17
Minimum, maximum	2, 48		2, 33		6, 39	4, 34
12 months						
<i>n</i>	157	-	40	-	61	4
Mean (SD)	39.8 (9.33)		42.1 (9.76)		39.8 (10.21)	28.5 (16.82)
Median	43		45.5		44	29
Minimum, maximum	4, 48		2, 48		10, 48	10, 46

Appendix 11 Secondary outcomes descriptives

TABLE 51 Unadjusted QuickDASH by treatment arm

Time point	Treatment arm			Total
	MUA	ACR	ESP	
Baseline				
<i>n</i>	192	197	96	485
Mean (SD)	57.0 (20.97)	61.7 (18.51)	59.4 (19.69)	59.4 (19.82)
Median	59	64	60	61
Minimum, maximum	0, 100	14, 100	14, 98	0, 100
3 months				
<i>n</i>	173	178	86	437
Mean (SD)	34.5 (23.95)	43.2 (24.01)	34.0 (23.98)	38.0 (24.32)
Median	30	41	32	34
Minimum, maximum	0, 91	0, 93	0, 96	0, 96
6 months				
<i>n</i>	171	169	75	415
Mean (SD)	21.98 (21.98)	26.1 (21.21)	25.9 (25.07)	24.36 (22.30)
Median	16	21	18	18
Minimum, maximum	0, 98	0, 91	0, 91	0, 98
12 months				
<i>n</i>	175	167	81	423
Mean (SD)	20.0 (23.16)	17.3 (21.39)	20.9 (22.77)	19.1 (22.40)
Median	11	9	14	11
Minimum, maximum	0, 98	0, 93	0, 89	0, 98

TABLE 52 Unadjusted Numeric Rating Scale for Pain by treatment arm

Time point	Treatment arm			Total
	MUA	ACR	ESP	
Baseline				
<i>n</i>	199	201	99	499
Mean (SD)	6.8 (2.23)	7 (1.89)	6.9 (2.37)	6.9 (2.13)
Median	7	7	7	7
Minimum, maximum	0, 10	0, 10	0, 10	0, 10
3 months				
<i>n</i>	178	178	88	444
Mean (SD)	3.8 (2.61)	4.5 (2.64)	3.5 (2.69)	4.0 (2.67)
Median	3	4	3	4
Minimum, maximum	0, 10	0, 10	0, 10	0, 10
6 months				
<i>n</i>	175	169	77	421
Mean (SD)	2.48 (2.43)	2.7 (2.34)	2.6 (2.76)	2.6 (2.46)
Median	2	2	2	2
Minimum, maximum	0, 10	0, 9	0, 9	0, 10
12 months				
<i>n</i>	179	174	86	439
Mean (SD)	2.2 (2.62)	1.6 (2.10)	2.2 (2.55)	2.0 (2.43)
Median	1	1	1.5	1
Minimum, maximum	0, 9	0, 9	0, 10	0, 10

TABLE 53 Unadjusted extent of recovery by treatment arm

Time point	Treatment arm			Total
	MUA	ACR	ESP	
Baseline				
<i>n</i>	198	201	99	498
Mean (SD)	83.8 (21.79)	86.2 (20.11)	89.2 (15.35)	85.9 (20.03)
Median	90	95	100	95
Minimum, maximum	0, 100	0, 100	50, 100	0, 100
3 months				
<i>n</i>	176	176	89	441
Mean (SD)	48.3 (36.35)	51.4 (35.94)	52.0 (36.54)	50.3 (36.18)
Median	50	55	55	50
Minimum, maximum	0, 100	0, 100	0, 100	0, 100
6 months				
<i>n</i>	174	171	78	423
Mean (SD)	29.6 (35.51)	32.3 (33.97)	35 (37.25)	31.7 (35.20)
Median	10	20	20	20
Minimum, maximum	0, 100	0, 100	0, 100	0, 100
12 months				
<i>n</i>	179	175	88	442
Mean (SD)	25.5 (33.99)	18.9 (31.00)	24.6 (31.71)	22.7 (32.5)
Median	5	0	10	4.5
Minimum, maximum	0, 100	0, 100	0, 100	0, 100

TABLE 54 Predominant shoulder problem (among patients who received their allocated treatment)

Problem	Treatment arm, <i>n</i> (%)		
	MUA (N = 164)	ACR (N = 162)	ESP (N = 99)
At the start of physiotherapy	N = 156	N = 152	N = 80
Pain	60 (38)	59 (39)	34 (43)
Stiffness	45 (29)	39 (26)	16 (20)
Pain and stiffness equally	51 (33)	54 (36)	30 (38)
At the end of physiotherapy^a	N = 150	N = 150	N = 78
Pain	37 (25)	39 (26)	15 (19)
Stiffness	98 (65)	82 (55)	52 (67)
Pain and stiffness equally	15 (10)	29 (19)	11 (14)

a For patients who attended two or more physiotherapy sessions.

Appendix 12 Adverse events

TABLE 55 Serious adverse events (summary by treatment arm)

Number of events	MUA (N = 2)	ACR (N = 8)	ESP (N = 0)
Type (n)			
Prolonged hospitalisation	0	2	0
Required hospitalisation	0	0	0
Other medically important condition	1	1	0
Not reported on SAE form	1	5	0
Relationship to trial treatments (n)			
Not related	1	2	0
Unlikely to be related	0	2	0
Possibly related	0	0	0
Probably related	0	1	0
Definitely related	1	3	0
Expectedness (n)			
Expected	0	2	0
Unexpected	1	2	0
Not reported on SAE form	1	4	0
Number of patients	MUA (N = 201)	ACR (N = 203)	ESP (N = 99)
Number of patients with one or more SAE	2 (1%)	7 (3%)	0 (0%)
Number of patients with one SAE	2	6	0
Number of patients with two SAEs	0	1	0

TABLE 56 Non-serious AEs (summary by treatment arm)

Number of events	MUA (N = 15)	ACR (N = 13)	ESP (N = 5)
Relationship to trial treatments (n)			
Not related	4	0	3
Unlikely to be related	2	2	0
Possibly related	4	5	1
Probably related	1	0	0
Definitely related	1	3	0
Not reported on AE form	3	3	1
Expectedness (n)			
Expected	9	7	3
Unexpected	3	3	1
Not reported through AE form	3	3	1
Severity (n)			
Mild	3	6	2
Moderate	5	2	1
Severe	2	2	0
Missing/not reported on AE form	5	3	2
Number of patients	MUA (N = 201)	ACR (N = 203)	ESP (N = 99)
Number of patients with one or more AE	14 (7%)	12 (6%)	5 (5%)
Number of patients with one AE	13	11	5
Number of patients with two AEs	1	1	0

Appendix 13 Treatment preferences

Among non-consenting patients ($n = 295$), ACR was the most popular treatment, followed by ESP (Table 57). Few patients gave MUA as their preferred treatment. Although clinicians did not have a preferred treatment for nearly half of these patients (45%), the agreed treatment was often ACR in line with patient preferences (43%). Average strength of any treatment preference was high for MUA, ACR and ESP (mean of 9 out of 10), but lower for individuals who wanted surgery but did not mind what surgery might be performed (Table 58).

In line with the above results, non-consenting patients thought that ACR was more likely to be effective than other treatments. More than half of these patients expected ESP to be fairly or very ineffective. Randomised patients, on the other hand, were more likely to evaluate trial treatments neutrally (i.e. as neither effective nor ineffective), although ACR was expected to be the most effective (Table 59).

Baseline preferences among randomised patients are shown in Table 60. Approximately half of participants had no treatment preference, and among those who did have a preference this was predominantly for surgery or for ACR in particular. Participants who had received physiotherapy prior to entering the trial only marginally preferred other treatments more and physiotherapy less.

After 12 months' follow-up, there was a trend for patients to change their preference to the treatment they had been allocated to (Table 61). Preference for ACR was much more common among patients who had received their preferred treatment than among those who initially preferred another treatment (50% vs. 31%), whereas the opposite was true for physiotherapy (13% vs. 23%).

TABLE 57 Treatment preferences of non-consenting patients

Preference	Patient preference (consent status CRF) (N = 281), n (%)	Detailed patient preference (optional preferences CRF) (N = 158), n (%)	Clinician advice (consent status CRF) (N = 271), n (%)	Agreed treatment (consent status CRF) (N = 270), n (%)
Any surgery	–	20 (13)	–	–
MUA	26 (9)	11 (7)	27 (10)	32 (12)
ACR	116 (41)	58 (37)	91 (34)	117 (43)
ESP	82 (29)	48 (30)	32 (12)	80 (30)
No preference	49 (17)	21 (13)	121 (45)	–
Other	8 (3)	–	–	41 (15)

TABLE 58 Strength of treatment preference of non-randomised patients (scale 1–10)

	Any surgery	MUA	ACR	ESP
n	14	11	54	34
Mean (SD)	5.1 (4.41)	9.0 (1.10)	9.0 (1.69)	9.0 (1.34)
Median (minimum, maximum)	3 (1, 10)	9 (7, 10)	10 (1, 10)	10 (6, 10)

TABLE 59 Treatment expectations

	Treatment arm, n (%)		
	MUA	ACR	ESP
Non-consenting patients			
Very ineffective	4 (3)	2 (1)	38 (26)
Fairly ineffective	12 (8)	4 (3)	43 (29)
Cannot decide	81 (56)	52 (36)	29 (20)
Fairly effective	24 (17)	35 (24)	26 (18)
Very effective	24 (17)	53 (36)	10 (7)
Randomised patients			
Very ineffective	22 (5)	29 (6)	42 (9)
Fairly ineffective	22 (5)	13 (3)	86 (18)
Cannot decide	231 (47)	208 (43)	252 (52)
Fairly effective	121 (25)	85 (17)	73 (15)
Very effective	90 (18)	151 (31)	35 (7)

TABLE 60 Treatment preferences for randomised patients at baseline

	N	Baseline preference, n (%)				
		MUA	ACR	Either surgery	ESP	No preference
Total						
All patients	499	35 (7)	76 (15)	86 (17)	39 (8)	263 (53)
Allocation						
MUA	199	20 (10)	28 (14)	36 (18)	12 (6)	103 (52)
ACR	202	8 (4)	38 (19)	29 (14)	16 (8)	111 (56)
ESP	98	7 (7)	10 (10)	21 (21)	11 (11)	49 (50)
Previous physiotherapy						
Had previous physiotherapy	306	25 (8)	48 (16)	57 (19)	20 (7)	156 (51)
Did not have previous physiotherapy	190	10 (5)	28 (15)	27 (14)	19 (10)	106 (56)

TABLE 61 Treatment preferences for randomised patients at 12 months

	N	Preference at 12-month follow-up, n (%)				
		MUA	ACR	Either surgery	ESP	No preference
Total						
All patients	416	102 (25)	150 (36)	40 (10)	76 (18)	48 (12)
Allocation						
MUA	166	81 (49)	31 (19)	25 (15)	10 (6)	19 (11)
ACR	166	16 (10)	102 (61)	11 (7)	22 (13)	15 (9)
ESP	84	5 (6)	17 (20)	4 (5)	44 (52)	14 (17)
Baseline preference						
MUA	29	9 (31)	7 (24)	4 (14)	6 (21)	3 (10)
ACR	63	16 (25)	31 (49)	5 (8)	7 (11)	4 (6)
Either surgery	70	17 (24)	33 (47)	6 (9)	9 (13)	5 (7)
ESP	31	5 (16)	8 (26)	0 (0)	12 (39)	6 (19)
No preference	220	54 (25)	70 (32)	24 (11)	42 (19)	30 (14)
Receipt of baseline preferred treatment ^a						
Received preferred treatment	103	23 (22)	51 (50)	9 (9)	13 (13)	7 (7)
Did not receive preferred treatment	90	24 (27)	28 (31)	6 (7)	21 (23)	11 (12)

a Excludes patients who had no baseline preference.

Appendix 14 Oxford Shoulder Score change scores

Patients' assessment of how their shoulder was by the end of the trial compared with 1 year previously revealed that the vast majority of patients felt 'much better' (i.e. data on other response categories were limited). 'Much better' was associated with a median OSS score change of 23, whereas 'slightly better' was associated with a median score change of 10 (Table 62). The trial effect size on which UK FROST was powered was half of this (4–5 OSS points).

TABLE 62 Anchoring of OSS change scores

Difference in OSS score between baseline and 12 months	How is your shoulder compared with 1 year ago?				
	Much better	Slightly better	About the same	Slightly worse	Much worse
<i>n</i>	359	42	20	9	6
Mean (SD)	22.7 (8.0)	9.4 (9.3)	3.4 (7.0)	4.9 (11.8)	-3.7 (6.56)
Median	23	10	0.5	1	-2
Minimum, maximum	-3, 42	-13, 34	-5, 18	-6, 26	-14, 5

Appendix 15 Outcomes for patients receiving no treatment

TABLE 63 Unadjusted trial outcomes by receipt of treatment

	OSS		QuickDASH	
	Received any treatment (N = 441)	Did not receive any treatment (N = 62)	Received any treatment (N = 441)	Did not receive any treatment (N = 62)
Baseline				
<i>n</i>	439	62	428	57
Mean (SD)	19.7 (8.2)	21.1 (8.3)	59.8 (19.7)	56.3 (20.7)
Median	20	20.5	61	58
Minimum, maximum	1, 48	3, 36	0, 100	23, 100
3 months				
<i>n</i>	409	38	400	37
Mean (SD)	30.3 (11.0)	29.0 (11.6)	37.9 (24.2)	38.9 (26.3)
Median	32	33	34	32
Minimum, maximum	2, 48	5, 47	0, 96	0, 91
6 months				
<i>n</i>	395	35	381	34
Mean (SD)	37.5 (10.0)	35.3 (11.1)	24.0 (22.0)	27.9 (25.1)
Median	40	39	18	20.5
Minimum, maximum	3, 48	6, 48	0, 98	0, 89
12 months				
<i>n</i>	405	41	385	38
Mean (SD)	40.1 (9.6)	36.3 (13.5)	18.3 (21.5)	27.0 (29.2)
Median	43	43	11	14
Minimum, maximum	2, 48	7, 48	0, 98	0, 93

Appendix 16 Health-related quality of life and quality-adjusted life-years

TABLE 64 Health-related quality of life: number of questionnaires returned and completed EQ-5D-5L scores

Treatment arm	Baseline, n (%)		3 months, n (%)		6 months, n (%)		12 months, n (%)	
	Complete	Missing	Complete	Missing	Complete	Missing	Complete	Missing
MUA (N = 201)	199 (99)	2 (1)	173 (86)	28 (14)	172 (85)	29 (15)	178 (88)	23 (12)
ACR (N = 203)	200 (98)	3 (2)	175 (86)	28 (14)	165 (81)	38 (19)	175 (86)	28 (14)
ESP (N = 99)	95 (96)	4 (4)	88 (89)	11 (11)	75 (76)	24 (24)	86 (87)	13 (13)

TABLE 65 Health-related quality of life: proportion reporting EQ-5D-5L levels 1-5 by dimension, treatment arm and time point (complete cases)

Dimension	Baseline, n (%)			3 months, n (%)			6 months, n (%)			12 months, n (%)		
	MUA	ACR	ESP	MUA	ACR	ESP	MUA	ACR	ESP	MUA	ACR	ESP
Mobility												
Level 1	159 (79.1)	146 (71.9)	69 (69.7)	148 (73.6)	130 (64.0)	67 (67.7)	145 (72.1)	128 (63.1)	55 (55.6)	144 (71.6)	129 (63.6)	66 (66.7)
Level 2	14 (7.0)	18 (8.9)	10 (10.1)	12 (6.0)	15 (7.4)	8 (8.1)	13 (6.5)	14 (6.9)	7 (7.1)	14 (7.0)	17 (8.4)	10 (10.1)
Level 3	18 (9.0)	28 (13.8)	10 (10.1)	10 (5.0)	20 (9.9)	9 (9.1)	10 (5.0)	14 (6.9)	8 (8.1)	14 (7.0)	17 (8.4)	4 (4.0)
Level 4	8 (4.0)	9 (4.4)	7 (7.1)	5 (2.5)	9 (4.4)	5 (5.1)	5 (2.5)	10 (4.9)	5 (5.1)	8 (4.0)	13 (6.4)	7 (7.1)
Level 5	0 (0.0)	1 (0.5)	0 (0.0)	0 (0.0)	2 (1.0)	1 (1.0)	0 (0.0)	1 (0.5)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Missing	2 (1.0)	1 (0.5)	3 (3.0)	26 (12.9)	27 (13.3)	9 (9.1)	28 (13.9)	36 (17.7)	24 (24.2)	21 (10.5)	27 (13.3)	12 (12.1)
Reporting probability	40 (20.1)	56 (27.7)	27 (28.1)	27 (15.4)	46 (26.1)	23 (25.6)	28 (16.2)	39 (23.4)	20 (26.7)	36 (20.0)	47 (26.7)	21 (24.1)
Self-care												
Level 1	17 (8.5)	12 (5.9)	10 (10.1)	65 (32.3)	39 (19.2)	28 (28.3)	95 (47.3)	75 (37.0)	35 (35.4)	106 (52.7)	116 (57.1)	51 (51.5)
Level 2	66 (32.8)	52 (25.6)	26 (26.3)	67 (33.3)	70 (34.5)	36 (36.4)	54 (26.9)	60 (29.6)	25 (25.3)	52 (25.9)	32 (15.8)	21 (21.2)
Level 3	80 (39.8)	94 (46.3)	45 (45.5)	29 (14.4)	47 (23.2)	15 (15.2)	14 (7.0)	27 (13.3)	10 (10.1)	13 (6.5)	21 (10.3)	9 (9.1)
Level 4	34 (16.9)	42 (20.7)	17 (17.2)	15 (7.5)	19 (9.4)	8 (8.1)	8 (4.0)	5 (2.5)	5 (5.1)	6 (3.0)	6 (3.0)	5 (5.1)
Level 5	3 (1.5)	2 (1.0)	1 (1.0)	0 (0.0)	2 (1.0)	3 (3.0)	1 (0.5)	1 (0.5)	0 (0.0)	3 (1.5)	1 (0.5)	2 (2.0)
Missing	1 (0.5)	1 (0.5)	0 (0.0)	25 (12.4)	26 (12.8)	9 (9.1)	29 (14.4)	35 (17.2)	24 (24.2)	21 (10.5)	27 (13.3)	11 (11.1)
Reporting probability	183 (91.5)	190 (94.1)	89 (89.9)	111 (63.1)	138 (78.0)	62 (68.9)	77 (44.8)	93 (55.4)	40 (53.3)	74 (41.1)	60 (34.1)	37 (42.1)

Dimension	Baseline, n (%)			3 months, n (%)			6 months, n (%)			12 months, n (%)		
	MUA	ACR	ESP	MUA	ACR	ESP	MUA	ACR	ESP	MUA	ACR	ESP
Usual activities												
Level 1	12 (6.0)	7 (3.5)	7 (7.1)	47 (23.4)	26 (12.8)	22 (22.2)	74 (36.8)	64 (31.5)	31 (31.3)	92 (45.8)	94 (46.3)	46 (46.5)
Level 2	58 (28.9)	35 (17.2)	19 (19.2)	70 (34.8)	67 (33.0)	35 (35.4)	67 (33.3)	60 (29.6)	25 (25.3)	53 (26.4)	48 (23.7)	22 (22.2)
Level 3	70 (34.8)	97 (47.8)	43 (43.4)	38 (18.9)	58 (28.6)	20 (20.2)	22 (11.0)	27 (13.3)	13 (13.1)	22 (11.0)	21 (10.3)	12 (12.1)
Level 4	41 (20.4)	53 (26.1)	24 (24.2)	19 (9.5)	22 (10.8)	10 (10.1)	10 (5.0)	14 (6.9)	4 (4.0)	9 (4.5)	10 (4.9)	6 (6.1)
Level 5	19 (9.5)	9 (4.4)	6 (6.1)	2 (1.0)	5 (2.5)	3 (3.0)	0 (0.0)	2 (1.0)	2 (2.0)	3 (1.5)	3 (1.5)	2 (2.0)
Missing	1 (0.5)	2 (1.0)	0 (0.0)	25 (12.4)	25 (12.3)	9 (9.1)	28 (13.9)	36 (17.7)	24 (24.2)	22 (11.0)	27 (13.3)	11 (11.1)
Reporting probability	188 (94.0)	194 (96.5)	92 (92.9)	129 (73.3)	152 (85.4)	68 (75.6)	99 (57.2)	103 (61.7)	44 (58.7)	87 (48.6)	82 (46.6)	42 (47.7)
Pain/discomfort												
Level 1	3 (1.5)	3 (1.5)	1 (1.0)	16 (8.0)	6 (3.0)	11 (11.1)	37 (18.4)	24 (11.8)	18 (18.2)	60 (29.9)	57 (28.1)	22 (22.2)
Level 2	22 (11.0)	11 (5.4)	9 (9.1)	83 (41.3)	61 (30.1)	39 (39.4)	85 (42.3)	87 (42.9)	34 (34.3)	69 (34.3)	73 (36.0)	43 (43.4)
Level 3	88 (43.8)	87 (42.9)	37 (37.4)	48 (23.9)	74 (36.5)	23 (23.2)	37 (18.4)	43 (21.2)	11 (11.1)	31 (15.4)	28 (13.8)	11 (11.1)
Level 4	66 (32.8)	83 (40.9)	35 (35.4)	27 (13.4)	28 (13.8)	11 (11.1)	12 (6.0)	12 (5.9)	7 (7.1)	14 (7.0)	14 (6.9)	7 (7.1)
Level 5	21 (10.5)	17 (8.4)	16 (16.2)	3 (1.5)	9 (4.4)	5 (5.1)	1 (0.5)	2 (1.0)	5 (5.1)	5 (2.5)	3 (1.5)	4 (4.0)
Missing	1 (0.5)	2 (1.0)	1 (1.0)	24 (11.9)	25 (12.3)	10 (10.1)	29 (14.4)	35 (17.2)	24 (24.2)	22 (11.0)	28 (13.8)	12 (12.1)
Reporting probability	197 (98.5)	198 (98.5)	97 (99.0)	161 (91.0)	172 (96.6)	78 (87.6)	135 (78.5)	144 (85.7)	57 (76.0)	119 (66.5)	118 (67.4)	65 (74.7)
Anxiety/depression												
Level 1	97 (48.3)	84 (41.4)	40 (40.4)	110 (54.7)	100 (49.3)	50 (50.5)	117 (58.2)	108 (53.2)	49 (49.5)	120 (59.7)	126 (62.1)	57 (57.6)
Level 2	47 (23.4)	64 (31.5)	25 (25.3)	35 (17.4)	38 (18.7)	18 (18.2)	33 (16.4)	33 (16.3)	13 (13.1)	31 (15.4)	20 (9.9)	15 (15.2)
Level 3	40 (19.9)	41 (20.2)	20 (20.2)	21 (10.5)	27 (13.3)	11 (11.1)	20 (10.0)	15 (7.4)	7 (7.1)	26 (12.9)	21 (10.3)	7 (7.1)
Level 4	10 (5.0)	7 (3.5)	6 (6.1)	8 (4.0)	7 (3.5)	5 (5.1)	3 (1.5)	9 (4.4)	2 (2.0)	3 (1.5)	5 (2.5)	6 (6.1)
Level 5	6 (3.0)	6 (3.0)	8 (8.1)	1 (0.5)	6 (3.0)	5 (5.1)	0 (0.0)	2 (1.0)	4 (4.0)	0 (0.0)	4 (2.0)	3 (3.0)
Missing	1 (0.5)	1 (0.5)	0 (0.0)	26 (12.9)	25 (12.3)	10 (10.1)	28 (13.9)	36 (17.7)	24 (24.2)	21 (10.5)	27 (13.3)	11 (11.1)
Reporting probability	103 (51.5)	118 (58.4)	59 (59.6)	65 (37.1)	78 (43.8)	39 (43.8)	56 (32.4)	59 (35.3)	26 (34.7)	60 (33.3)	50 (28.4)	31 (35.2)

Appendix 17 Missing data in health economics analysis

TABLE 66 Description of economic variables in UK FROST

Variable	Description	Missing values (%)				Range	Mean	SD
		Total	ESP	MUA	ACR			
Baseline variables								
age	Age at trial entry	0	0	0	0	30 to 70	54.25	7.72
sex	Male or female	0	0	0	0	1, 2	63% female	
eq5d_B	EQ-5D-5L at baseline	1.79	4.04	0.99	1.48	-0.37 to 1.00	0.43	0.26
OSS_B	OSS score at baseline	0.40	0	0.50	0.49	1 to 48	19.89	8.25
Diabetes	Diabetic yes/no at baseline	0	0	0	0	1, 3	70% not diabetic	
alloc	Treatment allocation	0	0	0	0	1, 3		
Outcome variables for HRQoL								
eq5d_3m	EQ-5D-5L at 3 months	13.32	11.1	13.9	13.8	-0.245 to 1.00	0.60	0.26
eq5d_6m	EQ-5D-5L at 6 months	18.09	24.2	14.4	18.7	-0.257 to 1.00	0.70	0.23
eq5d_12m	EQ-5D-5L at 12 months	12.72	13.1	11.4	13.8	-0.328 to 1.00	0.73	0.26
Outcome variables for costs (£)								
Cost_ESP	Costs of ESP ^a	0	0	0	0	59.80 to 768.40	279.46	148.8
Cost_MUA	Costs of MUA ^a	0	0	0	0	259.20 to 972.00	424.81	115.5
Cost_ACR	Costs of ACR ^a	0	0	0	0	877.30 to 3082.30	2170.46	431.1
Cost_PPP	Costs of physiotherapy ^b	0	0	0	0	0 to 975.20	209.65	152.9
Cost_add	Additional treatments ^c	0	0	0	0	0 to 167.97	2.83	21.0
Cost_further	Further treatments ^d	0	0	0	0	0 to 1521.87	41.41	204.2

Variable	Description	Missing values (%)				Range	Mean	SD
		Total	ESP	MUA	ACR			
Cost_other	Other treatments ^e	0	0	0	0	0 to 668	7.18	49.42
Cost_crossovers	Treatment after crossover ^f	0	0	0	0	0 to 125.01	0.50	7.87
Cost_Hosp_INP	Inpatient costs regarding complications ^g	0	0	0	0	0 to 4926.24	32.85	312.1
Cost_Hosp_OUP	Outpatient costs regarding complications ^h	0	0	0	0	0 to 875.07	19.37	82.71
Cost_GP_pr	Costs of GP visits (surgery)	33.00	37.40	31.80	32.00	0 to 822.80	57.26	110.6
Cost_GP_phone	Costs of GP visits (telephone)	34.20	38.30	32.30	34.00	0 to 197.60	6.33	23.01
Cost Nurse_pr	Costs of practice nurse	36.40	40.40	34.30	36.40	0 to 75.95	2.10	6.54
Cost_Nure_dis	Costs of district nurse	33.80	37.40	32.80	33.00	0 to 380	1.94	21.69
Cost_Physio_c	Costs of district physiotherapist	33.40	35.30	32.80	33.00	0 to 1214.40	56.27	183.1
Cost_OT_c	Costs of occupational therapist	16.90	16.20	16.40	17.70	0 to 282	0.67	13.79
Outcomes for cost-effectiveness								
Total_QALYs	Total QALYs over 1 year	26.6	35.3	22.4	26.6	-0.225 to 0.979	0.66	0.207
Total Costs	Total costs over 1 year	40.50	44.40	38.80	40.40	0 to 5732.54	1372.36	1095.99

OT, occupational therapist.

a For those who had ESP/surgery (MUA/ACR).

b Costs of PPP for those who had surgery (MUA/ACR).

c Any treatments the patient received before/during receiving randomised treatment.

d Any treatments the patient received after completing randomised treatment.

e Any non-trial treatments the patient had if they did not start/complete their randomised treatment.

f Cost of further treatments following crossover.

g Hospital inpatient stay costs related to complications.

h Hospital outpatient costs related to complications.

TABLE 67 Number and proportion of patients with complete data by treatment arm

Time point	Treatment arm, n (%)		
	ESP (N = 99)	MUA (N = 201)	ACR (N = 203)
Complete: HRQoL			
Baseline	95 (95.96)	199 (99.00)	200 (98.52)
3 months	88 (88.89)	173 (86.07)	175 (86.21)
6 months	75 (75.76)	172 (85.57)	165 (81.28)
12 months	86 (86.87)	178 (88.56)	175 (86.21)
Overall	64 (64.65)	156 (77.61)	149 (73.40)
Complete: costs			
3 months	78 (78.79)	164 (81.59)	158 (77.83)
6 months	71 (71.72)	155 (77.11)	150 (73.89)
12 months	77 (77.78)	161 (80.10)	158 (77.83)
Overall	55 (55.56)	123 (61.19)	121 (59.61)
Complete: both HRQoL and costs			
3 months	76 (76.77)	161 (80.10)	154 (75.86)
6 months	68 (68.69)	152 (75.62)	144 (70.94)
12 months	75 (75.76)	159 (79.10)	157 (77.34)
Overall	46 (46.46)	117 (58.21)	116 (57.14)

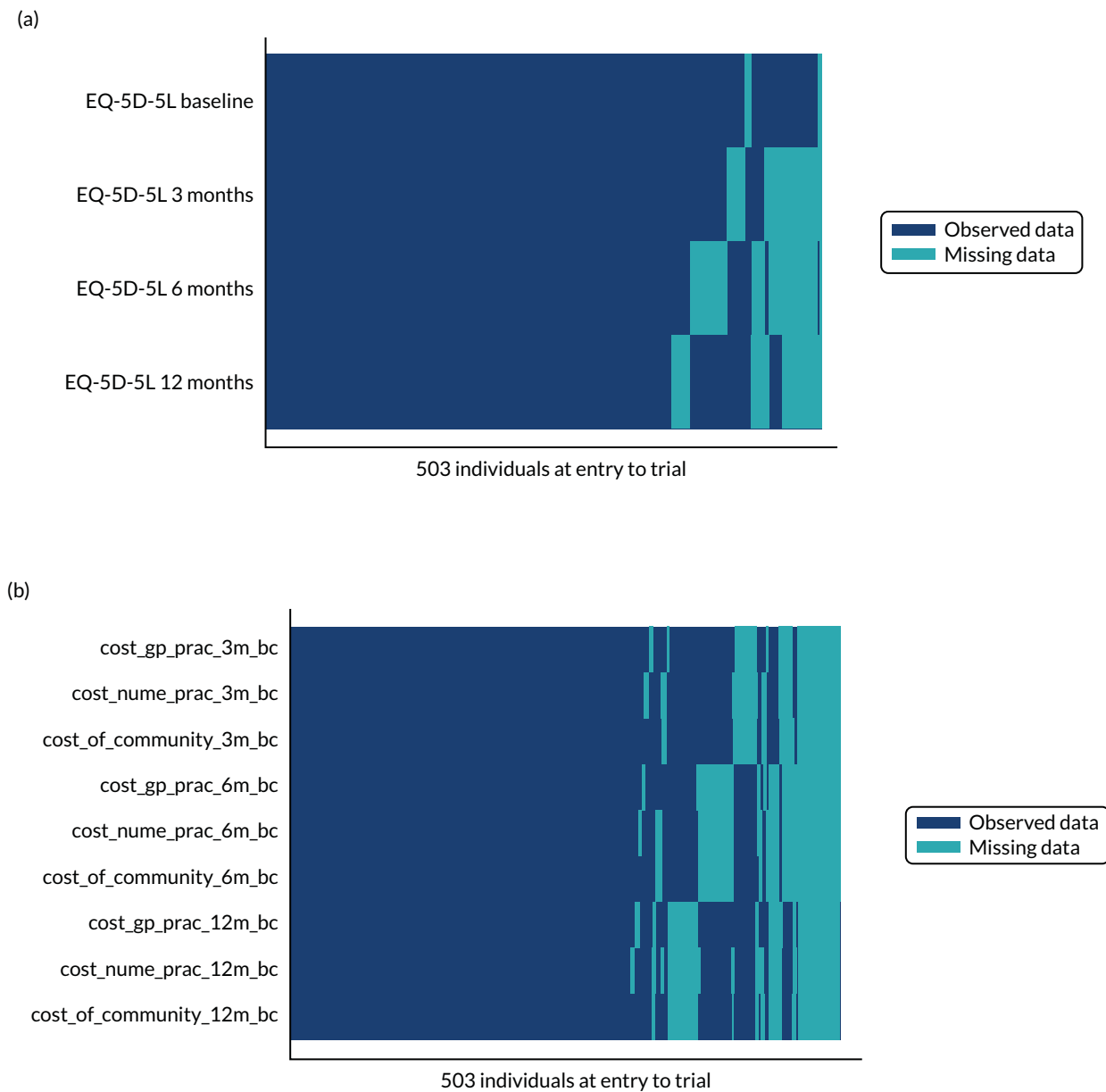


FIGURE 20 Pattern of missing data in UK FROST data set. (a) EQ-5D-5L data; and (b) cost data.

TABLE 68 Logistic regression for (1) missingness of costs and QALYs on baseline variables and (2) missingness between missing costs and QALYs and observed outcomes

	Odds ratio in logistic regression for missing data (95% CI)	
	Missing data on costs	Missing data on QALYs
Treatment allocation (MUA vs. ESP)	0.80 (0.48 to 1.32)	0.60 (0.34 to 1.05)
Treatment allocation (ACR vs. ESP)	0.85 (0.52 to 1.41)	0.71 (0.41 to 1.23)
Sex	1.26 (0.85 to 1.88)	0.87 (0.55 to 1.37)
Age	0.99 (0.97 to 1.01)	0.95 (0.93 to 0.98) ^a
Diabetes	1.11 (0.89 to 1.38)	1.06 (0.82 to 1.35)
EQ-5D-5L at baseline	0.28 (0.14 to 0.57) ^a	0.31 (0.14 to 0.67) ^a
QALYs at 3 months	0.003 (0.00 to 0.09) ^a	0.00 (0.00 to 0.50) ^a
QALYs at 6 months	0.007 (0.00 to 0.306) ^a	0.15 (0.0001 to 1.15)
Costs at 3 months	1.00 (0.99 to 1.00)	0.99 (0.99 to 1.00)
Costs at 6 months	1.00 (0.99 to 1.00)	1.00 (0.99 to 1.00)

a Statistically significant at 5% level.

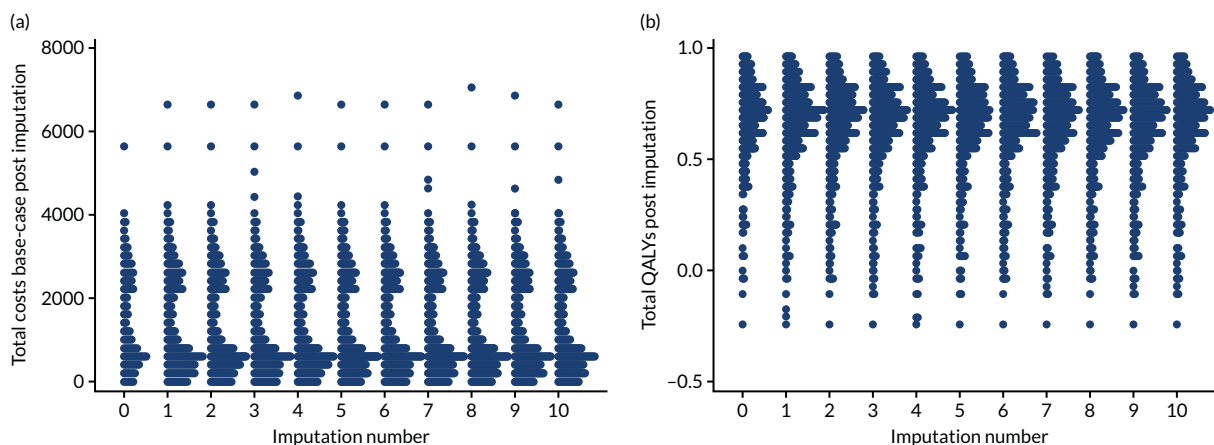


FIGURE 21 Comparison of the distribution of imputed values (imputations 1–10) with the observed data (imputation number 0) for (a) costs; and (b) QALYs. Individual values are represented by dots; the width of a row of dots represents the frequency of values in the distribution.

TABLE 69 Sensitivity analysis (scenario 6): summary for incremental analysis (ITT), cost-effectiveness results and uncertainty of different methods for handling missing data (MUA vs. ESP)

	Incremental cost (£) (95% CI)	Incremental QALYs (95% CI)	ICER (£ per QALY)	Probability cost-effective at £20,000/QALY (%)
MAR	276.507 (65.67 to 487.35)	0.0396 (-0.0008 to 0.0800)	6984	88
	228.605 (0.94 to 456.27)	0.0339 (-0.0138 to 0.0816)	6750	81
Same MNAR parameters in MUA and ESP				
-10% quality of life in both arms	228.605 (0.94 to 456.27)	0.0414 (-0.0041 to 0.0868)	5227	89
+10% cost in both arms	234.7271 (-6.91 to 476.36)	0.0339 (-0.0138 to 0.0816)	6935	80
-50% quality of life in both arms	228.605 (0.94 to 456.27)	0.0713 (0.0221 to 0.1206)	3204	99
+50% cost in both arms	259.2152 (-52.66 to 571.09)	0.0339 (-0.0138 to 0.0816)	7665	78
-10% quality of life and +10% costs in both arms	234.7271 (-6.91 to 476.36)	0.0413277 (-0.004 to 0.087)	5680	88
-50% quality of life and +50% costs in both arms	259.2152 (-52.66 to 571.09)	0.0710225 (0.0217 to 0.1203)	3650	98
Different MNAR parameters in MUA and ESP				
-10% quality of life in ESP	228.605 (0.94 to 456.27)	0.0559849 (0.010 to 0.102)	4083	96
-10% quality of life in MUA	228.605 (0.94 to 456.27)	0.0192851 (-0.0281 to 0.0667)	11,854	62
+10% cost in ESP	199.748 (-32.80 to 432.29)	0.0338503 (-0.0139 to 0.0816)	5901	82
+10% cost in MUA	261.540 (28.02 to 495.06)	0.0338673 (-0.0138 to 0.0816)	7722	79
-50% quality of life in ESP	228.605 (0.94 to 456.27)	0.144459 (0.101 to 0.188)	1582	99
-50% quality of life in MUA	228.605 (0.94 to 456.27)	-0.0390401 (-0.0895 to 0.0114)	-5856	3
+50% cost in ESP	84.318 (-171.7 to 340.42)	0.0337907 (-0.0139 to 0.0815)	2495	87
+50% cost in MUA	393.28 (130.9 to 655.60)	0.0338787 (-0.014 to 0.082)	11,608	71

EME
HS&DR
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 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