Title

Management of adults with primary frozen shoulder in secondary care: the UK FROST randomised controlled trial with economic evaluation

Keywords

Frozen shoulder; Physiotherapy; Manipulation under anaesthesia; Arthroscopic capsular release; Randomised controlled trial

Author list

Stephen Brealey[1], Matthew Northgraves[1], Lucksy Kottam[2], Ada Keding[1], Belen Corbacho[1], Lorna Goodchild[3], Cynthia Srikesevan[4], Saleema Rex[1], Charalambos Charalambous[5], Nigel Hanchard[6], Alison Armstrong[7], Andrew Brooksbank[8], Andrew Carr[4], Cushla Cooper[4], Joseph Dias[7], Ioana Donnelly[8], Catherine Hewitt[1], Sarah Lamb[4], Catriona McDaid[1], Gerry Richardson[9], Sara Rodgers[1], Emma Sharp[8], Sally Spencer[10], David Torgerson[1], Francine Toye[11], Amar Rangan[1,2,4]

[1] York Trials Unit, ARRC Building, Department of Health Sciences, University of York, York, Y010 5DD, UK

[2] The James Cook University Hospital, South Tees Hospitals NHS Foundation Trust, TS4 3BW, Middlesbrough, UK

[3] The Physiotherapy Practice, 17A Beach Road, South Shields, Tyne and Wear, NE33 2QA, UK

[4] Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Botnar Research Centre, Windmill Road, University of Oxford, Oxford, OX3 7LD, UK

[5] Department of Orthopaedics, Blackpool Victoria Hospital, Blackpool, FY3 8NR, UK, and School of Medicine, University of Central Lancashire, Preston, PR1 2HE, UK

[6] School of Health & Social Care, Teesside University, Middlesbrough, Tees Valley, TS1 3BX, UK

[7] University Hospitals of Leicester NHS Trust, Infirmary Square, Leicester, LE1 5WW, UK

[8] Glasgow Royal Infirmary, 84 Castle Street, Glasgow, G4 0SF, UK

[9] Centre for Health Economics, ARRC Building, University of York, York, Y010 5DD, UK
[10] Postgraduate Medical Institute, Edge Hill University, Ormskirk, Lancashire, L39 4QP, UK
[11] Physiotherapy Research Unit, Oxford University Hospitals NHS Foundation Trust, Oxford, UK, OX37HE, UK

Corresponding author
Prof Amar Rangan
Email: amar.rangan@york.ac.uk
Telephone number: 01642 854144

Conflicts of interest

Dr Kottam reports other grants from NIHR HTA during the conduct of this study. South Tees Hospitals NHS Foundation Trust receives educational grant to the department from DePuy Synthes (J&J limited). Institution also receives payment from DePuy Synthes (J&J limited) for Dr Kottam as a study Co-ordinator for the GLOBAL ICON Stemless Shoulder System Post Market Clinical Follow Up Study: CT 1401. These are outside and unrelated to the submitted work.

Professor Catherine Hewitt is a member of the NIHR HTA Commissioning Board.

Dr Catriona McDaid receives funding from the British Orthopaedic Association and is a member of the National Institute for Health Research Health Technology Assessment (HTA) and Efficacy and Mechanism Evaluation Editorial Board.

Professor Sarah Lamb reports grants from the NIHR Health Technology Assessment Programme during the conduct of the study and was a member of the following Boards:-
HTA Additional Capacity Funding Board 2012-2015
HTA Clinical Trials Board 2010-2015
HTA End of Life Care and Add on Studies 2015-2015
HTA Funding Boards Policy Group (formerly CSG) 2010-2015
HTA MNCH Methods Group 2013-2015
HTA Post-board funding teleconference (PG members to attend) 2010-2015
HTA Primary Care Themed Call board 2013-2014
HTA Prioritisation Group 2010-2015

NIHR CTU Standing Advisory Committee 2012–2016

Professor Rangan reports other grants from NIHR HTA, ORUK and H2020 during the conduct of this study. South Tees Hospitals NHS Foundation Trust receives educational grant to the department from DePuy Synthes (J&J limited). Institution also receives payment from DePuy Synthes (J&J limited) for Professor 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.

Abstract (499 words)

Background: Frozen shoulder causes pain and stiffness. It affects around 10% of people in their 50s and is slightly more common in women. Costly and invasive surgical interventions are used, without high-quality evidence.

Objectives: To compare the clinical and cost-effectiveness of three treatments in secondary care for adults with a frozen shoulder. To qualitatively explore their acceptability to patients and healthcare professionals.

Design: Multi-centre, pragmatic, three-arm, parallel, open randomised controlled trial (RCT) with unequal allocation (2:2:1). An economic evaluation and nested qualitative study.

Setting: Orthopaedic departments of 35 hospitals across the United Kingdom recruited from April 2015 with final follow-up in December 2018.

Participants: Adults (≥ 18 years) with unilateral frozen shoulder, characterised by restriction of passive external rotation in the affected shoulder to less than 50% of the opposite shoulder and plain radiographs excluding other pathology.

Interventions: Early Structured Physiotherapy (ESP) with a steroid injection. Manipulation under anaesthesia (MUA) with a steroid injection. Arthroscopic capsular release (ACR) followed by manipulation. Both surgical interventions were followed with post-procedural physiotherapy.

Main outcome measures: The primary outcome and end-point was the Oxford Shoulder Score (OSS) at 12 months post-randomisation. A difference of five points between ESP and MUA or ACR, or four points between MUA and ACR, was judged clinically important.

Results: The mean age of 503 participants was 54 years, 319 were female (63%) and 150 were diabetic (30%). The primary analyses included 473 participants (94%). At the primary end point of 12 months, participants randomised to ACR had on average statistically significantly higher (better) OSS than MUA (2.01 points, 95% confidence interval (CI) 0.10 to 3.91, p=0.04) and ESP (3.06 points, 95% CI 0.71 to 5.41, p=0.01). MUA did not have

statistically significant better OSS than ESP (1.05 points, 95% CI -1.28 to 3.39, p=0.38). No differences were deemed of clinical importance. Serious adverse events (SAE's) were rare but occurred in participants randomised to surgery (n=8 for ACR and n=2 for MUA). There was, however, one SAE in a participant who had non-trial physiotherapy. The base-case economic analysis showed that MUA was more expensive than ESP with slightly better utilities. The ICER for MUA was £6,984 per additional QALY and probably 86% cost-effective at the £20,000/QALY threshold. ACR was more costly than ESP and MUA, 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.

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.

Conclusions: None of the three interventions were clearly superior. ESP with a steroid injection is an accessible and low-cost option. MUA is the most cost-effective option. ACR carries higher risks and costs.

Future work: Evaluation in an RCT is recommended to address the increasing popularity of hydrodilatation despite the paucity of high-quality evidence.

Trial registration: ISRCTN48804508

Funding details: National Institute for Health Research Health Technology Assessment programme (project number 11/26/01).

Contents

Table of tables

Table of figures

List of Supplementary Material

Supplementary File 1 MUA Surgery Form

Supplementary File 2 ACR Surgery Form

Supplementary File 3 Frozen shoulder Information Sheet

Supplementary File 4 Frozen Shoulder Home Exercises

Supplementary File 5 Structured Physiotherapy Logbook

Supplementary File 6 Post-Procedural Physiotherapy Logbook

Supplementary File 7 52 week Complication Form

Supplementary File 8 Study Eligibility Form

Supplementary File 9 Trial Participant Information Sheet

Supplementary File 10 Trial Consent Form

Supplementary File 11 Baseline Form

Supplementary File 12 Consent Status Form

Supplementary File 13 Patient Preference Form

Supplementary File 14 Pre-treatment Form

Supplementary File 15 Change in Status Form

Supplementary File 16 Three Month Patient Questionnaire

Supplementary File 17 Adverse Event Form

Supplementary File 18 Serious Adverse Event Form

Supplementary File 19 Adverse Event Review Form

Supplementary File 20 Systematic Review Report

Supplementary File 21 Interview schedule_trial participants

Supplementary File 22 Interview schedule_health professionals

List of abbreviations

ACR	Arthroscopic Capsular Release
AE	Adverse Event
AUC	Area Under the Curve
CACE	Complier Average Causal Effect
CEAC	Cost-Effectiveness Acceptability Curve
CI	Confidence Interval/ Chief Investigator
CONSORT	Consolidated Standards of Reporting Trials
CRF	Case Report Form
DMEC	Data Monitoring Ethics Committee
ESP	Early Structured Physiotherapy
HEAP	Health Economics Analyses Plan
HRQoL	Health Related Quality of Life
HTA	Health Technology Assessment
ICER	Incremental Cost-Effectiveness Ratio
IQR	Interquartile Range
ITT	Intention-To-Treat
MAR	Not Missing At Random
MUA	Manipulation Under Anaesthesia
NHS	National Health Service
NICE	National Institute for Clinical Excellence
NMAR	Not Missing At Random
OSS	Oxford Shoulder Score
PPP	Post-Procedural Physiotherapy
PSSRU	Personal Social Services Research Unit
QALY	Quality Adjusted Life Years
RCT	Randomised Controlled Trial
SAE	Serious Adverse Event
SUR	Seemingly Unrelated Regression
TMG	Trial Management Group
TSC	Trial Steering Committee
UK	United Kingdom
UK FROST	United Kingdome FROzen Shoulder Trial

Plain English Summary (292 words)

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 one in ten people and is more common in women. Most patients are treated in the community. Those who do not improve, are offered treatments in hospitals. This includes costly and invasive surgical options. It's unclear which treatment provides better patient outcomes and is cost-effective.

UK FROST included 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, that 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 capsular release required further treatment and had slightly better shoulder function and pain outcomes than the manipulation procedure or early physiotherapy. This improvement, however, was unlikely to be of clinical benefit to patients.. Capsular release had slightly higher risks and substantially higher costs. Six serious complications were reported in patients who had a capsular release (mostly owing to co-existing health problems) and two in patients who had manipulation. Physiotherapy was the least expensive, but patients who had the manipulation had slightly better general health compared to physiotherapy. Early physiotherapy with steroid injection could be accessed quicker than the surgical alternatives. Manipulation costed more than physiotherapy but provided the most value for money. Patients in the study wanted early access to medical help to improve their shoulder problem.

Scientific Summary (2,3976 words)

BACKGROUND

Frozen shoulder occurs when the capsule, or the soft tissue envelope around the ball and socket shoulder joint, becomes inflamed then 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 United Kingdom (UK). When stiffness becomes more established, treatments include Physiotherapy with a steroid injection; Manipulation Under Anaesthesia (MUA); and Arthroscopic Capsular Release (ACR). With the intention of facilitating quicker recovery, more invasive and costly surgical interventions (MUA and ACR) are being used despite a lack of good evidence.

OBJECTIVES

To evaluate the effectiveness and cost-effectiveness of Early Structured Physiotherapy (ESP) versus MUA versus ACR for patients referred to secondary care for the treatment of primary frozen shoulder. A qualitative study to explore the acceptability of the different interventions to trial participants and health care professionals. A systematic review update to explore the trial findings in the context of existing evidence.

METHODS

Randomised Controlled Trial

Design

A pragmatic, multi-centre, superiority RCT comparing three parallel groups. The randomisation sequence was based on a computer generated randomisation algorithm provided by a remote randomisation service. Individual patients were allocated to MUA:ACR:ESP in the ratio of 2:2:1, stratified by the presence of diabetes, using random blocks sizes of ten and 15.

Eligibility criteria

Adults aged 18 year or older, presenting with a clinical diagnosis of frozen shoulder characterised by restriction of passive external rotation in the affected shoulder to less than fifty per cent 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 which necessitated hospital care; frozen shoulder secondary to other causes; contraindication of 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 National Health Service (NHS) hospitals in the UK across a range of urban and rural areas (April 2015 to December 2018).

Interventions

ESP: Up to 12 weekly sessions comprising essential 'focused physiotherapy' and optional supplementary physiotherapy. Focused physiotherapy included an information leaflet containing 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. Participants received supervised exercises and instructions on a graduated home exercise programme.

MUA: The affected shoulder was manipulated to stretch and tear the tight capsule and to improve range of movement. An intra-articular corticosteroid injection to the glenohumeral joint was used whilst the patient was under anaesthesia unless it was contraindicated. Post-procedural physiotherapy (PPP) was provided.

ACR: Arthroscopic release of the contracted rotator interval and anterior capsule was performed, followed by MUA to complete the release of the inferior capsule. Steroid injections were permitted at the surgeon's discretion. PPP was provided.

Outcome measures

The primary outcome was the Oxford Shoulder Score (OSS) at 12 months post-randomisation. The OSS is a 12 item patient reported outcome measure with a score range from 0 (worst) to 48 (best). This was also completed at three and six months post-

randomisation. Secondary outcomes, gathered at three, six and 12 months were the Quick-DASH (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 EuroQol 5 Dimensions (EQ-5D-5 L). Expected and unexpected complications and adverse events were also recorded.

Sample size

The minimum clinically important difference on the OSS was defined as a five point difference (standard effect size 0.42) between surgery and no surgery and four points (standard effect size 0.33) to distinguish between MUA and ACR. 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 OSS over 12 months and allowing for 20% attrition.

Analysis

Analyses were conducted for: ACR vs ESP, MUA vs ESP and ACR vs MUA using Stata Version 15 and two-sided statistical significance at the 0.05 level. The intention to treat (ITT) primary analysis was based on a linear mixed model incorporating OSS 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 OSS at baseline and included as further fixed effects: treatment arm, time, arm by time interaction, age, gender 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 endpoint of 12 months as well as an overall treatment effect over 12 months. The estimates are reported as mean differences between treatment groups with 95% confidence intervals and associated p-values. Continuous secondary outcomes were analysed using the same method as the primary outcome adjusting for the same covariates.

Pre-specified sensitivity analyses explored: the effect of non-compliance with ESP treatment using complier average causal effect (CACE) analysis; the effect of waiting times for intervention using additional data collected just prior and six months following treatment; the impact of missing data; and the effect of questionnaire return outside the intended follow-up time. The Data Monitoring Ethics Committee advised that employment status was included as a model covariate in a sensitivity analysis. Pre-specified sub-group 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 sub-group analysis for duration of symptoms at the time eligibility was confirmed.

Economic evaluation

Costs and health benefits were compared for the three groups over the 12 months. Hence discounting was not required. All costs were expressed in UK £ sterling at a 2017-2018 price base. Health outcomes were assessed in terms of quality-adjusted life years (QALYs), based on patients' health related quality of life (HRQoL) outcomes obtained from trial participants using the EQ-5D-5L at baseline, three months, six months and 12 months. Differences in mean costs and mean QALYs at 12 months were used to derive the incremental cost-effectiveness ratio (ICER) for surgery and non-surgical treatment. The base-case analysis was conducted on an ITT 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 semi-structured with open questions; 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 and reviewed by a second researcher. Data from trial participants was mapped against the International Classification of Functioning, Disability and Health framework.

Systematic review update

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 the previous HTA review. RCTs evaluating MUA, ACR, hydrodilatation or physiotherapy plus a steroid injection for treatment of primary frozen shoulder compared to 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, SD and number of participants) for each group was extracted, where available. The standardised mean difference (SMD) 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. A narrative and tabular summary 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 due to limited data and methodological and statistical heterogeneity.

RESULTS

Randomised Controlled Trial

Of 914 patients screened, 503 were randomised: MUA (n=201), ACR (n=203) and ESP (n=99). Follow-up rates were between 85% and 89%, and no evidence of differential dropout across the treatment arms. The primary analysis included all participants with OSS outcome data at one or more follow-ups (94%). Average shoulder function improved in all arms, with many participants (24%) regaining function to the top OSS at 12 months.

At the primary end point at 12 months, participants randomised to ACR had on average statistically significantly higher (better) OSS than MUA (2.01 points, 95% CI 0.10 to 3.91) and ESP (3.06 points, 0.71 to 5.41); there was no statistically significant difference between MUA and ESP 1.05 (-1.28 to 3.39).

For the short term follow-up at three months, ACR had lower (worse) OSS compared with the other two interventions (versus MUA -3.36 (-5.27 to -1.45); versus ESP -4.72 (-7.06 to -2.39)). There was no evidence for statistically significant differences in average OSS over the 12 months follow-up: MUA versus ESP 0.61 (-1.31 to 2.53); ACR versus ESP -0.23 (-2.15 to 1.70) and ACR versus MUA -0.84 (-2.41 to 0.72).

Mean differences were short of the minimal clinically important effect size of four (ACR versus MUA) to five (ACR or MUA versus ESP) OSS points. However, differences of that magnitude were included in the 95% CIs for the benefit of MUA and ESP compared with

ACR at three months, and ACR over ESP at 12 months. Sensitivity analyses did not substantially alter the results. There were no significant sub-group 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, outcomes for ESP compliers remained lower than those who complied in both surgery arms combined (-1.84 OSS points, 95% CI -4.41 to 0.74, p=0.157).

Of the secondary outcomes, QuickDASH and shoulder pain results followed a similar pattern to the OSS, in that statistically significant poorer outcomes were observed for ACR participants at three months but better outcomes at 12 months compared with MUA or ESP. There were no statistically significant differences between treatment arms in response to the global question on the extent of recovery.

In total, there were only ten serious adverse events (SAEs), reported for nine participants. All SAEs occurred in the surgical arms (n=8 for ACR and n=2 for MUA) although one participant in the ACR group had an SAE from non-trial physiotherapy. The events mainly related to serious medical complications such as chest infection or stroke. There were 33 non-serious adverse events, reported for 31 participants and were mainly expected and often related to persistent or worsening shoulder pain. There was no evidence for statistical differences in the proportion of non-serious adverse events (p=0.186).

Economic evaluation

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 [mean 0.0396 more QALYs per participant than ESP (95% CI -0.0008 to 0.0800)]. The resulting ICER for MUA was £6,984 per additional QALY. ACR was substantially more costly than ESP [on average £1,733.78 more per participant (95% CI 1,529.48 to 1,938.06)] for a slight benefit in utilities [mean 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 had slightly lower QALYs. MUA was the intervention most likely to be cost-effective at a £20,000 per QALY threshold (MUA 86% > ESP 14% > ACR 0%).

Qualitative Study

There were 44 interviews (mainly by telephone) undertaken with trial participants, 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 on getting their shoulder problems resolved which motivated them to participate in the trial. They thought that seeking early medical help and 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 had ACR described quicker recovery than they 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.

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 participants and mostly single centre studies. All studies were at high risk of bias to blinding of participants and clinicians, and outcome assessment. Due to considerable heterogeneity of the interventions and generally limited evidence for many of the comparisons, 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 ACR or Physiotherapy. The pooled effect favoured ACR, however, the second study provide little additional weighted information.

CONCLUSIONS

UK FROST has provided robust clinically relevant evidence that none of the three treatments were clearly superior on 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 when compared to the surgical interventions. MUA is the most cost-effective option with an ICER of £6,984 per additional QALY. Patients who receive ACR are least likely to need further treatment, but ACR is associated with relatively higher risks and costs.

To address the increasing popularity of hydrodilatation, and paucity of rigorous evidence for hydrodilatation's effectiveness, a high-quality RCT is recommended to compare hydrodilatation versus ESP with steroid injection versus MUA with steroid injection.

Trial registration

ISRCTN48804508.

Funding

This study was funded by the National Institute for Health Research Health Technology Assessment programme (project number 11/26/01).

Chapter 1 Introduction

Frozen shoulder

Frozen shoulder (also known as Adhesive Capsulitis) occurs when the capsule, or the soft tissue envelope around the ball and socket shoulder joint, becomes inflamed 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, neuro-surgery 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 as being similar to Dupuytren's disease.[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 four years or more from onset, around 40% of patients can have from mild to severe symptoms.[4] Figure 1 below illustrates the pathology of a frozen shoulder.

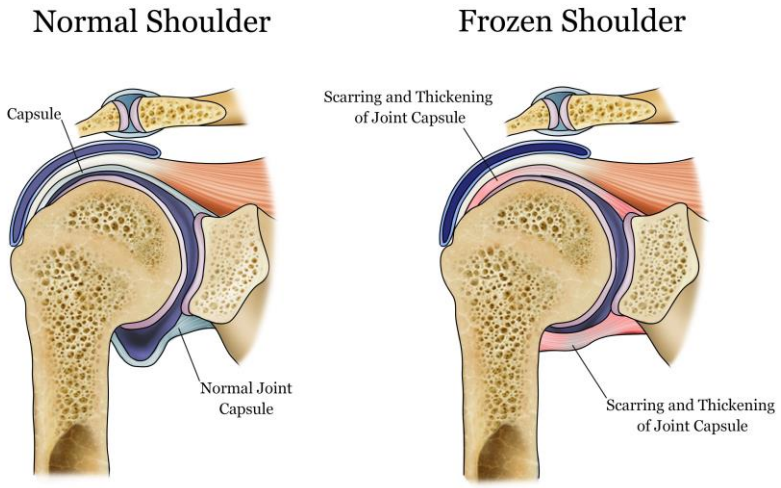


Figure 1: Diagram showing site of pathology of frozen shoulder. The image has been re-used from https://www.local-physio.co.uk/articles/shoulder-pain/frozen-shoulder/ with permission from the Copyright holders

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

(a) Painful phase, which may last three to nine months

(b) Adhesive phase, with stiffness lasting for four to six months

(c) Phase of resolution or 'thawing', lasting for five to 24 months.

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

The cumulative incidence of frozen shoulder is estimated at 2.4 per 1000 population per year based on a Dutch General Practice (GP) sample.[7] It most commonly affects individuals in their sixth decade of life and a large primary care based study in the United Kingdom (UK) found that frozen shoulder affected 8.2% of men and 10.1% of women of working age.[8] In contrast, an incidence of 1% has been reported by a UK shoulder surgeon in his hospital specialist practice.[9] This discrepancy in estimated prevalence can be explained by different populations used as the denominator 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

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] X-rays are reported as not being 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.[14] [15-18] 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.[19] This accords with other studies, which have found that GPs in the UK and United States of America (USA) lack confidence in making shoulder diagnoses.[20, 21].

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.[22] Generally, less invasive treatments are provided in a primary care setting in the UK for the earlier phases of the disease, particularly for control of pain. These may include oral analgesia; Physiotherapy; Acupuncture; and Glucocorticoid (steroid) injection.[22] Treatments utilised in secondary care, when stiffness becomes more established, were confirmed by a UK survey of health professionals conducted in 2009 as Physiotherapy; Manipulation Under Anaesthesia (MUA); and Arthroscopic Capsular Release (ACR).[23]

Physiotherapy treatment includes combinations of advice, exercises, therapist-applied mobilisation techniques, and thermo- and electrotherapies. The modalities of treatment recommended for use are described within the UK national physiotherapy guidelines for

frozen shoulder, which were based on a systematic review.[6] These are either provided in isolation, or as a supplement to other interventions such as intra-articular injection of corticosteroid or surgical interventions (MUA or ACR). Intra-articular corticosteroid injection helps improve 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 benefit with better improvement in pain, function and range of movement (up to 6-7 weeks) compared to placebo[13] and probably compared to isolated manual therapy and exercise.[6]

MUA is a procedure performed by the surgeon 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. The 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.

ACR is a 'keyhole' surgical procedure performed under general anaesthesia. The key holes 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 an MUA to complete and confirm full release of the contracted capsule. The ACR is also followed by physiotherapy for mobilisation of the arm and shoulder to restore mobility and function.

Rationale for the UK FROzen Shoulder Trial (UK FROST)

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 to other treatment options, [24, 25] and there is only limited RCT evidence available on ACR.[13, 26]

Systematic reviews have identified large gaps in the evidence-base and uncertainty in the effectiveness of treatments for frozen shoulder and a need for high quality primary research.[27] In a systematic review commissioned by the Health Technology Assessment (HTA)

Programme, 28 RCTs, one quasi-experimental study, and two case series were included.[13] The review found there were insufficient studies with a similar intervention and comparator to quantify effectiveness. Most studies had a high risk of bias, did not report adequate methods for 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 in 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 lack of good evidence.[13] There is a clear need for a well-designed high quality RCT to determine the effectiveness and cost-effectiveness of commonly used interventions for the treatment of frozen shoulder.

The findings of a national survey of health care professionals in the UK, conducted in 2009, were used to determine which are the most commonly used interventions that needed testing within an RCT in a secondary care setting. Physiotherapy, MUA and ACR were the more frequently used interventions that were recommended by healthcare professionals to be compared in a RCT. Only 6% of respondents at the time suggested Hydrodilatation as a comparator they could use in a trial, which did not make this a feasible intervention to test within 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 costlier surgical interventions i.e. MUA and ACR.[23] It is important to emphasize that whilst 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 patient experiences of a frozen shoulder is also limited,[19] participants were interviewed about their experience and acceptability of treatment, as were health professionals (physiotherapists and surgeons).

Aim and objectives

The strategic aim of UK FROST, underpinned by the key treatment uncertainties, was to provide evidence of the clinical and cost-effectiveness of three common interventions currently provided within the UK National Health Service (NHS) for the treatment of frozen shoulder in a hospital setting. The following objectives were defined in order to achieve this overarching aim:

1. The primary objective was to determine the effectiveness of ESP versus MUA versus 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) being 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 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 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 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 will be in addition to publishing the results of the study in key journals and publishing the HTA report.

Chapter 2 Trial design and methods

This chapter describes the trial design and methods to address the objectives about the clinical effectiveness of the healthcare interventions being compared. The methods of the health economic evaluation and the nested qualitative study are described in their respective chapters. The trial protocol has been published.[28]

Trial design

This was a pragmatic, multi-centre, stratified (diabetes present or not), superiority trial comparing three parallel groups (MUA versus ACR versus ESP, with unequal allocation [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.[29] To minimise diagnostic uncertainty, the clinical examination included the key diagnostic assessment of restriction of passive external rotation in the affected shoulder[30] for which there is evidence of good inter-rater agreement on whether restriction is present[31] and a high threshold (50% restriction) for inclusion. Plain radiographs (antero-posterior and axillary projections) were obtained routinely for all patients to see whether they were normal and could exclude glenohumeral arthritis and other pathology that could lead to similar clinical presentation (e.g. locked posterior dislocation).

Inclusion criteria

Patients, including diabetics, were eligible if:

- they were aged 18 year or older;
- they presented with a clinical diagnosis of frozen shoulder characterised by restriction of passive external rotation in the affected shoulder to less than 50% of the contralateral shoulder; and
- they had radiographs to exclude other pathologies.

Exclusion criteria

Patients were excluded if:

- they had a bilateral concurrent frozen shoulder;
- their frozen shoulder was secondary to trauma which 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; or
- they lacked the mental capacity to understand the trial.

Setting

The trial recruited from the orthopaedic departments of 35 National Health Service (NHS) hospitals in the UK across a range of urban and rural areas. This included 28 hospitals in England, six in Scotland and one in Wales. There were two additional hospitals in England that 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 have been listed in Appendix 1 .

Interventions

The components and standardisation of the surgical trial interventions were informed by a survey of 53 surgeons who were Principal Investigators (PIs) for two multi-centre shoulder surgical RCTs.[32, 33] The stand-alone physiotherapy and the post-procedural physiotherapy programmes were developed using evidence from a systematic review,[13] UK guidelines,[6] previous surveys of UK physiotherapists[34, 35] and consensus from expert shoulder physiotherapists in secondary care derived from a Delphi survey, which was specific to UK FROST.[19] Ethics approval for the latter was obtained from the School of Health and Social Care Research Governance and Ethics Committee of Teesside University on 23 May 2014 (REC reference 069/14). The development of the physiotherapy programmes are available on-line.[36] It is important to emphasize that whilst 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 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 pre-operative screening. In keeping with NHS waiting time targets, both surgical procedures were expected to be performed within 18 weeks of randomisation. These would be under general anaesthetic and were expected to be 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 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 number of frozen shoulder patients 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 used whilst the participant was under the same anaesthetic unless it was contra-indicated at the time of surgery. Post-operative analgesia including nerve blocks were provided as per usual care in the treating hospital. The details of the 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 intra-operatively to capsular release. The need for this was to be reviewed at another clinic appointment to allow assessment of outcome of the MUA and the need for any further intervention. Details of any further intervention were collected prospectively.

Arthroscopic capsular release with MUA

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 like posterior capsular release or subacromial decompression were permitted at the discretion of the operating surgeon. Steroid injections, which slightly increase the risk of infection and

morbidity, were permitted at the surgeon's discretion.[37] Post-operative analgesia including nerve blocks was provided as per usual care in the treating hospital. The details of ACR were collected prospectively on 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 six 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 with the following objectives:

1. To determine molecular and cellular abnormalities in tissue obtained during surgery in 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 (RN) posted a letter about the nested study, a patient information leaflet and a consent form. Written informed consent was performed 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 University of Oxford Musculoskeletal Biobank, and housed at the Botnar Research Centre for formal analysis. The biopsy material was small (2mm by 2mm), obtained with the use of arthroscopic graspers and not expected to have any significant effect on patient outcomes. The results of this study have been published.[38]

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*) containing education, 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 improves.[39, 40] Participants received supervised exercises and were provided instructions on 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 focussed physiotherapy package unless there was a specific clinical reason for them not to do so (e.g. a steroid injection might be withheld in a participant with currently uncontrolled diabetes; or in a participant with a stiff, but painless, and non-irritable shoulder).

Supplementary physiotherapy comprised of those interventions that were not essential, but which were 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 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 at each session (e.g. injection, advice and education, gentle active exercise) was recorded in the Structured Physiotherapy logbook (see *Report Supplementary Material 5*).

Post-procedural physiotherapy (PPP)

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 PPP differed from ESP in order to suit its very different context. As the research literature was uninformative, two essential 'focused physiotherapy' interventions were pre-specified based on established good practice. These were:

1. The provision of an information leaflet containing education, advice on pain management and function;
2. Instructions on 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 were interpreted as for ESP, provided optional, supplementary interventions. A steroid injection was to be avoided where possible during PPP. The PPP logbook (see *Report Supplementary Material 6*) was used to record the PPP given at each session.

Steroid injections

Steroid injections were administered with or without imaging guidance depending on the usual practice of the hospital site. Current evidence did not support the superiority of either approach.[41]

Modifications to interventions

There were no explicit criteria to modify, discontinue or crossover from the assigned trial treatment. The clinician and participant discussed whether to continue with the assigned treatment for reasons such as poorly controlled diabetes or the treatment no longer being required.

Adherence to interventions

Adherence to the trial treatments was explained in the Trial Site Manual and during Site Initiation Visits (SIVs). A requirement of the internal pilot was to check the feasibility of delivering the ESP programme. This was extended 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 by the Chief Investigator (CI), a Consultant Orthopaedic Surgeon, and the Lead Physiotherapist for treatment adherence who decided whether any action was required with a site. This was further monitored by the Trial Management Group (TMG), independent Trial Steering Committee (TSC) and Data Monitoring Ethics Committee (DMEC).

Concomitant care

The use of analgesia to ensure 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 in the management of a participant awaiting surgery. Specific home exercise programmes like 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 Oxford Shoulder Score (OSS), a patient-reported measure of functional limitation following shoulder surgery. The development and validation included patients with frozen shoulder[42] and 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).[43] It has been validated against the professionally endorsed Constant Score[44] and the SF-36 and responsiveness over a six month period following surgical intervention has been established.[45]

The OSS was completed by the participant at the hospital at baseline prior to randomisation. The questionnaire was posted to the trial participants at 3, 6 and 12 months after randomisation. The primary endpoint 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 collected at the hospital at the start of treatment. This was either the day of the operation or, for participants allocated to ESP, on the day when the steroid injection was given or at the first visit to physiotherapy, depending on which was first. The OSS was then posted to participants to complete 6 months from when treatment started.

Secondary outcomes

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

Quick Disabilities of Arm Shoulder and Hand (QuickDASH)

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, was used.[47] Scored from 0 to 100, an 8-unit improvement in scores has been defined as the minimum clinically important difference for patients with shoulder problems.[48] Validity and responsiveness for frozen shoulder has been established.[49]

EuroQol 5 Dimensions (EQ-5D-5 L)

The 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 5L version consists of the same five domains as the original EQ-5D-3L but with five levels rather than three to help overcome problems with ceiling effects and improve sensitivity.[52, 53] The EQ-5D-3L has been validated for a range of shoulder conditions.[54, 55] The 5L version 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 Numeric Rating Scale for pain,[56] a single 11-point numeric scale with 0 representing 'no pain' and 10 representing 'worst possible pain', a measure considered the most valid for this population.[57]

Extent of recovery

A simple subjective global question asked to what extent the participants' frozen shoulder symptoms in the past 24 hours affected their assessment of needing treatment. This informed the extent of resolution of symptoms 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 12 month 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 two 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. Additionally nerve, blood vessel, tendon or bone injury and complications related to steroid injection, including steroid flare and septic arthritis, were recorded.

Adverse events

Non-serious Adverse events (AEs) were classified as any untoward medical occurrence in a trial participant related to the affected shoulder up to 12 months from randomisation. Serious

AEs (SAEs) were 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 not listed above which may require medical or surgical intervention to prevent any of the above from occurring.

Sample size

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

There are data to suggest 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 of 12 points across different populations. This larger effect size was required to justify the greater costs and potential risks associated with surgery when comparing ESP with MUA and ESP with ACR.[43] A smaller difference of 4 points on the OSS (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: 200, ACR: 200, ESP: 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 the data contributed to the final analyses, was performed to confirm the feasibility of the trial and is explained below.

Internal pilot study

There were two phases to the internal pilot study.

Phase 1 (months 4 to 9)

It was important to critically test our assumptions after six months of recruitment by reviewing: the number of sites set up; and 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 was to review: a) whether the participating sites were being provided with enough training and documentation; b) the number of reasons why patients were not eligible for the trial; c) the length of time it took to consent a patient and reasons for not taking part; d) whether all clinicians at a site were actively taking part in the trial, and, if not, why not; and e) 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 six months who had recruited 24 patients i.e. evidence that sites could recruit the expected one participant/month.
- To ensure adequate progress was made with setting up other sites to recruit, to have twelve sites set up i.e. 50% of sites.

Phase 2 (months 10 to 27)

This phase of the internal pilot continued for a further 18 months and was reviewed at six monthly intervals with the independent oversight committees. Patients were likely to have already suffered with a frozen shoulder for several months and received physiotherapy in primary care before referral to hospital. There was concern that this could impact on patient consent and adherence to the ESP intervention which were threats to both the feasibility and validity of the trial. Evidence from simulation work found 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 has 90% power and effect sizes of around 0.3 to 0.4, 20% to 30% non-compliance in the ESP group 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 group 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 was also monitored:

- reasons for patient non-consent into the trial, their treatment preferences and to informally discuss this with willing patients;
- whether all 25 sites were set up and had recruited 250 patients (50% of our target); and
- monitor waiting times at sites from randomisation to intervention and to consider the need to substitute 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 2009/2010 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 50% of frozen shoulder patients presenting in secondary care met the inclusion criteria, and of whom 40% 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 one 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 recruitment was achieved to target.

Patients who had been referred for a frozen shoulder to an outpatient hospital clinic were identified by the RN or assessing clinician. In the clinic, a designated individual within the shoulder team (e.g. surgeon, physiotherapist) completed the Study Eligibility Form (see *Report Supplementary Material 8*) to confirm whether the patient was eligible; and when applicable, approached the patient about the study. The RN then provided an information sheet (see *Report Supplementary Material 9*) and answered any questions. The patient was able to consent at that time or take up to a week to decide. When the patient consented (see *Report Supplementary Material 10*), the patient was asked to complete the Baseline Form (see *Report Supplementary Material 11*). The RN completed the Consent Status Form (see *Report Supplementary Material 12*) to confirm status.

When the patient did not consent, a further section on the Consent Status Form was completed by the RN to briefly record the reason given and the treatment plan. 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 SIV and a Trial Site Manual was prepared which included guidance on consenting patients into the trial and how to answer questions that might arise during consent. 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 PIs and RNs 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 (YTU), University of York. The unit of randomisation was the individual patient who were allocated to the trial interventions MUA:ACR:ESP in the ratio of 2:2:1, stratified by the presence of diabetes,[62] using random blocks sizes of 10 and 15. The RN 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 RN then informed the treating clinician and 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 about treatment allocation after randomisation. The statistician was blind to group 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 vs ESP, MUA vs ESP and ACR vs MUA according to the principle of intention to treat. All analyses were

conducted in Stata Version 15[63] 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, gender, 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, randomisation to follow-up and analysis of the trial was presented in a CONSORT flow diagram.

Baseline characteristics

All participant baseline characteristics were summarised descriptively by trial arm, both for participants 'as randomised' and 'as analysed'. The 'as analysed' population comprised all participants included in the primary analysis (i.e. patients who have complete data for the baseline covariates and outcome data for at least one post-randomisation time point). No formal statistical comparisons were undertaken between groups. Continuous measures were summarised using n, mean, standard deviation, median, minimum and maximum, while 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, optimal or sub-optimal 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 group. 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. Grades/bands and experience of treating surgeons and physiotherapists were presented.

Missing data

Item-level missing data for individual outcomes (OSS and QuickDASH) was 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 close proximity for some participants, OSS data from these CRFs were used as a substitute if data for one was available 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 were illustrated graphically.

The primary analysis was conducted on intention to treat (ITT) basis, including patients in the groups to which they were randomised. The primary analysis compared OSS between treatment groups 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, gender 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 multi-centre trial, surgeons or physiotherapists were not specifically adjusted for.

For the modelling of repeated measurements, the best fitting (based on AIC and BIC information criteria), 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 endpoint 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 groups with 95% confidence intervals and associated p-values.

Data were assumed missing at random. 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 non-compliance with participants' allocated treatment, secondary Complier Average Causal Effect (CACE) analysis was carried out. This retains the initial randomised assignments but overcomes problems of per protocol analysis. Given the three active treatments under investigation with different adherence criteria and multiple alternative treatment pathways for each participant, not all comparisons were suitable for CACE analysis. Therefore only compliance with ESP (minimum of 8 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 group 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 three- and six-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 three, six 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 more than 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 excluding data from such questionnaires.

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. Upon their 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.

Sub-group Analyses

In order to explore differences in treatment response for different participant populations, three planned exploratory sub-group analyses were conducted. One exploring the influence of whether the participant was diabetic (yes/no), one exploring whether 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 Trial Steering Committee proposed a further sub-group analysis based on the length of frozen shoulder symptoms at baseline (using the median of less/more than nine months as cut-off). For each analysis, a treatment group by sub-group 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 sub-group / treatment group pairing.

Analysis of Secondary Outcomes

QuickDASH, Pain, Extent of Recovery

Continuous secondary outcomes were reported descriptively (unadjusted mean, standard deviation, 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 for 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 adverse event data, these data sets were reviewed, and a single list of serious and non-serious adverse events compiled to avoid duplication in reporting. These events were then summarised by type for each treatment group. A logistic regression model was used to determine treatment group differences in having experienced at least one adverse event if the number of participants with one or more events exceeded 10 in each arm. The same covariates 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 descriptively explored by trial arm as well as for patients who had and had not received prior physiotherapy and patients who did and did not receive their allocated intervention. Any change in preferences were explored by tabulating participant preferences at 12-month follow-up against baseline preferences and against their allocated treatment.

OSS 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.

OSS Subdomains

Exploratory and confirmatory factor analysis of OSS from a population of patients with rotator cuff tears in the UKUFF trial identified reliable OSS subdomains of pain (items 1, 8, 11 and 12) and function (items 2, 3, 4, 5, 6, 7, 9, 10).[64] To explore the nature of shoulder outcomes further, descriptive statistics and associated graphs were presented for OSS sub-domains of pain and function by allocated arm at each time point.

Outcomes for Participants Receiving No Treatment

OSS and QuickDASH scores were summarised descriptively at baseline and all follow-up points for participants who did and did not receive any treatment as indicated on their change in status form (See *Report Supplementary Material 15*). Where available, 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 management of the frozen shoulder was updated.[13] The updated review focussed 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 undertaken to inform the design of UK FROST.[65] 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. The review protocol has been registered (PROSPERO 2019 CRD42019122999).

Data Management

A central database at YTU was used to manage data collection including the sending out and return of participant questionnaires (see *Report Supplementary Material 16*) and hospital CRF. This included automated email 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 YTU, 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 RN 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 regarding data quality of the postal questionnaires were made on return to YTU although a Trial Co-ordinator checked whether the participant had given extreme responses to either the last EQ5D-5L question and/or given a free text response to indicate the participant could be at harm. When this occurred, the PI, RN and CI were notified by email. After this initial check, all postal questionnaires and hospital CRFs passed through a process of scanning in the Teleform software, second checking and validation against predetermined rules.

Active and systematic follow-up of all randomised participants by post included pre-notification reminders, two- and four-week letter reminders and the option to complete an abridged questionnaire (a minimum of the OSS and EQ-5D) via telephone after 6 weeks. At 12 months, the primary endpoint, an unconditional incentive of £5 was included. If the patient agreed at the time of consent, text messages were sent on the day that the participant was sent the postal questionnaire[66] and newsletters were circulated to trial participants.[67] Trial participants could entirely withdraw from the study at any time for any reason but any data collected up to that point was included in the analysis. The participant could agree to being withdrawn from only postal questionnaire collection or only hospital CRF collection.

Essential trial documentation were kept with the Trial Master File and Investigator Site Files allowing the evaluation of the conduct of the trial and quality of the data produced. The documentation will be retained for a minimum of five years after the conclusion of the trial. The postal questionnaires and hospital CRFs will be stored for a minimum of five years after the conclusion of the trial as paper records; and a minimum of 20 years in electronic format.

Adverse Event Management

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

Once received, causality and expectedness was determined by the CI. SAEs that were unexpected and related to the trial were notified to Research Ethics Committee (REC) within 15 days for a non-life threatening event and within seven days for a life-threatening event. For non-serious AEs, the central office were notified within five days of the event being known. All (S)AEs were reported to the DMEC, TSC and TMG. Expected adverse events 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 like nerve, tendon, bone or joint; recurrent stiffness requiring further treatment; transient hyperglycaemia, steroid flare or joint sepsis following corticosteroid injection; or injuries related to heating or cooling of tissues. Follow-up reports a month later (see *Report Supplementary Material 19*) were reviewed by the CI to ensure that adequate action has been taken and progress made.

Ethical approval and monitoring

Ethics committee approval and any changes to the project protocol

NRES Committee North East – Newcastle and North Tyneside 2 approved the study on the 18 November 2014 (REC reference 14/NE/1176). Health Research Authority (HRA) 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 REC approval have been listed (see Appendix 2).

Trial management group (TMG)

The day-to-day management of the trial was overseen by the TMG who met on a quarterly basis. 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 (S)AEs and reasons for participant withdrawal. The review of progress was undertaken at a

participating site level and, as necessary, feedback was given to the PI and Research Nurses at each site.

Trial steering committee (TSC)

A TSC was appointed by the funding body to provide overall supervision for the trial and to advise on its continuation. Membership is listed in the Acknowledgement section.

Data monitoring ethics committee (DMEC)

The DMEC was appointed by the funding body with access to the unblinded comparative data as provided by a statistician at YTU 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. Membership is listed in the Acknowledgement section.

Patient and public involvement

Two patients who had previously received treatment for a 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, patient facing data collection forms and the consent process for trial participation. The need to develop a leaflet to provide general information about what is a frozen shoulder was identified following a qualitative study of patients with frozen shoulder using semi-structured interviews.[19] 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 of the meetings when necessary. The study steadily stayed on target with recruitment and 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 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 frozen shoulder patients 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 for various platforms. This includes contributing to the lay summary for this report, journal publications, web-based outputs such as updating the entry about management of the frozen shoulder on Wikipedia

and helping develop content for other appropriate webpages. These patients will also meet with local (SHRUG –shoulder research users group) and national shoulder patient groups (BESS patient liaison group) to ensure the current evidence base for treatment options are available and appropriately disseminated 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. Then summarises recruitment, the flow of participants through the trial, characteristics of participants at baseline, results of analyses of the primary and secondary outcomes, as well as the integration of findings into existing literature.

Summary findings of the internal pilot

The objectives of Phase 1 of the internal pilot (months 4 to 9) was to have a minimum of four sites recruiting during the six months who had recruited 24 patients i.e. evidence that could recruit the expected one participant/month. To ensure adequate progress had been made with setting up sites to recruit, twelve sites were to be set up i.e. 50% of total sites. At the end of month nine, we had recruited 20 patients (i.e. 83% of the target). This was in spite of not starting recruitment until month 7 and only having three of the four pilot sites set up. There were, however, two sites at which we were waiting on approval, and 16 of 26 sites with whom 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 there were four patients who met the exclusion criteria, ten non-consenting patients, and 20 consenting patients. Early data illustrated that all the reasons for non-consent were due to treatment preferences, rather than patients being too busy or not wanting to be involved in research. The time taken to consent was ranging from a minimum of 15 minutes to a maximum of one 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 he didn't see a sufficient number of patients to take part, and a further 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 group did not exceed 20% to 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 had either 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 i.e. the participant had attended eight sessions or more (n=19); or had attended fewer than eight sessions but both the participant and physiotherapist were satisfied with their progress (n=9); or the patient attended fewer than eight sessions and declined to attend further 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% to 30%. The oversight committees agreed that this was acceptable non-compliance and that the trial should continue with all three treatment groups.

Another aspect of the feasibility of the trial that was reviewed at month 27 was non-consent into the trial. This is because there was concern that patients would often have already had physiotherapy in primary care and 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 the key hole surgery (45%, n=67). For patients who were randomised into the trial, the majority had no treatment preference (53%, n=159), over a third preferred surgery (39%, n=116) and the remainder preferred physiotherapy (8%, n=25). Despite the preferences for surgery, this did not 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 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 were prioritised the trial participants for the surgical procedures.

In short, UK FROST was being delivered on time and to target, with an acceptable amount of non-compliance in 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 key molecular processes and changes seen in 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 NRES Committee, Newcastle and North Tyneside (14/NE/1176), appropriate informed consent was sought from UK FROST participants randomised to receive the Arthroscopic Capsular Release intervention. For a small sample of 16 patients who consented to the study, the shoulder capsular tissue and a venous blood sample was collected. Findings from analysis of the capsular tissue samples were then compared with data available in the Oxford Tissue Biobank of findings in healthy and diseased rotator cuff tendon tissues.

Inflammation signatures differed between tissues from frozen shoulder and tendon tears. Compared to tendon tear tissues, frozen shoulder capsular tissue showed reduced expression of nuclear factor-kB (NFkB) response genes including TNFA, IL6 and IL8; and increased expression of IL10, CD14, CD163 and C1QA mRNA. Fibroblast activation markers podoplanin (PDPN), CD106 (VCAM-1), CD248 and fibroblast activation protein (FAP) were highly expressed in adhesive capsulitis and torn tendons compared with healthy tendon. Fibroblast activation marker CD90 was however significantly reduced in adhesive capsulitis compared with healthy and diseased tendon tissues. Proresolving receptors mediating resolution of inflammation including ALX/FPR2, CMKLR1 and GPR32 were highly expressed in frozen shoulder capsular tissue.[38]

This study in similarly aged patients has provided some insight as to why the inflammation ultimately resolves in frozen shoulder but persists in tendon tears. This study suggests the phenotypes of fibroblast subsets populating diseased shoulder tissues differ between conditions with self-limiting and 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 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 sites randomised at least one patient. Appendix 3 presents the number of patients screened and randomised at each site, as well as the number 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 in Figure 2. Of 914 screened patients, 503 were randomised into the UK FROST trial. Reasons for exclusion were not meeting eligibility criteria (n=95), non-consent (n=295) or other reasons (n=21). The most frequent reason for exclusion was if frozen shoulder symptoms were secondary to trauma that required hospital care. Where patients provided information as to why they were not willing to join the trial, the most frequent reasons related to already having had physiotherapy and wanting 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 no evidence of differential dropout in any of the treatment arms. The primary analysis at 12 months follow-up included all participants with OSS outcome data at one or more follow-ups, therefore 94% of participants could be included in the analysis.

Table 1: Reasons for Exclusion

	Number excluded	Percent of total excluded (n=411)
Trial Exclusion Criteria (n=116, more than one reason possible)		
a) Bilateral concurrent frozen shoulder	20	4.9
b) Secondary to trauma (i.e. trauma to the shoulder that required hospital care (e.g. fracture, dislocation, rotator cuff tear)	23	5.6
c) Secondary to other causes (e.g. recent breast surgery or radiotherapy)	16	3.9
d) Any of the trial treatments are contraindicated (e.g. un fit for anaesthesia or corticosteroid)	16	3.9
e) Patient not resident in a catchment area of a participating site	20	4.9
f) Patient lacks mental capacity and unable to understand the trial or instructions for treatment	5	1.2
g) Other reason	21	5.1
Patient non-consent (n=295 – Grouped free text information from Screening form)		
Patient wants surgery	79	19.2
Patient does not want surgery	40	9.7
Patient wants physiotherapy	22	5.4
Patient does not want physiotherapy	48	11.7
Patient wants steroid injection	2	0.5
Patient wants clinician to decide	3	0.7
Patient wants no further treatment	2	0.5
Could not travel to trial site	1	0.2
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
Patient non-consent (n=295 – Selection of possible reasons from list of Preference form if agreed to complete, more than one reason possible)		
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 didn't want to take part	10	2.4
Other	29	7.1
Did not agree to complete Preference form	109	26.5

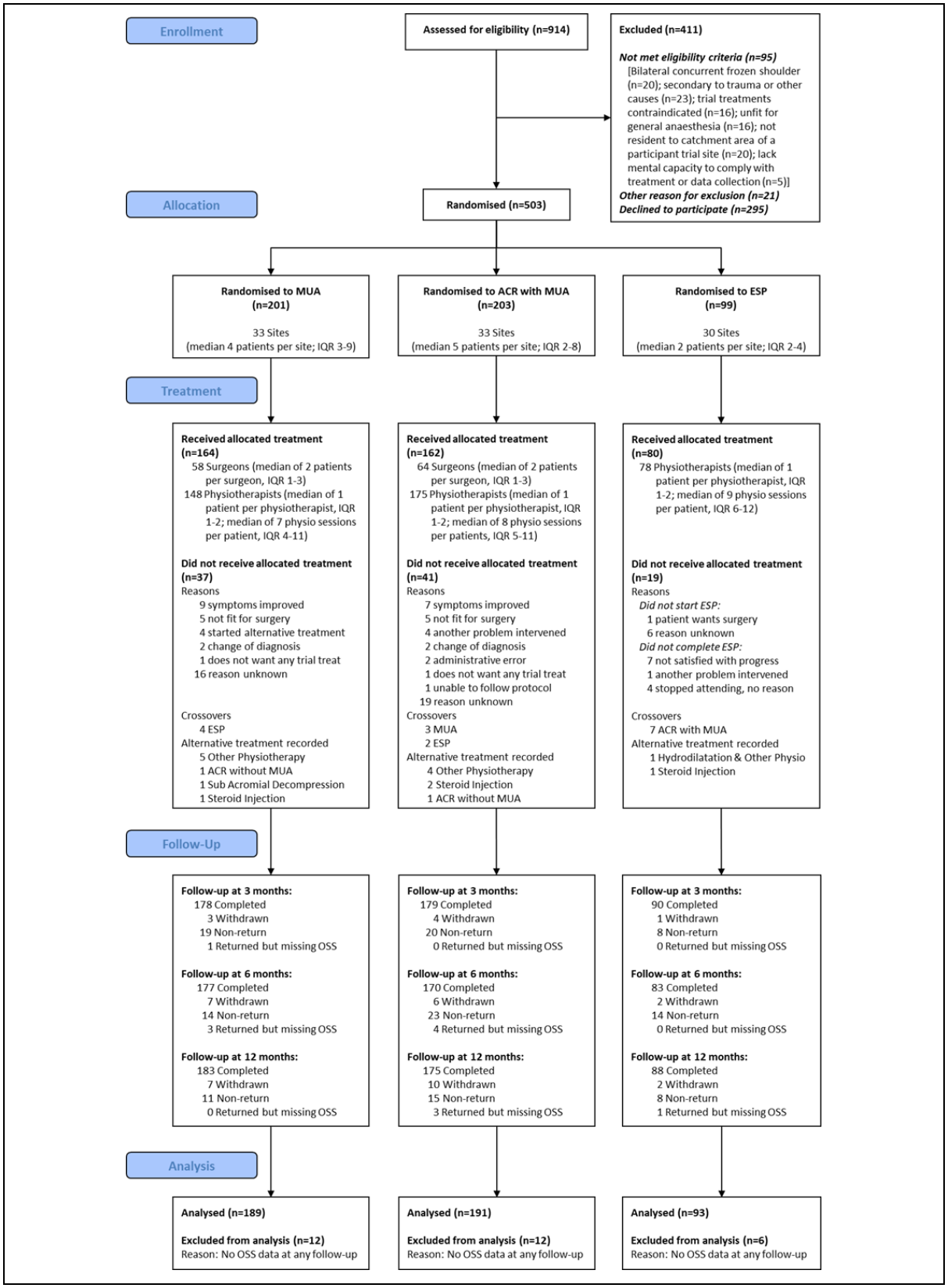


Figure 2: CONSORT Flow Diagram

Baseline characteristics

Eligible patients who did and did not consent to participate in the trial were comparable in their baseline characteristics (Table 2). Demographic and clinical characteristics of participants at baseline are presented in Table 3, comparing the profile of the total number of patients randomised (n=503) to that of participants who were 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 comparable, with the exception of a greater number of participants currently in paid work in the MUA arm and some group imbalance in having had a similar shoulder problem on the opposite side to the reference shoulder.

Table 2: Baseline Characteristics of different populations

Characteristic	Eligible but non-consenting N=295	Eligible and randomised N=503
Gender, n (%)		
Female	200 (68%)	319 (63%)
Age (years)		
N	293	503
Mean (SD)	53.7 (8.0)	54.3 (7.7)
Median (min, max)	53 (32, 82)	54 (30, 77)
Diabetic, n (%)		
No	219 (74%)	353 (70%)
Type I	23 (8%)	29 (6%)
Type II	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)		
N	288	495
Mean (SD)	10.5 (7.0)	10.9 (9.2)
Median [IQR]	9 [6, 12]	8 [6, 12]
min, max	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%)

Table 3: Baseline Characteristics of randomised participants

	As Randomised N=503			As Analysed N=473		
Characteristic	**MUA**	**ACR**	**ESP**	**MUA**	**ACR**	**ESP**
Gender, n (%)						
Female	129 (64%)	126 (62%)	64 (65%)	121 (64%)	117 (61%)	62 (67%)
Age (years)						
N	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 (min, max)	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 I	12 (6%)	12 (6%)	5 (5%)	12 (6%)	11 (6%)	5 (5%)
Type II	48 (24%)	48 (24%)	25 (25%)	46 (24%)	45 (24%)	22 (24%)
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]
min, max	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%)
X-rays, 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	33 (16%)	28 (14%)	15 (15%)	31 (16%)	26 (14%)	14 (15%)
Left school at 16	75 (37%)	74 (37%)	37 (37%)	70 (37%)	71 (37%)	34 (37%)
Left education at 18	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%)
Type of employment, n (%)						
Unskilled manual	17 (8%)	15 (7%)	8 (8%)	16 (8%)	13 (7%)	7 (8%)

	As Randomised N=503			As Analysed N=473		
Characteristic	**MUA**	**ACR**	**ESP**	**MUA**	**ACR**	**ESP**
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 General Practitioner, 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 physio-therapy 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]
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]

	As Randomised N=503			As Analysed N=473		
Characteristic	**MUA**	**ACR**	**ESP**	**MUA**	**ACR**	**ESP**
Similar shoulder problem on the same side, n (%)						
Yes	19 (9%)	26 (13%)	12 (12%)	17 (9%)	24 (13%)	12 (13%)
No	178 (89%)	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
Min, Max	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
Min, Max	0, 100	14, 100	14, 98	0, 100	14, 100	14, 98
Pain NRS (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
Min, Max	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
Min, Max	0, 100	0, 100	50, 100	0, 100	0, 100	50, 100

Intervention delivery

The criteria for having completed each of the three trial interventions were agreed and documented in the statistical analysis plan (SAP). For MUA and ACR, this constituted the completed receipt of the respective 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 sooner.

From Table 4, 82% of patients completed MUA, 80% of patients completed ACR and 81% of patients completed ESP. Overall, sixteen participants (3%) crossed over to a different trial treatment, and 17 (3%) received an alternative non-trial treatment. Where participants did not

start or complete any trial or non-trial treatment, these were classed as 'no treatment recorded' (n=64, 13%).

The profile of treating surgeons and physiotherapists for 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 a month. Physiotherapists delivering ESP or PPP were most frequently Band 6, treating between two to three frozen shoulders per month.

Table 4: Completed Treatment

Treatment Received	Randomised Treatment		
	MUA N=201	ACR N=203	ESP N=99
Trial Treatments			
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%)	-	-
Hydrodilatation & Other Physio	-	-	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 recorded for patient

Waiting times to the start of each randomised intervention varied considerably. From Table 5, ESP patients received their first physiotherapy session or steroid injection within a median of 14 days, whereas patients waited for a median of 56.5 days for MUA and a median of 71.5 days for ACR. In the ESP group, a steroid injection was received by 70 patients within an average of 12.8 days since randomisation. Nearly half (46%, n=32) were administered on the day of randomisation.

Following completion of their randomised treatment, a number of patients received further treatment, 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.

Table 5: Time to Start / End of Treatment

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

Table 6: Further Treatment (any treatment following completion of trial treatment)

	Sub-group: Randomised & 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

	Sub-group: Randomised & Completed Treatment		
	MUA	**ACR**	**ESP**
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%)**

As part of the surgical treatments, optimal release was reported as achieved in 92% of MUA procedures and 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 7: Fidelity (Surgery and Injection)

	MUA			ACR			ESP	
		% of randomised to MUA	% of randomised and completed MUA		% of randomised to ACR	% of randomised and completed ACR	% of randomised to ESP	% of randomised and completed ESP
	N	N=201	N=164	N	N=203	N=162	N=99	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%	79 (80%)	69 (86%)

The number of delivered physiotherapy sessions is presented in Table 8. 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). Individual therapeutic elements delivered as part of ESP and PPP sessions are summarised in Appendix 5.

Table 8: Fidelity (Physiotherapy)

	MUA		ACR		ESP	
Number of physiotherapy sessions	Randomised & completed MUA	Randomised to MUA	Randomised & completed ACR	Randomised to ACR	Randomised & completed ESP	Randomised to ESP
N	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
Min, Max	0, 18	0, 18	0, 18	0, 18	2, 15	0, 15

Primary Outcome

Descriptives

The Oxford Shoulder Score (OSS) was the trial primary outcome and was collected using questionnaires at baseline and then at 3, 6 and 12 months post-randomisation. Where OSS data from either the 6-month post-randomisation or 6-month post-treatment questionnaires were available, and the two questionnaires had been sent to patients within 28 days, 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 split by pain and function sub-domains). By 12 months 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 the ability for the trial to detect clinically meaningful differences at the primary end point was reduced.

Table 9: Unadjusted OSS by Treatment Arm – Follow-up since Randomisation

	MUA	ACR	ESP	Total
Baseline				
N	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
Min, Max	2, 48	1, 37	2, 42	1, 48
N (%) max score (48)	1 (1%)	0 (0%)	0 (0%)	1 (<0.5%)
3 Months				
N	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
Min, Max	5, 48	2, 48	4, 48	2, 48
N (%) max score (48)	2 (1%)	1 (1%)	2 (2%)	5 (1%)
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
Min, Max	3, 48	7, 48	6, 48	3, 48

	MUA	ACR	ESP	Total
N (%) max 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
Min, Max	4, 48	2, 48	4, 48	2, 48
N (%) max score (48)	44 (24%)	45 (26%)	17 (19%)	106 (24%)

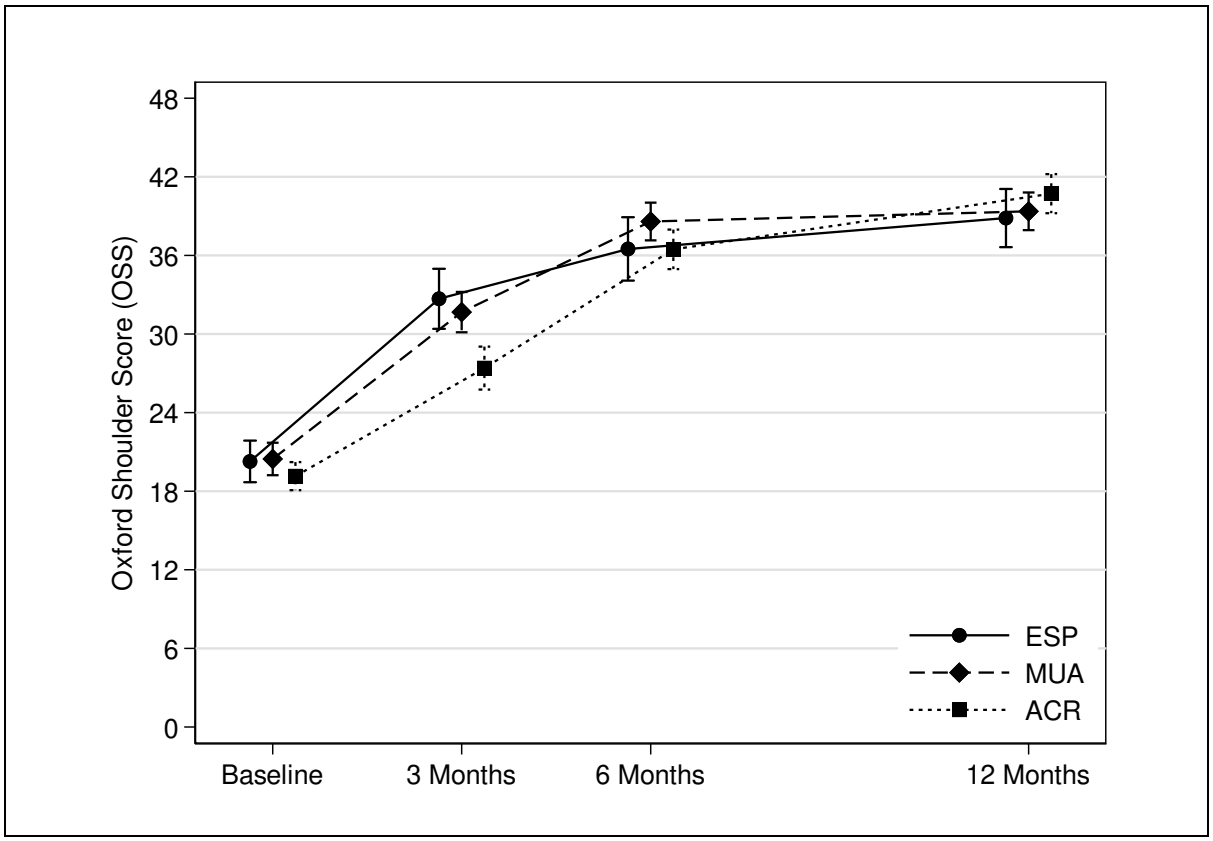


Figure 3: Unadjusted Mean OSS and 95% CIs by Treatment Arm

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, gender, 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 MUA (2.01 points, 95% CI 0.10 to 3.91) and ESP (3.06 points, 95% CI 0.71 to 5.41). Although statistically significant,

mean estimates were short of the sought minimal clinically important effect size of 4 to 5 OSS points (the trial was powered for differences of 4 points for comparing MUA with ACR and 5 points for comparisons with ESP).

For the short term follow-up at 3 months post-randomisation, ACR was shown to have lower (worse) outcomes compared with the other two interventions. Mean differences for all treatment comparisons are illustrated in Appendix 7. There was no evidence for 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])

	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)	p-value
	MUA	**ESP**	**Difference**	
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**	**Difference**	
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**	**Difference**	
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, gender, diabetes, OSS at baseline (fixed effects), and site (random effect)
[b] primary endpoint 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 presented together with OSS scores at baseline and 12 months post-randomisation follow-up in Figure 4. 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 for 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 ESP (95% CI for 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.

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). The 95% CI interval still included the minimal clinically import difference for this comparison of 5 OSS points.

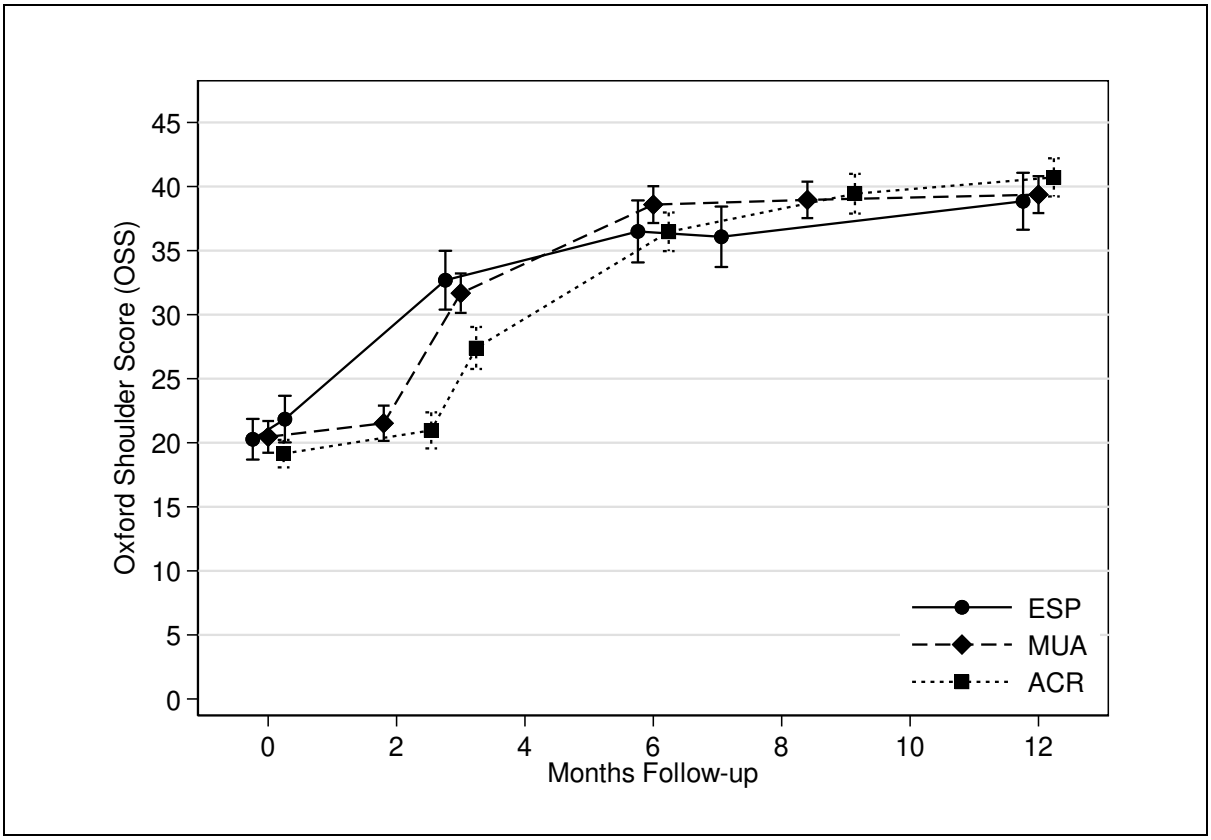


Figure 4: Unadjusted Mean OSS and 95% CIs by Treatment Arm (using scores at baseline, follow-up before treatment and 6 months after treatment; and 3, 6 and 12 months post-randomisation)

Table 11: Unadjusted OSS by Treatment Arm – Pre-/ 6 months post-treatment

	MUA	ACR	ESP	Total
Pre-treatment				
N	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
Min, Max	3, 46	1 , 42	6, 42	1, 46
Post-treatment (6 months)				
N	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
Min, Max	6, 48	2, 48	6, 48	2, 48

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])

	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)	p-value
	MUA	ESP	Difference	
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	Difference	
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

	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)	p-value
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**	**Difference**	
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, gender, diabetes, OSS at baseline (fixed effects), and site (random effect)

Analysis incorporating treatment compliance

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

Table 13: Comparison of baseline characteristics by treatment compliance

Characteristic	MUA		ACR		ESP	
	Completed treatment	Did not complete treatment	Completed treatment	Did not complete treatment	Completed treatment	Did not complete treatment
	n=164	N=37	n=162	N=41	n=80	N=19
Gender, 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 (min, max)	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 I	12 (7%)	-	8 (5%)	4 (10%)	4 (5%)	1 (5%)
Type II	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 (min, max)	8 (2, 60)	7.5 (2, 48)	9 (2, 96)	9 (0, 36)	9 (3, 36)	8 (2, 72)
Previous physio-therapy for affected shoulder, n (%)						
Yes	100 (61%)	25 (68%)	99 (61%)	25 (61%)	49 (61%)	10 (53%)
No	64 (39%)	12 (32%)	63 (39%)	14 (34%)	31 (39%)	8 (42%)
Missing	-	-	-	2 (5%)	-	1 (5%)
OSS (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 (min, max)	20 (2, 48)	20 (3, 36)	19 (1, 37)	19 (4, 35)	20 (2, 42)	18 (4, 34)

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.

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 other treatment arms (-1.84 OSS points, 95% CI -4.41 to 0.74, 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.

Missing data

Possible predictors of missing OSS data at 3, 6 or 12 months 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	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	Mean (SD)	20.1 (8.2)	18.5 (8.6)	0.19
6-month follow-up		**N=430**	**N=73**	
Age	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	Mean (SD)	20.2 (8.2)	18.2 (8.4)	0.06

		Not missing	Missing	p-value
Month-3 OSS	Mean (SD)	30.4 (10.9)	26.2 (11.9)	0.05[a]
12-month follow-up		**N=446**	**N=57**	
Age	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	Mean (SD)	20.1 (8.2)	18.3 (8.7)	0.11
Month-3 OSS	Mean (SD)	30.4 (10.8)	25.9 (13.4)	0.05
Month-6 OSS	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 (Appendix 9). Results were similar to those observed in the primary analysis.

Sub-group analyses

The possibility of differential treatment effects were explored for sub-groups based on diabetes status, receipt of previous physiotherapy and baseline treatment preference, and additionally length of frozen shoulder symptoms at baseline following advice from the trial oversight committee. Interaction terms between treatment allocation and sub-groups 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 were statistically significant, although the study was not powered to detect such interactions and the number of participants in some of the sub-groups in each treatment arm was very low.

Possible trends are illustrated in Figure 5, Figure 6, Figure 7, Figure 8 and descriptive tables in Appendix 10. Diabetic patients tended to have poorer outcomes compared with non-diabetic patients at all time-points, and especially at 3 months follow-up for patients in the ACR arm. Patients who had previous physiotherapy tended to have worse outcomes if randomised to ESP, especially at 3 and 6 months follow-up, whereas patients who indicated a prior

preference for physiotherapy tended to have better outcomes if randomised to ESP and worse outcomes if randomised to either surgical treatment. Participants who reported frozen shoulder symptoms for 9 months or more prior to entering the trial, tended to have worse outcomes at 3 months if randomised to ACR and better outcomes at 3 months if randomised to ESP.

Table 15: Sub-group analyses summary

Sub-groups	MUA n	ACR n	ESP n	Treatment Comparison	Contrast	95% CI	p-value of allocation interaction with sub-group
Diabetes							
n= 150 diabetic	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
n= 353 not diabetic	141	143	69	ACR vs MUA	0.43	-2.98 to 3.85	0.80
Previous Physiotherapy							
n= 308 had previous physiotherapy	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
n= 192 did not have previous physiotherapy	76	77	39	ACR vs MUA	1.21	-2.02 to 4.44	0.46
Patient Treatment Preference							
n= 131 allocated to preferred treatment	56	64	11	MUA vs ESP	2.10 [a]	-4.32 to 8.52	0.81
					1.28 [b]	-4.66 to 7.22	
n= 105 allocated to non-preferred treatment	40	27	38	ACR vs ESP	3.11 [a]	-3.50 to 9.73	0.65
					2.18 [b]	-3.73 to 8.09	
n=263 no treatment preference	103	111	49	ACR vs MUA	1.01 [a]	-3.84 to 5.87	0.87
					0.90 [b]	-2.70 to 4.50	
Length of symptoms at baseline							
n=249: < 9 months	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
n=246 >= 9 months	93	106	47	ACR vs MUA	0.41	-2.73 to 3.56	0.80

[a] Allocated to non-preferred vs preferred treatment

[b] No treatment preference vs allocated to preferred treatment

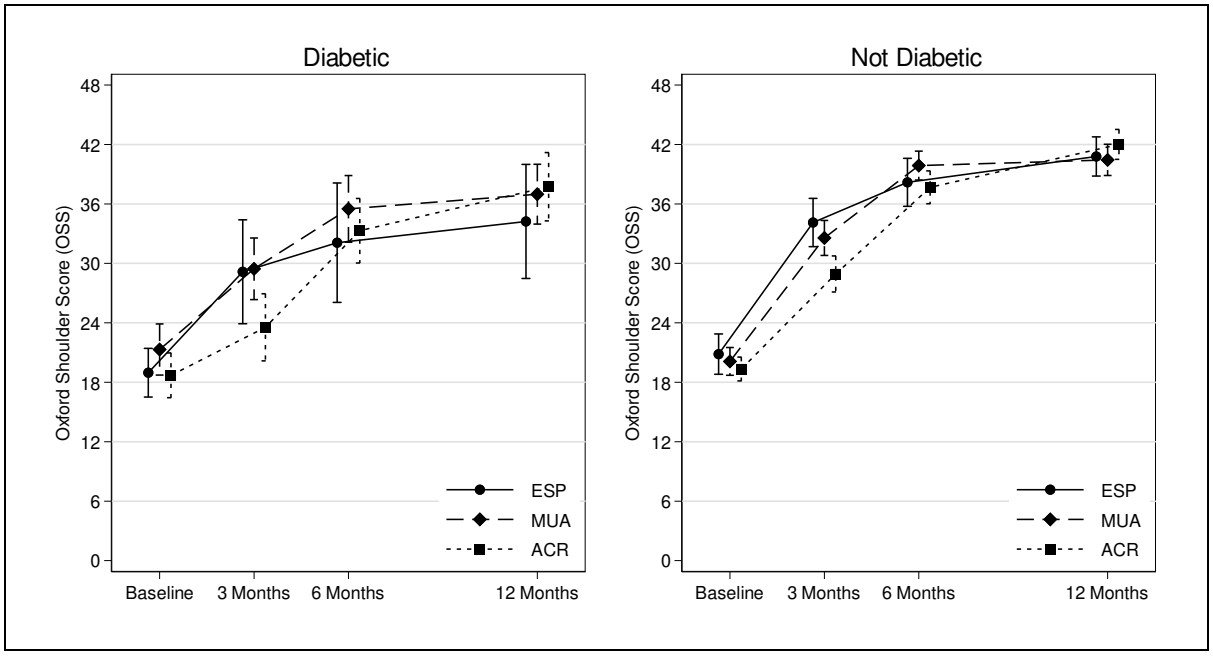


Figure 5: Unadjusted Mean OSS Function Items and 95% CIs by Treatment Arm and Diabetes

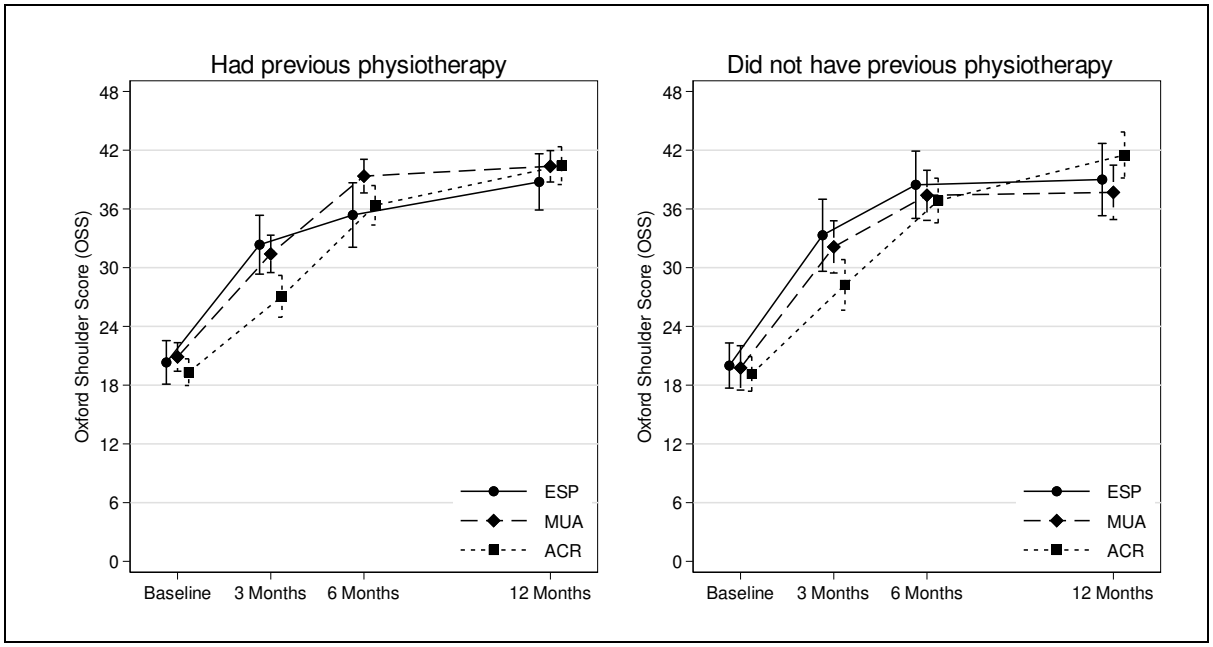


Figure 6: Unadjusted Mean OSS Function Items and 95% CIs by Treatment Arm and Previous Physiotherapy

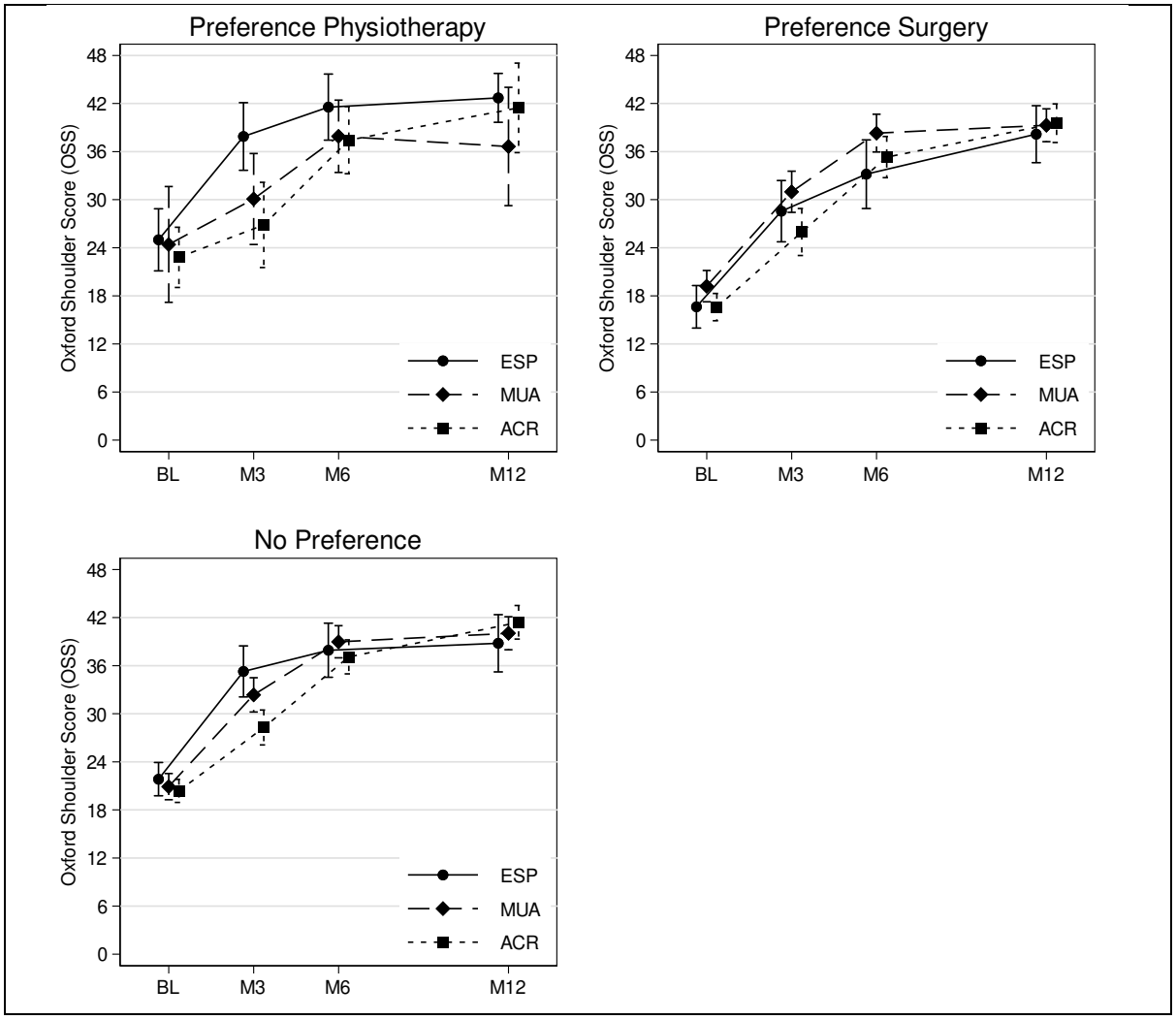


Figure 7: Unadjusted Mean OSS Function Items and 95% CIs by Treatment Arm and Baseline Preference

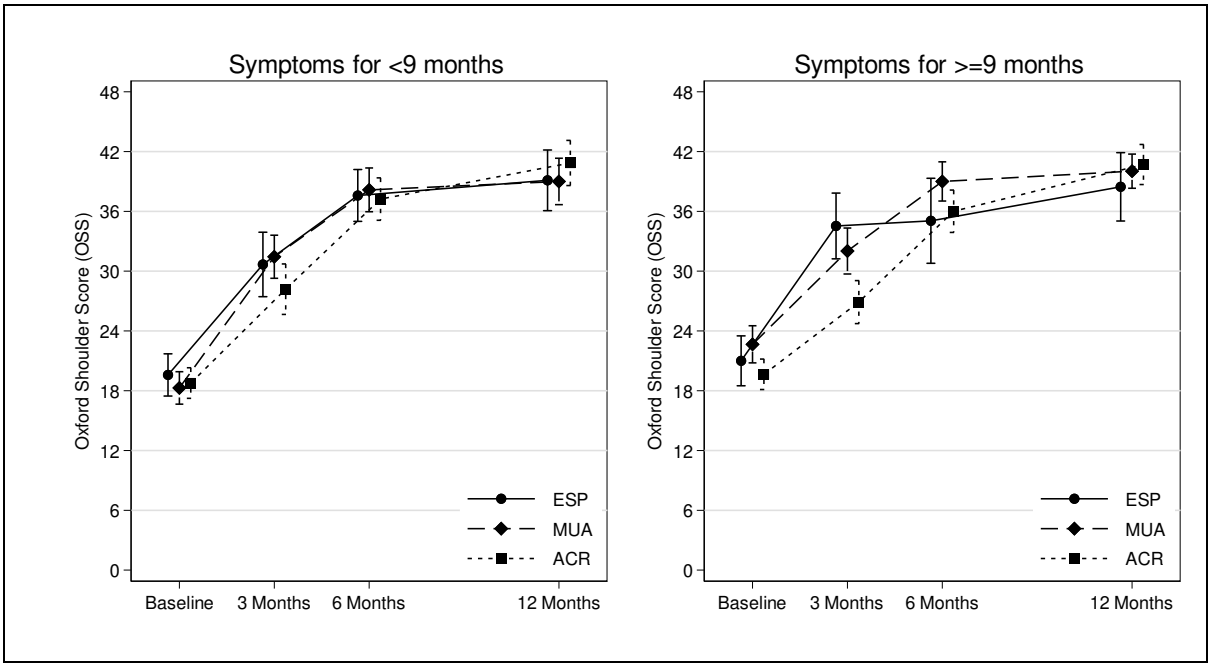


Figure 8: Unadjusted Mean OSS Function Items and 95% CIs by Treatment Arm and Length of Symptoms

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 (note that many patients only recently had or were still waiting for their surgery at this point) but better outcomes at 12 months post-randomisation compared with MUA or ESP (Table 16, Table 17). Unadjusted means are presented and illustrated in the tables and figures below (Appendix 11, Figure 9, Figure 10).

QuickDASH

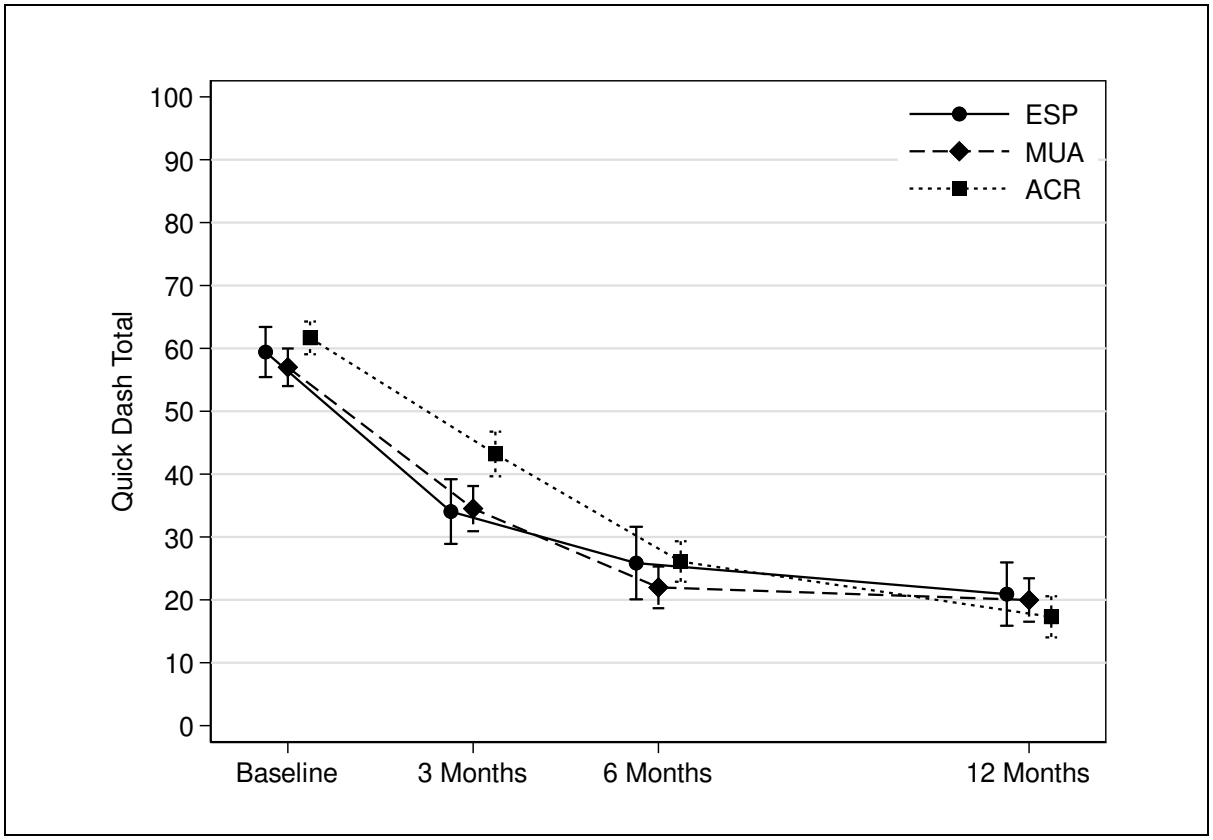


Figure 9: Unadjusted Mean *Quick* DASH scores and 95% CIs by Treatment Arm

Table 16: Estimated Mean Quick DASH Differences by Treatment Arm[a]

	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)	p-value
	MUA	**ESP**	**Difference**	
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**	**Difference**	
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**	**Difference**	
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, gender, diabetes, Quick DASH at baseline (fixed effects), and site (random effect)

Shoulder pain

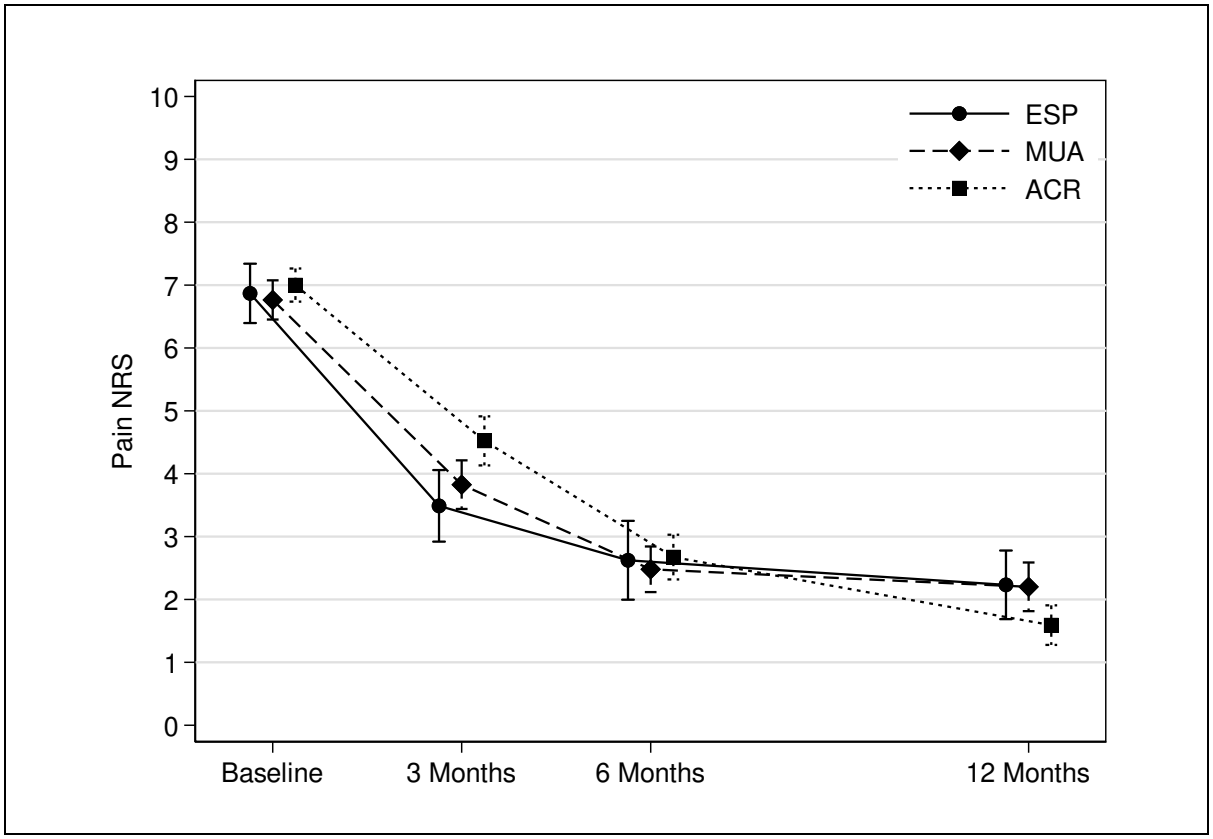


Figure 10: Unadjusted Mean Shoulder Pain NRS and 95% CIs by Treatment Arm

Table 17: Estimated Shoulder Pain NRS Differences by Treatment Arm[a]

	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)	p-value
	MUA	**ESP**	**Difference**	
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**	**Difference**	
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**	**Difference**	
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, gender, diabetes, Pain NRS at baseline (fixed effects), and site (random effect)

Extent of recovery

There was no evidence for statistically significant differences between treatment arms for 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 0-100 VAS). Descriptives are presented in Appendix 11 (Table 53), illustrated in Figure 11, and results of the analysis given in Table 18.

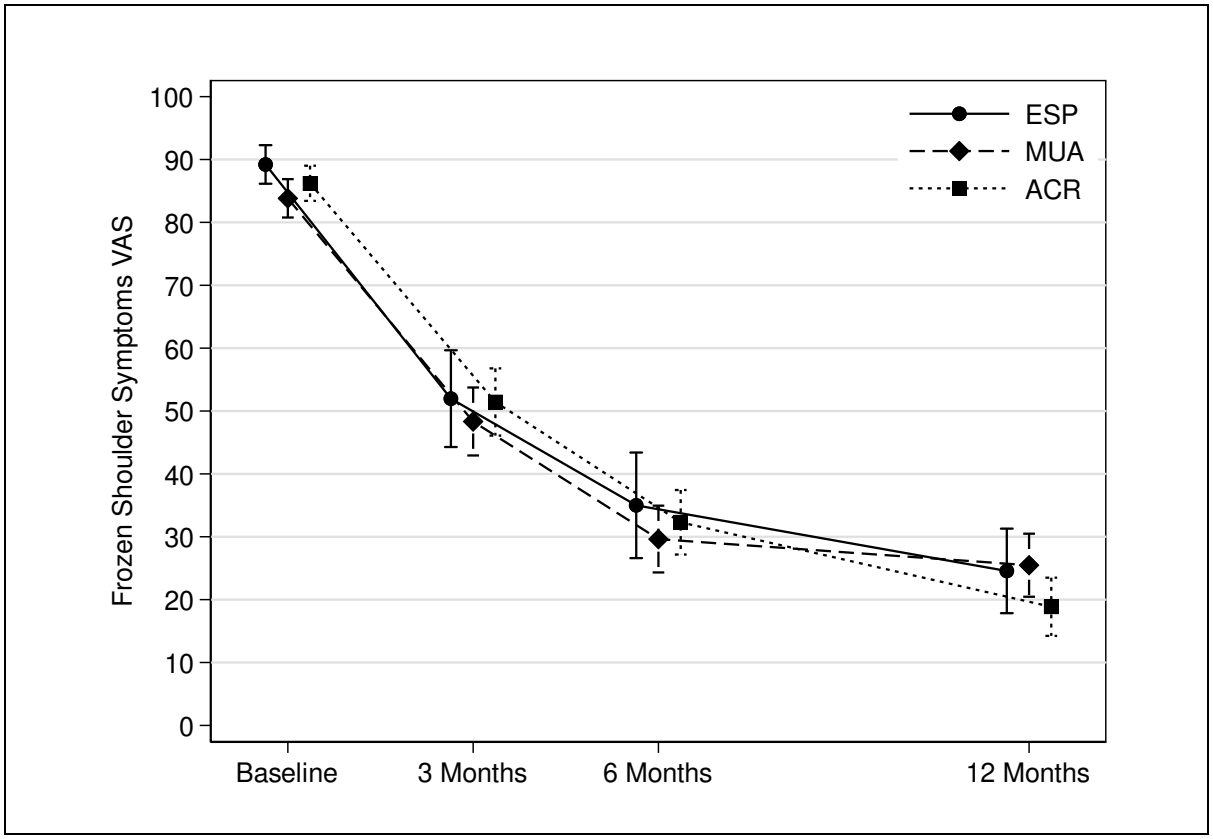


Figure 11: Unadjusted Mean Extent of Recovery[a] VAS and 95% CIs by Treatment Arm

[a] Scale: 0-100, 100 equivalent to maximum belief that symptoms require further treatment

Table 18: Estimated Mean Frozen Extent of Recovery[a] VAS Differences by Treatment Arm[b]

	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)	p-value
	MUA	**ESP**	**Difference**	
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**	**Difference**	
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**	**Difference**	
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: 0-100, 100 is equivalent to the maximum belief that symptoms require further treatment

[b] linear mixed covariance pattern model adjusted for age, gender, 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 are presented in Appendix 11 (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 compared with the other groups.

Complications / Adverse events

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

In total, there were only ten serious adverse events (SAEs), reported for nine patients (summarised in Appendix 12, itemised list available in Table 19). All SAEs occurred for patients randomised to the surgical arms (n=8 for ACR and n=2 for MUA). However, one SAE in the ACR group was for a participant who had none 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 for formal analysis.

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

Source	Treatment		Description	Type	Relatedness	Potentially long lasting consequences
	Allocated	Received				
SAE CRF	ACR	ACR	Elevated blood sugars	Prolonged Hospitalisation	Probably related	No
SAE CRF	ACR	ACR	Decreased oxygen saturation	Prolonged Hospitalisation	Not related	No
SAE CRF	ACR	ACR	Hypoglycaemic seizure whilst under anaesthetic	Prolonged Hospitalisation	Unlikely 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 related	No

[a] Event identified following review rather than AE CRF, not all information available

There were 33 non-serious adverse events, 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, itemised list available in Table 20). There were sufficient patient numbers experiencing one or more adverse event to allow for a valid statistical comparison between the two surgical arms, which confirmed no evidence for statistical differences in the proportion of non-serious adverse events (p=0.186).

Table 20: Non-serious adverse events (itemised)

Source	Treatment		Description	Relatedness	Expectedness	Severity
	Allocated	Received				
AE CRF	ESP	ESP	Persistent pain	Not related	Expected	U/K
AE CRF	ESP	ESP	Long head 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

Source	Treatment		Description	Relatedness	Expectedness	Severity
	Allocated	Received				
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	U/K
AE CRF	MUA	MUA	Ipsilateral face swelling, face flushed and neck and face hot	Possibly related	Unexpected	Moderate
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	U/K
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 + needles to 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 identified following review rather than AE CRF, not all information available

Other Analyses

Treatment Preferences

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

OSS Change Scores

Details of participants' comparative assessment of their symptoms at baseline and 12 months with reference to their change OSS score 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

OSS and Quick DASH scores for participants who did (n=441) and 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 months follow-up, however the proportion of participants with available data was much lower in this group as well (e.g. 66% of valid OSS scores at 12 months versus 92% of participants who did receive treatment).

Systematic review – integrating the new evidence

A systematic review was undertaken to assess the effectiveness of MUA, arthroscopic capsular release, 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 largest and most robust evidence.

Due 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 in size (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 as 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 within the NHS.

Overview

A prospective economic evaluation was conducted alongside the UK FROST trial with the aim of estimating the cost-effectiveness of the three most commonly used interventions for the management of the frozen shoulder in secondary care. The three interventions compared in the study were: Early Structured Physiotherapy (ESP) with an intra-articular steroid injection compared with manipulation under anaesthesia (MUA) with a steroid injection or arthroscopic (keyhole) capsular release followed by manipulation (ACR). Both surgical interventions were followed with a programme of post-procedural physiotherapy (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 UK £ sterling at a 2017-2018 price base. Health benefits were expressed in terms of quality-adjusted life-years (QALYs), based on patients health related quality of life (HRQoL) assessed using the EuroQol-5 Dimensions, five-level version (EQ-5D-5L).[68, 69] Differences in mean costs and mean QALYs at one year were used to derive an estimate of the cost-effectiveness of surgery and non-surgical treatment.

The base-case analysis was conducted on an intention-to-treat (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, 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 Clinical Excellence (NICE) guidelines were applied to all methods used for this economic analysis.[70]

Owing the impact of missing data, the base case analysis was conducted as an imputed analysis;[71] the choice of method to handle missing data (multiple imputation) was grounded in the assumed missing data mechanism (missing at random) which in turn was supported by the UK FROST dataset. The impact of alternative assumptions about the missing data mechanism was carefully assessed in sensitivity analyses.

Methods

Data sources

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

Data relating to surgical care were collected via Surgical Forms that were specifically designed for the trial. Similarly, Physiotherapy Logs were completed by physiotherapists providing patient care. These were used to record the essential components of physiotherapy at each session for each participant, and were used to estimate the cost of ESP, and the costs of post procedural physiotherapy following MUA or ACR.

Resource use from primary health care consultations was collected using participant questionnaires only. All resource use recorded from participants was split into "shoulder-related" and "non-shoulder related" and was collected at three, six and 12 months. The base case analysis was based on shoulder related resource use. Hospital stays and hospital outpatient appointments were recorded by two sources (patient questionnaires and hospital forms). Our health economic analysis plan (HEAP) indicated that where 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: (i) Complication form; and (ii) Change in Status form. These forms recorded any hospitalisation from discharge after initial treatment up to 12 months.

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

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

Data on health benefits, expressed in terms of HRQoL, were elicited from participants by means of the EuroQol-5 Dimensions, five level version (EQ-5D-5L) measure at baseline, three, six and 12 months.

Measurement of resource use and costs

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

Surgery (MUA and ACR)

An accurate record of procedures at hospital level (e.g. centres in the trial) was put in place in order to record 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 cost of MUA and ACR. The post-procedural physiotherapy (PPP) form was used to cost post-surgical physiotherapy care for MUA and ACR participants.

Non-surgery (ESP intervention)

The Structured Physiotherapy (SP) form was used for patients receiving ESP. This form recorded information on the physiotherapy sessions (i.e. duration of the session and staff band of the physiotherapist delivering the session). Information on physiotherapy visits was also available from participant questionnaires at three, six and 12 months. As stated in our HEAP, PPP and SP forms were used as primary source for the base-case analysis. As part of the ESP treatment, patients were offered an intra-articular steroid injection at the earliest opportunity. In order to cost injections, we collected information on the type of steroid used (e.g. methylprednisolone acetate, triamcinolone acetonide, triamcinolone hexacetonide or other steroid); the dose of steroid used (i.e. 20mg, 30mg, 40 mg or other dose); whether local anaesthetic or image guidance was used; and the job title of person injecting (i.e. specialist registrar, associate specialist, consultant, physiotherapist Band 6, or physiotherapist Band 7).

Healthcare consultations and hospital care

Data on healthcare consultations and hospital care were used to assess whether participants allocated to the MUA and ACR experienced different levels of resource use from the ESP group. The costs of healthcare consultations consists of all the costs of visits to both primary and secondary healthcare 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 three, six 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. received before/during receiving randomised treatment), further treatments (i.e. treatments received after completing randomised treatment), other treatments (i.e. any non-trial treatments the patient had if they did not start/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 health care resource use by the associated unit costs.

Costs relating to both surgical interventions was 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) data[72] for hospital-based health staff. These unit cost estimates were inclusive of components for staff salaries, salary on-costs, overheads and capital overheads. Drug tariff per milligram for medications (i.e. anaesthesia, antibiotics and steroid injections) were obtained from the British National Formulary.[73] In order to cost length of stay, we used NHS Reference costs[74] taking the weighted average inpatient bed-day for all major and intermediate shoulder procedures (footnote to HRG codes).

Costs relating to the ESP intervention were based on staff time. Physios cost per hour was estimated using PSSRU 2018 based on data for hospital-based health staff (Bands 5 to 8). The full course of ESP was up to 12 sessions, exceptionally, physiotherapist may decide that more than 12 sessions were needed. The costs relating to the ESP intervention comprised of the costs of the physiotherapy sessions and the cost of the steroid injection, which was obtained from the British National Formulary.

The use of other hospital-based care was valued by applying unit costs extracted from national tariffs.[72, 74] Similarly, costs for the primary care and community-based services were estimated by applying unit costs from national tariffs[72, 74] 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 National Statistics to occupational information derived from self-reported work status information.[75]

Costs were estimated in UK pound sterling and based on the financial year 2017-2018. Table 21 details the unit costs used in the analysis. The total cost comprises three main components: (1) the cost of the initial intervention (i.e. ESP, MUA and ACR); (2) hospital stays and hospital outpatient appointments after initial intervention; (3) and 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. Sensitivity analysis were used to explore the impact of a broader perspective (i.e. private care costs and productivity costs) in cost-effectiveness results.

Mean (SE) costs by cost category and mean (SE) total cost were estimated per each treatment group using regression analysis to control for patients' covariates (i.e. age, gender, treatment group, baseline Oxford Shoulder Score (OSS) and diabetes (yes/no).

Health Related Quality of Life (HRQoL) and quality-adjusted life-years (QALYs)

This economic evaluation took the form of a cost utility analysis, where health outcomes were assessed in terms of quality-adjusted life years (QALYs). HRQoL was expressed in terms of utilities, which were obtained from trial patients using the EQ-5D-5L at baseline, three months, six months and one year. The EQ-5D-5L consists of two principal measurement components. The first is a descriptive system, which defines HRQoL in terms of five dimensions: 'mobility', 'self-care', 'usual activities', 'pain/discomfort' and 'anxiety/depression'. Responses in 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.

According to the responses to the EQ-5D classification system, a health status can be defined and a single index utility 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.[76] However, at the time of this analysis, the most recent guidance issued by NICE regarding the EQ-5D-5L (https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-

technology-appraisal-guidance/eq5d5l_nice_position_statement.pdf), recommends the use of the mapping function (i.e. crosswalk) developed by van Hout et al. 2012[69] 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 scores –0.281 to 1.0, with 0 representing death and 1.0 representing full health; values below 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.[77] Therefore, utility values were adjusted using a univariate generalised linear model, including group as a fixed factor and baseline EQ-5D 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 into QALYs for each patient using the area under the curve method (AUC), following the trapezium rule which assumes linear interpolation between follow-up points.[78]

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

Missing data

Missing data occur frequently in randomised controlled trials (RCTs), irrespective of how well designed the data collection is. This is a major concern for within-trial cost-effectiveness analyses (CEA), as costs and QALYs, the main outcomes in CEAs, are cumulative measures collected over the trial follow-up. Therefore, missing data at one follow-up time point (e.g. one dimension response missing to 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 amount 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 to different results and decisions on the value for money of the assessed interventions. Complete-case (CC) 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.[79] Therefore, it was decided prior to the analysis that CC would be presented only for comparison purposes. Besides, the analysis of the missing pattern of the UK FROST dataset would also support this decision, as results suggest data is not missing completely at random (MCAR; assumption driving the CC mechanism).

An alternative method to address missing data in CEAs alongside RCTs is multiple imputation (MI)[79] which has been recommended as the appropriate method to reflect the uncertainty in the results of the economic evaluation attributable to missing data.[70] The main assumption that drives the MI mechanism is that the data are "missing at random" (MAR). That is, the missing values in the dataset may depend on the value of other observed variables in the dataset, but conditional on those values the data are missing at random. A major concern is that the chance of data being missing maybe directly linked 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 is 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 analysis conducted within RCT,[80, 81] we conducted descriptive analyses of the missing data to explore whether MAR assumption is plausible given the actual missing data mechanism of UK FROST dataset. We assessed the amount of missing data by trial arm at each follow-up period, explored missing data patterns using graphical tools, and investigate the association between missingness and baseline variables /observed outcomes by means of logistic regressions.

Based on the results of the descriptive analyses, we could conclude that MAR is a plausible assumption fitting UK FROST dataset. Therefore, MI was selected to handle missing data for the base-case analysis. MI using chained equations[82] and predicted mean matching were carried out on the EQ-5D-5L at three, six and 12 months as well as the total cost estimates. Predicted

mean matching is a semi-parametric imputation approach, and ensures that observed data were 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, that no imputed values will lie outside the bounds of the original data distribution. The MI 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, gender, 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. MI by chained equations was performed for a total of 60 imputations. The estimates obtained from each imputed dataset are combined to generate mean estimates of costs and QALYs, variances and CI using Rubin's rules.[83]

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. Additionally, 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 cost-effectiveness rations (ICERs) when appropriate.[84] SUR allows to jointly estimates outcomes, so brings efficiency gains over unrelated Ordinary Least Squares (OLS).[85] In the bivariate model incremental costs and QALYs are simultaneously estimated from two separate OLS regressions, assuming correlation between the error terms in both regressions.[86] The SUR model used the same set of covariates as the mixed-effect regression model used for the clinical

effectiveness analysis (age, gender, baseline EQ-5D score, baseline OSS score and diabetes (yes/no)).

The cost-effectiveness results were expressed in terms of incremental costs-effectiveness ratios (ICER). This was 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 WTP threshold for an additional QALY ranges from £20,000 to £30,000.[70] 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.[87, 88] 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.

In order to compute the probability of each intervention being cost-effective at a given cost-effective threshold, the seemingly unrelated regression was conducted within a bootstrapping approach on five imputed datasets 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 thresholds applied by NICE of £20,000 to £30,000/QALY, and a threshold of £13,000/QALY as suggested by recent research.

Sensitivity Analyses and uncertainty

The uncertainty around the cost effectiveness results was explored by means of 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: to: (1) including non-shoulder costs (ITT); and (2) adopting a broader perspective that includes productivity costs and private care costs. A further number of sensitivity analyses were conducted to explore the impact of missing data in cost-effectiveness estimates: (3) restricting the analyses to complete cases following (ITT); (4) imputing QALY data at aggregated level rather than at the index-score level; (5) mix model approach; and (6) missing not at random scenario.

Results

Study population

The baseline study population for the economic analysis was 503 patients. A total of 99 patients were allocated to the ESP intervention; 201 to MUA and 203 to ACR. A total of 19 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 crossover from their initial randomisation. This involved patients crossing from ESP to ACR (n=7), MUA to ESP (n=4), ACR to ESP (n=2) and ACR to MUA (n=3).

Health care resource use and costs

Costs of delivering Surgery (MUA and ACR)

Detailed resource use and costs of both surgical interventions are given in Table 21. Costs relating to surgical procedures are based on time in theatre, delivery of anaesthesia and injections and length of hospital stay. In order 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 that 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 that crossed over to ACR with surgical form available (n=5).

The mean cost of MUA was £424.81 (SD=115.55). For 97% of the cases MUA was delivered as a day case, only 3% of the cases required hospitalization (only one night). The average duration of the MUA intervention was 25.11 minutes (SD=14.20).

The mean cost of ACR was £ 2,170.46 (SD=431.11). For 90% of the cases ACR was delivered as a day case; 10% of the cases required hospitalization being on average 2.8 nights (median=1; min=1; max=31) in hospital. The average duration of an ACR was 76.61 min (SD=24.22).

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

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

Intervention - Features	**MUA (N=168)**	**ACR (N=170)**
Theatre time (minutes)[b] - 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 [c] – N (%) Length of stay (nights) – Mean (SD)	5 (3%) 1.2 (0.45)	17 (10%) 2.8 (7.31)
Patients had PPP within their allocated group Number of physio sessions (PPP) - Mean (SD) - Max	160 (80%) 6.42 (4.95) - 18	159 (78%) 6.65 (4.81) - 18
Patients had injection during PPI PPP – N (%)	162 (97%)	46 (27%)
Costs (in £)	**MUA Mean (SD)**	**ACR 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 (DC/LOS)	218.52 (8.03)	1,590.81 (398.29)
Cost Surgical procedure[c]	**424.81 (115.55)**	**2,170.46 (431.11)**
Cost Surgical procedure _Sensitivity analysis	*428.57 (242.45)*	1,308.26 (413.43)
Cost of PPP[d]	**213.61 (157.13)**	**209.44 (152.95)**
Cost surgical Procedure PLUS physio	**638.42 (204.75)**	**2,379.90 (457.88)**

Non-surgery (ESP intervention)

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

Table 22 Costs related to ESP intervention

	ESP* (N=92)
Cost Steroid injection – mean (SD) *	42.96 (31.82)
Cost physiotherapy – mean (SD) *	217.11 (146.85)
MEAN (SD) cost – ESP intervention	**260.07 (155.07)**

Hospital costs related to complications and additional/further/other treatments the patient had from discharge after initial treatment up to 12 months are shown in Table 23.

Table 23 Cost related to complications and additional treatments by trial allocation

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

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 dataset. 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, median estimates are smaller than means for all resource use, which suggests that the distributions were skewed to the right.

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

Resource use were 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 PSS 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, £2,271.09 (902.50) in the ACR arm.

Total costs estimates shown here are unadjusted means, and relates to complete cases, therefore there is limited value in interpreting differences between treatments. Mean differences for each surgical treatment versus ESP and corresponding 95% CIs, adjusted for patient covariates, and taking into consideration the correlation between costs and QALYs are shown in the cost-effectiveness section.

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	Mean (SD)	Median	Missing (%)	N	Mean (SD)	Median	Missing (%)
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)
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 Therap.	**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)

Table 25 Unit costs used for the analysis (£, 2017-18 prices)

Item	Unit cost (£)	Source
PRIMARY AND COMMUNITY CARE		
GP visit at GP practice	37.40	Curtis and Burns [5]
GP visit at home	93.60	Curtis and Burns [5]
GP by phone	15.20	Curtis and Burns [5]
Nurse visit at GP practice	10.85	Curtis and Burns [5]
District/ community nurse	38.00	Curtis and Burns [5]
Occupational therapist visit	47.00	Curtis and Burns [5]
Physiotherapist visit [b]	57.25	DH [7]
HOSPITAL CARE		
Inpatient stay (shoulder) [c]	258.00 – 449.00	DH [7]
Inpatient stay (non-shoulder)	384.22	DH [7]
Day case visit (shoulder) [c]	420.00 - 2,512.00	DH [7]
Outpatient visits (shoulder)	125.01	DH [7]
Outpatient visit (non-shoulder)	123.93	DH [7]
Hospital physiotherapy visit	54.91	DH [7]
Other health service visit	74.11	DH [7]
Consultant surgical	108.00	Curtis and Burns [5]
Associate specialist	105.00	Curtis and Burns [5]
Speciality Registrar	43.00	Curtis and Burns [5]
Foundation doctor FY1	32.00	Curtis and Burns [5]
Foundation doctor FY2	28.00	Curtis and Burns [5]
Physiotherapist B5	35.00	Curtis and Burns [5]
Physiotherapist B6	46.00	Curtis and Burns [5]
Physiotherapist B7	55.00	Curtis and Burns [5]
Physiotherapist >8 [d]	72.00	Curtis and Burns [5]
Nurse B5	37.00	Curtis and Burns [5]
Nurse B6	45.00	Curtis and Burns [5]
Nurse B7	54.00	Curtis and Burns [5]
MEDICATIONS		
Depomedrone 40mg	3.44	BNF [6]
Depomedrone 80mg	6.88	BNF [6]
Triamcinolone 40mg	17.88	BNF [6]
Triamcinolone 80mg	35.76	BNF [6]
Bupivacaine 0.5% (10ml)	0.915	BNF [6]
General anaesthesia *	30.99	BNF [6]
Antibiotics	6.11	BNF [6]
PRIVATE CARE		
Private Non-NHS physio	50.00	https://www.capitalphysio.com
Private osteopath	42.50	https://www.nhs.uk/conditions/osteopathy
Private chiropractitioner	55.00	https://www.nhs.uk/conditions/chiropractic
Community care service	49.00	Averaged of three above
Private hospital - night	337.00	DH [7]

[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 HRG 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 band >8. Hence unit cost for physio>8 is estimated as averaged 8a(£66) and 8b(£78)

Table 26 Costs for cases with complete data by trial allocation and cost category (£, 2018-19 prices)

Costs	MUA Mean (SE) (£)	ACR Mean (SE) (£)	ESP Mean (SE) (£)
MUA^	349.46 (191.91)	5.55 (55.78)	0
ACR^	0	1,762.32 (934.61)	113.42 (495.67)
ESP^	6.90 (59.08)	1.25 (13.02)	260.07 (155.07)
Physio (Hospital)	175.88 (163.90)	174.63 (161.73)	6.98 (36.55)
Physio (Community)	43.80 (146.04)	66 (201.67)	61.87 (211.11)
Further treatments[a]	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 surgery	60.33 (113.86)	64.77 (120.87)	33.78 (70.61)
GP on the phone	8.15 (31.22)	6.69 (16.93)	1.49 (7.18)
Nurse at 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)
Total Costs (NHS) Shoulder (a)	**834.02 (752.66)**	**2,271.09 (902.5)**	**599.06 (359.23)**
Total Costs (NHS) non-shoulder^ (b)	182.12 (228.98)	195.76 (304.22)	241.82 (366.23)
Productivity costs (c)	1,995.29 (2,999.85)	3,735.61 (5,031.35)	1,308.70 (3,165.177)
Private care costs (d)	31.23 (117.63)	21.40 (111.22)	40.00 (144.51)
Total Broader costs (a + b + c +d)	**3,200.98 (3,824.39)**	**5,377.18 (4,240.28)**	**1,475.05 (2,367.87)**

^The cost of the intervention includes the costs of injections. [a]As mentioned in the text, non-shoulder costs were restricted to primary care data.

Health Related Quality of Life and QALYs

Regarding the extent of completeness of EQ-5D questionnaires at each time point during trial follow-up, it was shown a balanced decrease in complete EQ-5D questionnaires throughout the trial follow-up. Response for the complete follow up (i.e. baseline, three, six and 12 months) was available for 369 (73%) participants: 156 (78%) in MUA, 149 (73%) in ACR, and 64 (65%) in ESP. The extent of incomplete EQ-5D due to missing data strengthened the justification for using the imputed datasets as the base-case (see Appendix 16).

The proportion of participants who reported the EQ-5D-5L levels (1 to 5) by dimension, group and time point are provided (Appendix 16). Comparing self-care levels between baseline and 12 months we found that there were more people who report no problems across all the treatments, but there was a slightly higher percentage in the ACR group. When looking at usual activities, all groups had a similar increase over 12 months in the number of participants who class themselves as being in Level 1. A lower percentage of participants were in level 1 for pain/discomfort after 12 months in the ESP group compared to the MUA and ACR groups. 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 scores (utilities) for the different follow-up assessments is illustrated in Figure 12. Patients allocated to MUA started from a higher utility value compared to patients allocated to ACR and ESP.

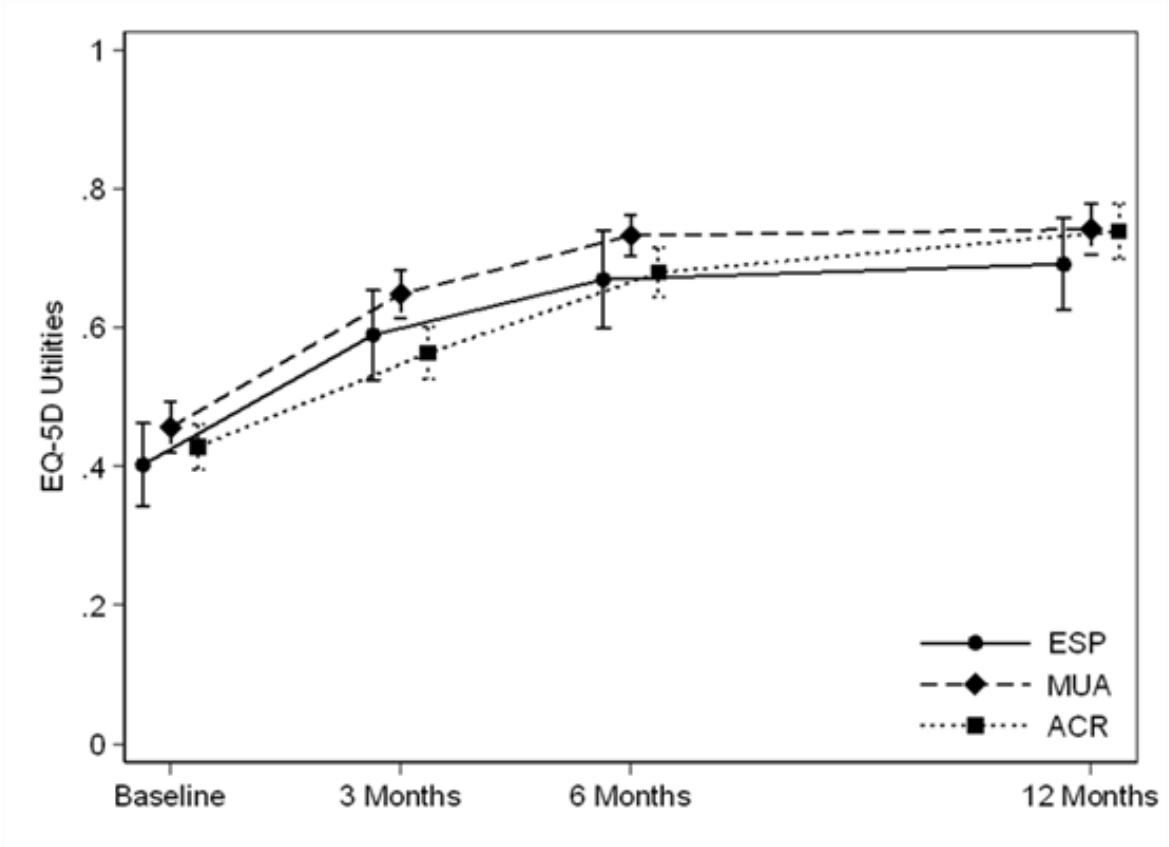


Figure 12 EQ-5D-5L scores distribution at the different time points over the 12 months

Table 27 summarises the mean EQ-5D scores reported at each follow-up point for all the available cases. Adjusted analysis shows that patients allocated to ACR and MUA had similar utility values at 12 months follow up [ACR (mean 0.739) vs MUA (mean 0.734)]. Similarly, patients allocated to the surgical groups had better utility values compared to ESP (mean 0.693). QALYs estimates at one year follow up (adjusted for baseline utility) shows that patients allocated to MUA accrued more QALYs that the other two groups: MUA (0.6765) > > ESP (0.6492) >ACR (0.6475).

Table 27 HRQoL: EQ5D-5L summary scores (available cases) at each time point adjusted for baseline utility. QALYs estimates (complete cases) adjusted for baseline utility

Follow-up point	MUA, mean (SE)/(CI)	ACR, mean (SE)/(CI)	ESP, mean (SE)/(CI)
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)	**0.6765 (0.651 to 0.702)**	**0.6475 (0.621 to 0.674)**	**0.6492 (0.609 to 0.690)**

As for total costs, HRQoL and QALYs estimates shown in this section are of limited value, as these estimates corresponds exclusively to patients with complete EQ-5D data (i.e. baseline, three, six and 24 months). 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 the cost-effectiveness section.

Missing data

The UK FROST study collected data on EQ-5D-5L at three, six and 12 months. Healthcare resource use was elicited from patients by postal questionnaire at three, six and 12 months; and from health professionals by hospital forms at 52 weeks after randomisation. A description of economics variables in UK FROST can be found in Appendix 17.

Overall, the proportion of participants with complete economic data remained similar between treatment groups (Appendix 17): 46.46% in the ESP group, 58.21% in the MUA group and 57.14% in the ACR group. In all groups, 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 (lost to follow-up 6 months but remained subsequently). Hence, Inverse Probability Weighting (IPW) would be inappropriate under such pattern. Similarly, CCA would be, as a minimum, inefficient because it would discard observed data from individuals with some missing outcomes.

Figures representing the pattern of missing data are shown in Appendix 17. As discussed above, missing data is shown to be non-monotonic, since individuals with missing data are 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 at baseline is associated with missing cost and QALY data (Appendix 17). Baseline age was also found to be a significant predictor of missing data on HRQoL. This suggests that the data is unlikely to be MCAR. The other baseline covariates (gender and diabetes) were associated with missingness but not

statistically significant at 5%. However, diabetes was significant predictor of costs and QALYs at 6 months and 1 year, which would support both CD-MCAR and MAR assumptions.

We also explored whether 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 month). Most regressions produced statistically insignificant results ($p>0.05$) results 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 CD-MCAR.

Therefore, data are assumed to be MAR and MI by chained equations (MICE) was selected to handle missing data for this economic analysis. In the analysis, missingness is assumed to depend on baseline covariates (gender, diabetes, age, EQ5D at baseline and OSS score at baseline) and observed costs and QALYs but independent of unobservable costs at QALYs at one 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. CCA, which is not valid under MAR, is presented for comparison only.

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

Cost effectiveness analysis

Base-case analysis

A bivariate regression, in the form of a seemingly related regression, conducted in the imputed dataset, 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 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 ESP (mean difference 0.0103; 95% CI -0.0304 to 0.0510). Overall, ACR had worse (non-significant) QALYs compared to MUA at the 12 month follow up (mean difference -0.0293; 95% CI -0.0616 to 0.0030).

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

	Adjusted difference in means with SUREG[a]	95% confidence limits		
Difference in cots (£)				
MUA vs ESP	276.507	(65.67 to 487.35)		
ACR vs ESP	1,733.78	(1,529.48 to 1,938.06)		
ACR vs MUA	1,457.26	(1,282.73 to 1,631.79)		
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 cost-effective at £13,000/QALY**	**Probability cost-effective at £20,000/QALY**	**Probability cost-effective at £30,000/QALY**
MUA	6,984	0.7942	0.8632	0.8978
ACR	> 100,000	0.0000	0.0002	0.002
ESP	-	0.2058	0.1366	0.1002

[a] Compared with ESP, as it is the alternative with lower costs and health outcomes

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 to physiotherapy, MUA intervention mean cost 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 6,984 per additional QALY. ACR is significantly more costly than ESP [on average £1,733.78 more expensive per participant (95% CI (1,529.48 to 1,938.06)]; and despite the QALY gained 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 ACR being a cost-effectiveness use of NHS resources when compared with ESP (i.e. ICER above recommended NICE threshold). Similarly, ACR surgical interventions is dominated by MUA, with higher mean costs and lower QALYs.

The corresponding CEACs, showing the probability of each treatment being cost-effective across a range of thresholds is shown in Figure 13. The probability that MUA surgery was cost-effective is 0.88 for 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 was cost-effective dos not exceed 0.002.

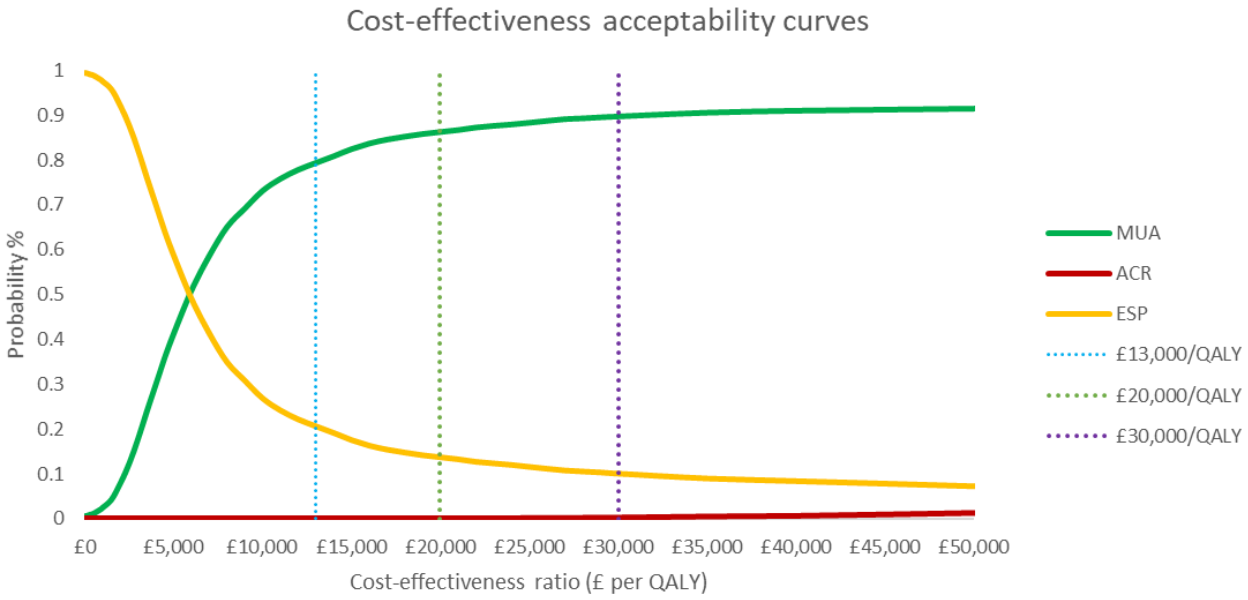


Figure 13 Base case cost-effectiveness acceptability curves

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. In contrast, cost-effectiveness results were sensitive to a wider perspective scenario; suggesting the ICER from a wider perspective was higher than the thresholds that NICE normally consider for reimbursement decisions. Regarding ACR, this continued being dominated by MUA in both scenarios.

Table 29 Sensitivity analysis (Scenario 1 and Scenario 2): Summary for incremental analysis (ITT), cost-effectiveness results and uncertainty under different costs scenarios

		MI of costs (shoulder – NHS perspective) and QALYs analysis with SUREG Base-Case analysis	MI of costs (shoulder and non-shoulder – NHS perspective) and QALYS analysis with SUREG SA (Scenario 1)	MI of costs (broader perspective^) and QALYS analysis with SUREG SA (Scenario 2)
MUA vs ESP				
Difference in cots (£)	**Mean**	276.5	162.76	1,031.86
	SE	107.4462	112.83	595.33
	95% CI	65.67 to 487.35	-58.39 to 383.91	-136.92 to 2,200.65
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	6,984	10,485	27,522
Probability that MUA is cost-effective [a]		0.88	0.77	0.36
ACR vs ESP				
Difference in cots (£)	Mean	1,733.78	1,555.48	4,109.96
	SE	104.147	112.42	647.75
	95% CI	1,529.48 to 1,938.06	1,335.14 to 1,775.82	2,836.20 to 5,383.73
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 ^		0.030	0.008	0.000
ACR vs MUA				
Difference in cots (£)	Mean	1,457.26	1,392.72	3,078.10
	SE	88.90998	91.41	548.27
	95% CI	1,282.73 to 1,631.79	1,213.56 to 1,571.87	1,999.07 to 4,157.13
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 ACR is cost-effective surgery		0.00%	0.00%	0.00%

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

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

Table 30 Sensitivity analysis (Scenario 3. Scenario 4 and Scenario 5): Summary for incremental analysis (ITT), cost-effectiveness results and uncertainty under different missing data assumptions

		Complete case analysis with SUREG	MI of costs and utilities followed by SUREG	Mixed model with adjustment for covariates
Difference in cots (£)	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	5,395.58	8,562
Probability that MUA is cost-effective		0.48	0.89	0.76

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

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

Conclusion

This economic analysis 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 for the ITT approach showed that MUA was the intervention most likely to be cost-effective. The resulting ICER for MUA was £6,984 per additional QALY when compared to ESP; over common threshold values of a QALY, the probability that MUA was cost-effective was high (> 85% for NHS perspective). The finding indicates that ACR is dominated by MUA (higher mean costs and lower QALYs) and ACR showed very low probability of being cost effective (<5% for NHS perspective).

The positive cost-effectiveness profile of MUA is robust to plausible departures of 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 £20,000 per QALY threshold, probabilities ranging across scenarios from 48% (CCA) to 99%. The only exemption was when we used the societal perspective to estimate the costs. ACR appeared as 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 to methodological guidelines; and 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 the frozen shoulder in terms of QALYs gained calculated using the EQ-5D-5L. Compared to ESP, MUA intervention 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 for was 6,984 per additional QALY. The probability that MUA is cost-effective is above 85%, whilst the probability that ESP is cost-effective did not exceed 20%. ACR was significantly more costly than ESP [on average £1,733.78 more expensive per participant (95% CI (1,529.48 to 1,938.06)]; and despite the QALY gained 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 ACR being a cost-effectiveness use of NHS resources when compared with ESP (i.e. ICER above recommended NICE threshold). Despite the ACR group having fewer additional interventions, ACR was dominated by MUA, with higher mean costs and lower QALYs. Therefore, given the worse outcomes observed in the ACR group compared to the other two groups, along with its greater costs makes this treatment difficult to justify. The results of the base-case 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 in interpreting these results. The first relates to the issue of missing data. Although the use of hospital forms improved the amount of incomplete data, the presence of missing data was unavoidable. We followed a comprehensive analysis to explore whether MAR assumption is plausible given the actual missing data mechanism of UK FROST dataset. Our analysis showed that MAR is a plausible assumption fitting UK FROST dataset; therefore, MI was selected to handle missing data for the base-case analysis. Furthermore, the results were robust to alternative assumptions on the pattern of missing data, showing that the positive cost-effectiveness profile of MUA is robust to plausible departures from 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

view of potential functioning. The clinical results showed that nearly fifty per cent of the patients were only three points off being on perfect health in the OSS, which in turns should influence positively on 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 the QALYs observed in the ACR group were lower compared with the other two interventions at three months, which is consistent with the results of the OSS. Conversely, the ACR group had higher QALYs and OSS scores at 12 months. Moreover, while MUA had marginally higher estimates of OSS scores compared with ESP and ACR for the average treatment effect over 12 months, this also applied to MUA group accruing more QALYs over the duration of the study. The extra additional cost for an 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, where the most significant risk is fracture and it is very unlikely this happened beyond follow-up. To note that sensitivity analysis to explore whether results were sensitive to under reporting complications did not change the positive cost-effectiveness results in favor 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 generalizable to other healthcare 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 commonly on the trial feasibility, participation, data collection process, and effects of trial interventions.[89, 90] A comprehensive understanding of the subjective experiences to complement the quantitative evaluation of clinical and process outcomes in a trial would better inform patient-centred care and evidence-based practice.[90] The use of qualitative research methods in stand-alone studies or as part of clinical trials is gaining momentum in a wide range of musculoskeletal conditions.[91-93] However, there are very few published qualitative studies in people with frozen shoulder.[19, 94] 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 FROST[28] to provide trial participant and professional-centred insights to guide clinical decision making.

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

Methods

The UK FROST trial participants, surgeons, and physiotherapists who 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 endpoint of the trial.[8] This allowed for post-surgical recovery and time 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 via post or email. Non-respondents were reminded at the second and fourth week of the invitation. On receiving signed consent for participation and audio-

recording of interviews, 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 semi-structured with open questions about their experience of living with frozen shoulder and the treatments in the trial. An interview schedule (*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 (*Supplementary Material 22*) covered routine clinical management of frozen shoulder; the experience of treating participants in the trial; personal treatment preferences; and barriers and enablers for positive treatment outcomes in frozen shoulder.

We planned to interview to the point of theoretical saturation[95] until no further useful categories emerged. We proposed to interview up to 45 trial participants and 15 healthcare professionals.

The data were analysed in two ways:

1. The interviews were analysed using constant comparative methods.[96, 97] This involves comparing similarities and differences and developing themes with a shared essence. The data was 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[98] to organise the analysis.

2. The ICF[99, 100] (International Classification of Functioning, Disability, and Health) 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 activities and participation component and '5' its chapter on 'Self-care'. A second-level category 'd510' depicts a self-care problem 'washing oneself'. The third and fourth levels of 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 are 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 component; 34 in activities and participation and 8 in environmental factors) identified in a previous study on chronic shoulder conditions including the frozen shoulder.[101]

Results

Sixty interviews (Trial participants: 44; Surgeons: 8 and Physiotherapists: 8) were completed between August 2016 and January 2018. All interviews with the trial participants were conducted via telephone. The majority (75%) of the interviews with surgeons and physiotherapists were telephone-based and a few were face-to-face (Surgeons: 2; Physiotherapists: 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 Early Structured Physiotherapy (ESP), Manipulation under Anaesthesia (MUA) or Arthroscopic Capsular Release (ACR). All participants who were interviewed had their allocated treatments, except for one participant allocated to MUA and another allocated to ACR who received ESP.

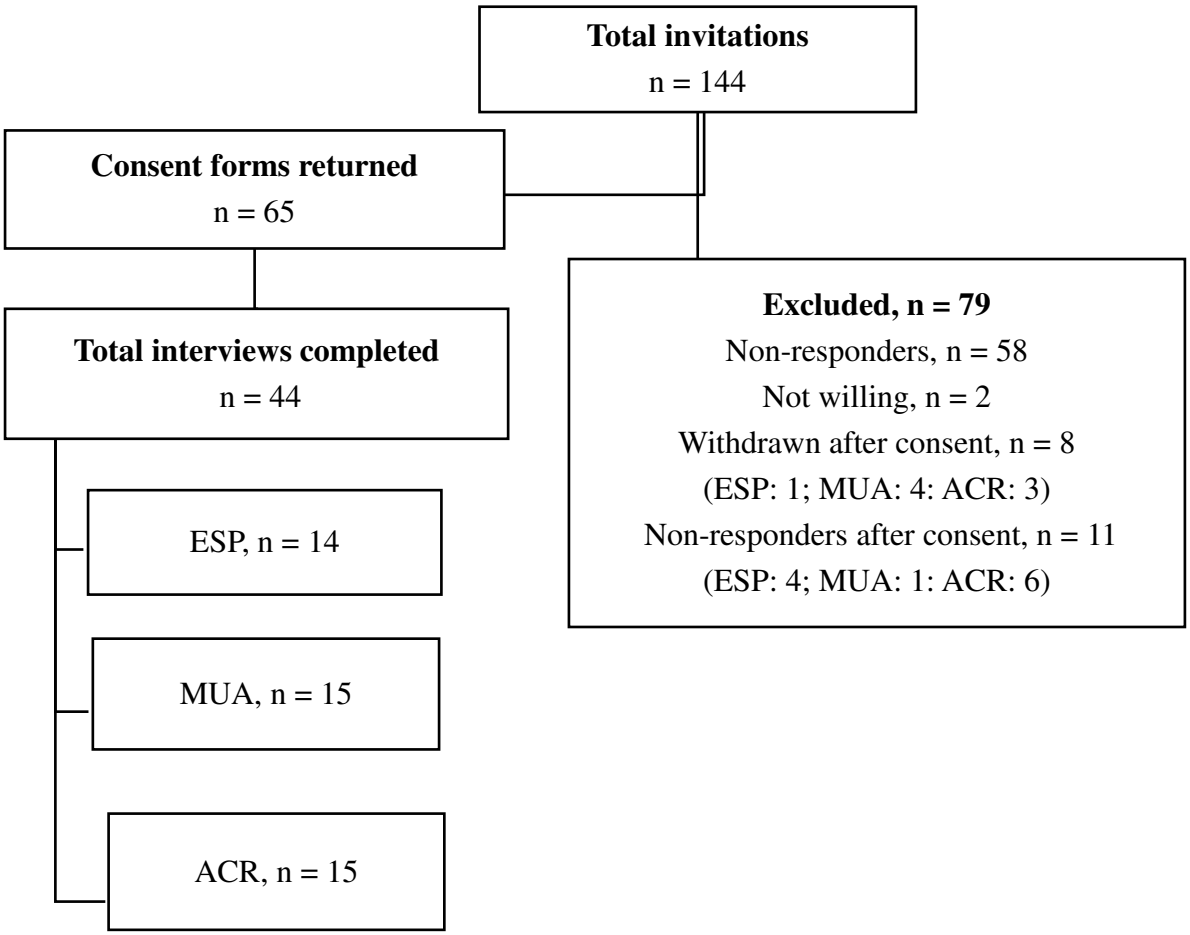


Figure 14 Flow diagram of the trial participant interviews

Table 31 Characteristics of the trial participants interviewed

UK FROST arms	Male	Female	Age Median (IQR[a]) in years	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

[a]IQR: Interquartile range

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

Living with frozen shoulder

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

61 years, Female (ESP): *"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".*

53 years, Female (ACR): "*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".*

Trial participants had mixed experiences with 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 in participating in leisure, work, and social activities. Table 32, Table 33 and Table 34 present the narratives of the trial participants mapped to 19 ICF categories (5 in body functions and structures; 12 in activities and participation; 2 in environmental factors) from a previous study.[15]

Table 32 Participant-reported problems mapped to ICF categories: Body structures and functions

Trial participants' quotes	ICF categories
"It was like a stabbing pain; it was very severe" – 73 years, Male (ACR).	b280-Sensation of pain
"I couldn't move my shoulder at all. So it was stuck to my side" -59 years, Female (ESP)	b710-Mobility of joint functions
"Definitely lost my strength in my arm"- 66 years, Male (MUA).	b730-Muscle power functions
"I didn't sleep at all"- 48 years, Female (ESP).	b134-Sleep functions
"I was generally getting very depressed. I would have happily, towards the end, I would happily have them amputate the arm"- 50 years, Male (MUA).	b152-Emotional functions

Table 33 Participant-reported problems mapped to ICF categories: Activities and participation

Trial participants' quotes	ICF categories
"Carrying was probably a bit of a problem because I couldn't move that arm so well" 64 years, Female (ESP)	d430-Lifting and carrying objects
"Reaching things from the tall shelves in the kitchen, reaching stuff out of the top of the wardrobe…And its things you take for granted really"-57 years, Male (MUA)	d445-Hand and arm use
"I couldn't drive my car even. I couldn't change gear"-55 years, Male (ESP)	d475-Driving
"I couldn't lift my hand up, my arm up, you know, to wash my	d510-Washing oneself

hair or anything in the bath. And eventually I couldn't get into the bath properly" -76 years, Male (ESP).	
"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" -53 years, Male (ACR).	d530-Toileting & d520-Caring for body parts
"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" -58 years, Female (ACR).	d540-Dressing
"I couldn't open anything, I couldn't use a tin opener, so I couldn't cook neither, I couldn't do anything"-55 years, Female (ACR)	d630-Preparing meals
"Doing housework was nigh impossible"-54 years, Female (MUA)	d640-Doing housework
"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 seven months". -55 years, Male (MUA)	d850 Employment
"Well I decorate cakes. I do novelty cakes, all kinds of decorating. That literally stopped". -64 years, Female (ACR) *"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".* -66 years, Male (MUA)	d920 Recreation and leisure
"I couldn't sleep on my left side anymore and if I turned over in the night and tried to sleep on that side it instantly woke me up." -57 years, Male (MUA)	d410 Changing basic body position

Table 34 Participant-reported problems mapped to ICF categories: Environmental factors

Trial participants' quotes	ICF categories
"…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". -72years, Male (ESP) *"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"*- 53years, Female (MUA)	e310 & e320- Immediate family and friends -Facilitators

Participants described they put off making a GP appointment until the symptoms worsened and some described delays in NHS care. In retrospect, participants thought a quicker NHS care pathway in terms of diagnosis and further specialist referrals was important. Participants also emphasised seeking early medical help and referrals by their GP.

64 years, Female (ESP): "*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".*

50 years, Male (MUA): "*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".*

Some participants reported that they had a range of treatments such as pain killers, physiotherapy, acupuncture, steroid injections before the trial, whilst others had no treatment at all. Participants felt that the pain killers and injections used before the trial did not help them. Similarly, physiotherapy treatments had not helped to increase range of movement, partly because of the difficulty in exercising due to pain.

59 years, Male (MUA): "*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".

Participants were concerned that they were stuck with the disability from frozen shoulder and were *eager* on getting it sorted. This was the main motivation for them to enter the UK FROST trial.

59 years, Female (ACR): "*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 pain killers 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".*

Improvements in outcomes and participant satisfaction following the trial

Trial participants considered pain relief and return of shoulder movements and function as important treatment outcomes.

62 years, Female (ESP): "*Going back to normal… When you had nil pain and full flexibility and movement within your and shoulder. No sleepless nights and that".*

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

76 years, Male (ESP): "*When I went to the surgeon I was injected into my shoulder and the pain down my arm that more or less went straightaway. 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".*

56 years, Female (MUA): "*So at the beginning I said the pain was ten and now after all my physios, I'd say it was, I'd say it was about two now".*

55 years, Female (ACR): "*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".

Trial participants in all treatment arms reported increased shoulder movements.

68 years, Male (ESP): "*Virtually full movement. My shoulder is fine as far as movement is concerned".*

58 years, Male (MUA): "*Basically had all my full movement back".*

58 years, Female (ACR): "*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".*

Trial participants described that the physiotherapy sessions (ESP and post-procedural) helped to improve their shoulder movements.

64 years, Female (ESP): *"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 twelve weeks was up, I virtually had full movements with no pain or anything, it was brilliant!"*

45 years, Female (MUA): "*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".*

53 years, Female (ACR): "*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".*

ACR arm participants felt their recovery in terms of pain and movements was quicker than they expected. Some experienced improvements as early as one to two weeks of physiotherapy after surgery.

44 years, Male (ACR): "*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"*.

Trial participants in all treatment arms said that their ability to do routine activities improved.

54 years, Male (ESP): "*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"*.

52 years, Female (MUA): "*My little everyday things have come back; I have come back, yes"*.
59 years, Female (ACR): "*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"*.

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

45 years, Female (ESP): "*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"*.

55 years, Male (MUA): "*There's still a wee bit of pain there, but it's nothing. You see, I'm not concerned about it"*.

73 years, Male (ACR): "*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"*.

Trial participants were satisfied with the UK FROST treatments they received.

64 years, Female (ESP): "*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"*.

56 years, Female (MUA): "*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".*

61 years, Male (ACR): "*Very satisfied. I have no complaints at all".*

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

50 years, Female (ESP): "*...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 okay, and they didn't".*

59 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".*

Trial participants' adherence to home exercises

Participants found that the exercises were difficult to begin with but eased off on subsequent sessions.

57 years, Male (MUA): "…*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".*

However, they were aware of the benefits of exercises and persevered to do their home exercises regularly.

48 years, Female (ESP): "*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".

Most participants said they did not continue their home exercises after the trial because they felt they 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.

54 years, Male (ESP): "*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".*

Trial participants' treatment preferences before and after the trial

Participants had mixed treatment preferences before the trial. Some MUA and ACR participants felt physiotherapy would be ineffective because it didn't work for them previously or felt that physiotherapy would be difficult to do with their painful shoulder. MUA was perceived as less invasive and ACR as an effective treatment. A few preferred physiotherapy to avoid the risks of surgery. Some didn't have any particular preference at all. Three participants with diabetes felt it would take longer to recover after ACR because of diabetes.

53 years, Male (ACR): "*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".*

55years, Female (MUA): "*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 evasive surgery which means cutting open, an actual surgical procedure which I wasn't that keen on to be honest with you".*

76 years, Male (ESP): "*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".*

55 years, Male (MUA): "*Just I didn't think, as I said, I didn't know if any better was than the other, I just went along with what was there".*

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

61 years, Female (ESP): *"Oh no I would have the same treatment. As I say I was only there for ten weeks and I mean, on my eleventh week I was still hell of a lot better".*

53 years, Female (MUA): "*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, Male (ACR): "*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".*

A few ESP and MUA arm participants wanted to choose ACR for a permanent and quicker solution for their shoulder problem.

68 years, Male (ESP): "*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".*

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 while some expected to get their shoulder problem treated quickly.

62 years, Female (ESP): "*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".*

53 years, Female (ACR): "*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*".

Trial participants found the trial questionnaires simple and relevant to complete. A few felt that the questions were lengthy and repetitive. Some had difficulties answering questions on health status compared to the previous month.

64 years, Female (ESP): "*It was quite easy, they were simple questions, and I just sort of flew through the questionnaire with no problem".*

57 years, Male (MUA): "*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".*

Trial participants said that the physiotherapists who delivered the physiotherapy sessions (ESP and post-procedural) were supportive and helpful.

41 years, Female (ESP): *"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".*

50 years, Male (MUA): "*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".*

They also liked to see the same physiotherapist for developing good connections and rapport throughout the programme.

53 years, Male (MUA): "*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".

MUA and ACR participants said that the surgical procedures were both explained and went well.

58 years, Female (MUA): "*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".*

44 years, Male (ACR): *"It was just a case of just sitting around, reading books, talking with other people, some people had been there 2-3 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."*

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

64 years, Female (ESP): "*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…"*

53 years, Male (ACR): *"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".*

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.

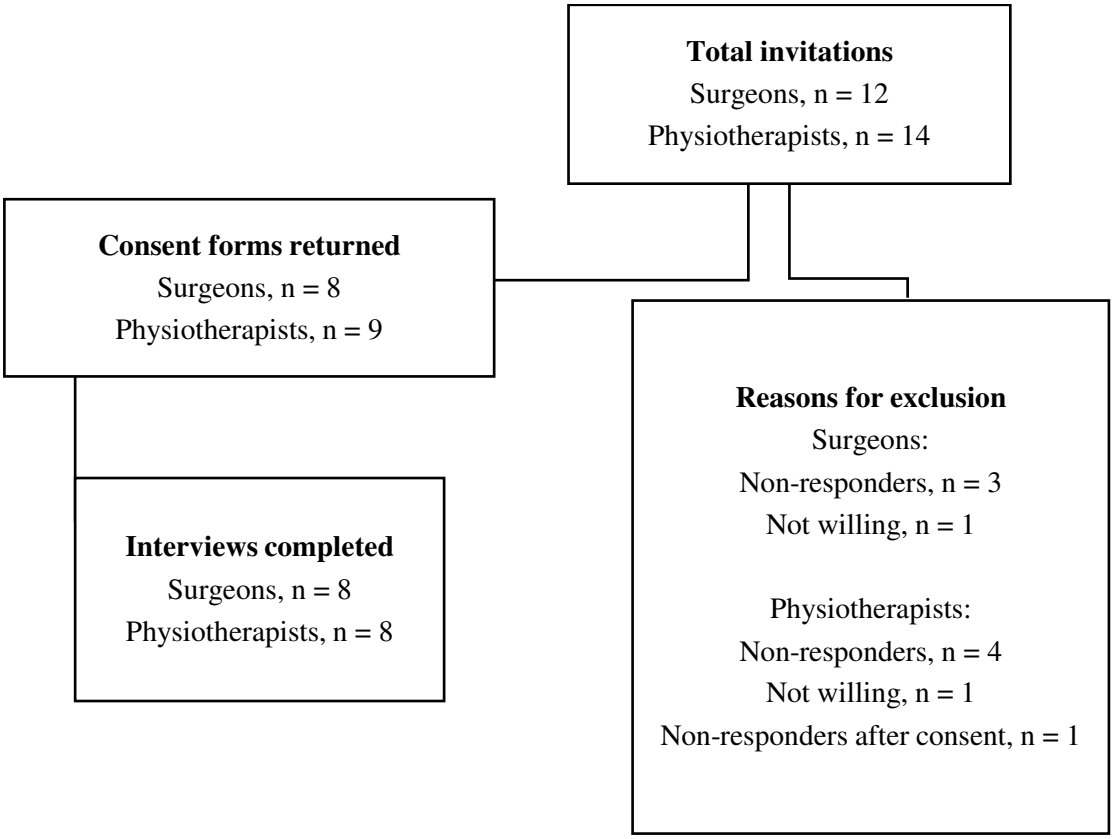


Figure 15: Flow diagram of surgeons' and physiotherapists' interviews

Table 35 Characteristics of the surgeons and physiotherapists interviewed

	Gender		Years of experience in treating shoulder conditions
	Male	**Female**	**Median (IQR[a]) years**
Surgeons	7	1	11 (7-16.25)
Physiotherapists	--	8	13.5 (12-15.75)

[a]IQR: Interquartile range

The four themes from the interviews with surgeons and physiotherapists are described below:

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 severity and duration of symptoms) in their routine practice. During the early painful phase, surgeons and physiotherapists thought that pain control by steroids would be the priority to initiate exercises. Physiotherapists also provided patient education as part of their treatment plan.

Surgeon 6: "*A humeral steroid injection is beneficial for pain relief. It allows them to do the physiotherapy. So, I do use that a lot".*

Physiotherapist 1: "*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 5: "*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".*

Surgeons felt that exercises during the early painful phase might aggravate the pain but thought that exercises after surgery are important for the recovery.

Surgeon 2: "*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 3: *"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 timescales and the pain. That's one compliment of the physiotherapies is some health professionals who can talk to them on a regular basis".*

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

Physiotherapist 4: "*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".*

In people with severe or long-standing 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 for faster recovery, low risk of humeral fractures, and familiarity with the procedure.

Surgeon 6: "*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".*

Physiotherapist 2: "*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".*

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

Physiotherapist 2: "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 5: "*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…*"

Physiotherapists also commented on the feasibility of the UK FROST physiotherapy programmes within 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.

Physiotherapist 2: "*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 4: "*... they* (Trial participants) *were seen with the 24 hours' post-surgery and they had twelve sessions which is a luxury because in our Trust, that is never, not going to happen and that never used to happen".*

Physiotherapist 7: *"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".*

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 effective as surgery.

Surgeon 3: "*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 7: "*My expectation is that physio won't work. My expectation is that the other two are probably equivocal".*

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

Surgeon 3*: "I expected my patients to get better on whichever arm they chose".*

Physiotherapist 6*: "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 8: "*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."*

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

Surgeon 5: "*I'd definitely have a hydro-dilatation group because part of your trial is trying to work out if the cheaper operation is better than the more expensive operation and hydro dilatations probably gained quite popularity since we started the trial design and it reflects current practice".*

Physiotherapist 6: "*It needs hydro-dilatation 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".*

Factors that influence treatment outcomes

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

Surgeon 3: "*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".*

Physiotherapist 5: "*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 7: "*I think pain is always the predominant thing with these patients. They just want someone to do something to help with the pain".*

Surgeons and physiotherapists felt that diabetes negatively affects treatment outcomes.

Surgeon 8: "*Well my experience with diabetics they have been very bad for a hell of a long time, okay, 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".

Physiotherapist 2: "*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".*

They also felt that participants' engagement with treatments and positive expectations lead to better outcomes.

Surgeon 3: "*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".*

Physiotherapist 5: "*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".*

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, they came with fixed ideas of what treatment they wanted in the trial.

Surgeon 6: "*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 okay with that".*

Physiotherapist 8: "*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".*

Surgeons and physiotherapists described that some people declined the trial because their previous physiotherapy did not work and therefore did not want to be randomised to the ESP.

Surgeon 8: "*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*".

Physiotherapist 6: *"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".*

Physiotherapists said that the UK FROST interventions were well-received. Surgeons and physiotherapists described that participants were surprised with the number of post-procedural physiotherapy sessions they received in the trial. Surgeons and physiotherapists described a few participants randomised to the ESP felt they were not improving.

Physiotherapist 2: "*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 4: "*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."*

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) on the trial interventions. The key findings were:

Trial participants described that frozen shoulder had a major impact on all aspects of their life. They were keen on getting 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, shoulder movements, and function. Participants who had ACR described quicker recovery than they 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: 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

Though 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 in this topic.[19, 102, 103] The problems due to frozen shoulder as described in our participants' interviews were mapped to the ICF biopsychosocial framework of disability.[99, 100] This is the first time that the ICF has been specifically used to describe functioning and disability due to frozen shoulder. Our findings support the range of problems reported in previous studies in people living with chronic shoulder pain.[19, 101-104] Our results on participants' concerns in seeking early diagnosis and referrals are comparable to a previous qualitative study in people with frozen shoulder.[19] However, in the context of variable prognosis (from self-resolving to resistant/chronic cases) of frozen shoulder, a screening tool to identify the sub-group of patients who might benefit from early referral would be helpful. Factors such as chronicity,

severity, diabetes, inability to cope with functional restriction, or pain tolerance could be incorporated to predict the need for further treatment.

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

Participants and health professionals described pain relief and improvements in function and range of motion as the main treatment outcomes to be achieved in frozen shoulder. Their priorities resonate with similar results in a previous survey of 225 health care professionals[105] and other studies.[13, 19, 106, 107]

Steroids help in pain relief and to initiate the 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 maintain the range of movements and to avoid long-standing symptoms. There is moderate evidence to support the efficacy of steroids on pain, function and disability when compared to placebo[22, 108] and additional benefit with shoulder exercises.[109] Steroids are also reported to be a potentially cost-effective option.[13]

Commitment to adhere 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.[110-112] During the early recovery phase, UK FROST participants were motivated by treatment benefits[112] and had self-determination[113] 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.[114]

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[115, 116] and health professionals.[117-119] The preferences of surgeons and physiotherapists were mainly based on their clinical experience in routine practice.[117-119] It would be highly unlikely if experts did not acquire personal preferences, especially in conditions where treatment decisions are expertise-based due to 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%.[120] Our interview findings indicate that the presence or absence of diabetes did not influence the trial participants' experience of the trial interventions. These support the main trial findings which indicate no significant between-group differences in the mean Oxford Shoulder Scores 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[121, 122] 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 a cost-effective alternative to combat the NHS waiting lists for surgery. However, the available evidence on the effects of hydrodilatation is inconclusive.[26, 123, 124] A meta-analysis[123] of seven small randomised controlled trials concluded that hydrodilatation combined with corticosteroid has no significant clinical effect on pain, disability and shoulder

movements compared to corticosteroid alone. A further RCT[26] in 50 participants with severe frozen shoulder found ACR compared with hydrodilatation improved Oxford Shoulder Score at six months. Despite the lack of sufficient evidence, hydrodilatation appears to be growing in popularity and is being increasingly used by shoulder surgeons.[125] This resonates with the views of the UK FROST surgeons. More large-scale and high-quality randomised controlled trials evaluating the clinical effectiveness and safety of hydrodilatation compared to other treatments is essential to make recommendations and guide evidence-based practice.[126]

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.[126, 127] Often, people with prolonged symptoms and those resistant to conservative treatments of at least six months are seen in secondary care and recommended for surgery as the final treatment option.[27, 127, 128] Of the two surgical procedures used in UK FROST, evidence shows that surgeons commonly perform ACR[126] for the controlled procedure of capsular release[129] and improved clinical outcomes.[130] 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

The interviews were conducted with participants who took part in UK FROST and therefore may not be relevant outside this context. Secondly, 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. Thirdly, 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 differently if face-to-face interviews were conducted. Fourthly, the response rate to participate was low across all arms. Those who did not participate might have reported different experiences.

Strengths

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 in clinical practice:

1) Frozen shoulder has a major impact on all aspects of an individual's life. A better understanding of patient 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 while some perceived physiotherapy as a low-risk intervention. These personal treatment preferences should be well understood by health professionals and opportunities should be provided to patients to address their preferences during the process of 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 have 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 complementing 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 within the NHS. More primary qualitative studies on people with frozen shoulder are needed to integrate patient perspectives in informing patient-centred care and shared decision making.

Chapter 6 Discussion and conclusion

UK FROST is the largest RCT to date that evaluates three commonly used options to treat the frozen shoulder. The trial was sufficiently powered to draw strong conclusions about the effectiveness of the treatments being compared. 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 emphasize that whilst physiotherapy is a common treatment in NHS practice, the ESP intervention was a specifically designed, standardised and new physiotherapy pathway for UK FROST trial that was based on best available evidence and expert consensus. The pragmatic, multi-centre 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 treatment options within the trial follow-up period. The primary analysis perspective is from the NHS and will have direct applicability to informing future policy and commissioning decisions in the UK. In this discussion, we begin with summarising the main results and then explore potential risks of bias that might challenge trial validity and applicability. We conclude by discussing application of the trial 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) and ESP (3.06 points, 95% CI 0.71 to 5.41) based on ITT analysis. Although statistically significant, mean estimates were short of the minimal clinically important effect size of four to five OSS points (the trial was powered for differences of four points for comparing MUA with ACR and five 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 three months, and 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 ESP and were similar to final follow-up scores by eight months. Analyses of the data incorporating all available time points for each participant (day of treatment, six months post-treatment, three, six 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 for ACR and ESP at 12 months (3.26 points in favour of ACR, 95% CI 1.18 to 5.35). The 95% CI interval still included the minimal clinically important difference for this comparison of five 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 length 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 three months but better outcomes at 12 months post-randomisation compared with MUA or ESP. There were no statistically significant differences between 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 compared with the other groups.

Fidelity of treatment

Of the participants randomised to their allocated treatment, 82% completed MUA, 80% completed ACR and 81% completed ESP. Only sixteen 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, there were only ten SAE's reported for nine participants, eight of whom were randomised to ACR and two who 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 co-morbidities or surgery in general, rather than being specifically related to the trial procedures. As an example, a stroke was diagnosed three months after ACR. Furthermore, of the eight SAE's in participants randomised to ACR there were two participants who did not have an ACR (one had an MUA and the other had a none trial physiotherapy treatment). Only one of the two participants allocated to MUA who had an SAE, actually received an MUA and the other 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 adverse events, 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 the frozen shoulder was updated.[13] The updated review focussed 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 undertaken to inform the design of UK FROST.[65] 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 at high risk of bias to blinding of participants and clinicians, and outcome assessment.[24-26, 131-135] Three trials were at high risk of bias for incomplete outcome reporting,[25, 131, 133] two trials for selective reporting[132, 133] and two trials for 'other' biases.[24, 131] Due to considerable heterogeneity of the interventions and generally limited evidence for many of the comparisons, only two trials were pooled in a meta-analysis, UK FROST and one other trial,[132] 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. Whilst it did not include hydrodilatation, evidence of hydrodilatation's effectiveness from four trials was inconclusive.[24, 26, 133, 135]

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 £6,984 per additional QALY. ACR was more costly than ESP [on average £1,733.78 more expensive per participant (95% CI 1,529.48 to 1,938.06)]. 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 had slightly lower QALYs. Therefore, given the limited differences in outcomes observed in the ACR group compared to the other two treatment options, along with its much greater costs, it is difficult to justify this 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 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 this analyses was the days off work at £113.80 a day. During the twelve month follow-up, participants allocated to ESP had a median of no days off work, MUA participants a median of six days off work, and ACR patients had a median of two weeks off work. This potentially could be related to quicker access to treatment for ESP participants, and may be important to patient decision-making. The analysis was also limited to a 12 month follow-up. However, as the results of the OSS at 12 months shows that fifty per cent of the participants were only five points off 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 on getting their shoulder problem resolved which motivated them to participate in the trial. They thought that seeking early medical help and 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 had ACR described quicker recovery than they 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 commented that some people who had received previously ineffective physiotherapy did not want 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 that there was comparability in the characteristics of the three treatment groups. There was a greater number of participants currently in paid work in the MUA arm and some group imbalance in having had a similar shoulder problem on the opposite side to the reference shoulder. A sensitivity analysis of the primary outcome that included employment status as an additional covariate, found results were similar to those observed in the primary analysis. The use of unequal random allocation reflected differential treatment effect expectations. The larger number of participants who were allocated to surgery compared with physiotherapy was to allow for a larger effect size to justify the greater costs and potential risks associated with surgery.[28]

To help ensure 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 conducted by consultant surgeons for both surgical procedures and most operating surgeons routinely performed both procedures up to once a month. Physiotherapy was delivered by qualified physiotherapists who were predominantly Band 6 across both ESP and PPP, and treating two to three frozen shoulder patients per month. It is unlikely that not including students or assistants in delivering physiotherapy introduced a bias as this was applied consistently across all treatment arms. The number of physiotherapy sessions across the three trial arms were 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 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 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.[28] There was a ceiling effect at 12 months follow-up, in that 24% of participants had regained full function (top OSS score). While it is encouraging that participants across all three treatment groups 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, the blinding of participants and clinicians to treatment allocation was not possible or desirable in this pragmatic trial. The statistician and health economist were blind to group allocation until after data were hard locked and no further changes could be made. The lack of any sub-group effect of participant baseline preferences on treatment outcome (using the OSS) may in part mitigate against concerns of 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 patients starting ESP around 14 days and for MUA and ACR around 57 days and 72 days post-randomisation, respectively. This could benefit the ESP group at the three-month follow-up when it was being compared with participants who had not yet received a trial intervention or were recovering from a surgical procedure. In order to account for differential waiting times, participants also completed the OSS on the day of treatment and six months later. Reassuringly, 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 if all three trial arms were delivered at similar times.

A further potential threat to study validity is non-compliance because the treatments were not delivered as planned in all participants. This could dilute the treatment effect observed in the intention-to-treat (ITT) primary analysis. There were only sixteen participants (3%) who crossed over to a different trial treatment, and 17 (3%) who received an alternative treatment that was not a trial intervention (e.g. steroid injection only). There were, however, around 20% of participants who did not complete their treatment across all three trial arms,

according to our defined criteria. This was expected, as the natural history is for the frozen shoulder to resolve,[136-138] particularly for participants awaiting MUA or ACR, who did not receive their allocated treatment due to waiting times of 57 and 72 days respectively. For ESP patients, despite encouraging up to 12 sessions of physiotherapy based on existing evidence,[39, 40] we used strict criteria to define 'compliers', as they had 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 less than two points on the OSS. Interestingly, unadjusted OSS at 12 months found that participants who complied with ESP scored on average five points higher on the OSS than those who did not, which is potentially clinically important. Finally, a steroid injection was delivered for all completed MUAs and 80% of patients randomised to ESP, compared with 28% of ACRs who had a steroid injection at the discretion of the surgeon. This could be argued to be a bias against ACR, but is consistent with our finding from a survey of 53 surgeons when developing the trial protocol that only 30% routinely provide a steroid injection with ACR.[28] It therefore reflects clinical practice.

Applicability of results

Characteristics of the trial population

Of 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.[136, 138] There were comparable characteristics between patients who did and did not consent to take part.

The consent rate of eligible patients was 63%. Nearly a 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. For patients who did take part, around half had no preference for a treatment but the majority of the remainder preferred surgery. This preference for wanting surgery could be explained by trial participants having already had symptoms for around eight to nine months at the time of enrolment. Moreover, nearly two-thirds of trial participants had previous physiotherapy for their affected shoulder. Whilst recognising that the trial was not powered to detect statistically significant effects between treatment allocation and sub-groups, none were found when exploring the effect on treatment outcome of participants previously having had physiotherapy or not; their treatment preferences; or duration of symptoms.

Finally, 30% of trial participants had diabetes, a common complaint in people with a frozen shoulder, ranging between 10% and 36%.[120] Diabetics tended to have poorer outcomes at all time-points, which is why we stratified for this at randomisation.[62] The sub-group analyses, however, showed that whether participants were diabetic or not 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 there is 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. Those surgeons who did operate were mostly consultants, as would be expected. Although trial physiotherapy had to be delivered by qualified physiotherapists (i.e. not 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 ensured 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 was delivered on average around 57 days and 72 days post-randomisation. Both ESP

and PPP were encouraged to be delivered as soon as possible, particularly PPP within 24 hours of surgery.

It is important to emphasize that whilst 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 groups of physiotherapy were developed using evidence from various sources[6, 13, 19, 34, 35] and consensus from an expert Delphi study,[36] which encouraged the delivery of up to 12 treatment sessions. For the ESP intervention a steroid injection was to be offered at the first opportunity, whilst for 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 be 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.[139] In the context of the trial, standardised and structured physiotherapy protocols were applied to ensure their rigorous and optimal delivery for all three treatment groups.

When designing UK FROST, a national survey of health care professionals found that only 5% used hydrodilatation to treat a frozen shoulder.[65] UK FROST therefore focused on the more urgent comparisons between ESP and 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 another upper limb orthopaedic surgical trial (ProFHER-2), we found that 52 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 is 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 high risk of bias, that compared MUA with hydrodilatation,[24, 133] and two further trials that compared it with physiotherapy and steroid injection[135] and ACR.[26] Whilst the applicability of UK FROST needs to be considered in the

context of the increasing popularity of hydrodilatation, there is a paucity of rigorous evidence to support its use.

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

Application of the trial results to clinical practice

The characteristics of the trial participants and their duration of symptoms was as expected. Therefore, at the primary end-point of 12 months on the primary outcome, it was encouraging to find that participants in all three treatment groups had improved considerably from when they were enrolled into the trial. Whilst participants did a little better in the ACR group, the mean differences between treatment options were not of the magnitude of the minimally clinically importance difference that we sought. In contrast, at the earlier time-point of three months follow-up, participants in the MUA arm did a little better than ACR participants, and the ESP participants approached a clinically important difference in its favour compared with ACR. The timing of when these interventions were delivered could explain these findings. The analyses, however, that attempted to account for variation in waiting times, illustrated that the differences between treatment options were smaller when compared with the primary analyses except for a further benefit in favour of ACR compared with ESP, and still less than the minimally clinically important difference. The findings on the secondary outcomes illustrated a similar pattern. Therefore, there is evidence of potential early benefits of ESP compared with ACR. Whilst it could be argued that this is confounded by waiting times and the surgical procedures being performed in more selective participants whose frozen shoulder had not resolved naturally, pragmatically, this reflects that there is quicker access to this intervention in clinical practice with waiting lists for surgery likely to be longer than during the trial. Importantly, participants in the nested qualitative study commented on the need to get their frozen shoulder resolved and therefore ESP offers quick access to an effective treatment. Otherwise, the evidence is inconclusive as to whether any of the three treatment

options are superior on the 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 in UK FROST. Whilst this might be potentially challenging in routine care in the NHS, effective delivery of ESP could prevent the 'opportunity cost' of using theatre resources for an MUA or an ACR and prevent the need for post-procedural physiotherapy.

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

All three treatments were similar in their completion at around 80%. There was an optimal release in 92% and 98% of participants, who had an MUA or ACR procedure, respectively. Overall, for those allocated to ACR, this was a more definitive treatment with further treatment required in 4% of participants. Nearly twice as many participants in the MUA group (7%) required further treatment and even more in the ESP group (15%). Serious complications were rare, although the ACR group was relatively less safe (4%). Only two participants allocated to MUA had a serious complication (1%). One of the participants in the ACR group diagnosed with a deep vein thrombosis actually received non-trial physiotherapy. There was, therefore, only a marginal difference in the safety profile between MUA and ESP for which in the latter group there were 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 single trials being available, often in a single centre with small sample sizes, considerable presence of bias, and heterogeneity in treatment interventions. None of the included trials helped to produce conclusive findings about the effectiveness of the interventions evaluated in UK FROST. Whilst there has been an increase in the popularity of hydrodilatation in clinical practice, and an increase in research in this area, evidence of its effectiveness in the systematic review was inconclusive. In this context, it will need to be considered as a treatment option for patients with a frozen shoulder.

Finally, ESP was the least expensive intervention, as most participants did not require a surgical procedure. MUA was the second most expensive treatment option, with participants spending a third of the time in theatre compared with ACR, for which the latter was by far the most expensive option. MUA, however, accrued 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 £20,000/QALY threshold if commissioners of services would be willing to pay £6,984 per additional QALY.

Conclusion

UK FROST has provided robust clinically relevant evidence that none of the three treatments were clearly superior on 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 when compared to the other two interventions. MUA produced the best QALYs overall. At a modest additional cost, it is the most cost-effective option to the NHS with an ICER of £6,984 per additional QALY. Patients who receive ACR are 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 are also lengthier since the trial. This may have meant that participants with a more resistant frozen shoulder were those operated on. It also could be challenging to implement the ESP pathway in clinical practice to the same optimal timing of access and 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 its inclusion 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 whilst physiotherapy was perceived as a low-risk alternative. Given patient preferences for different treatment options and the trial findings, we propose the RCT is a three-arm trial that compares hydrodilatation versus ESP with steroid injection versus MUA with a steroid injection followed with PPP as the latter 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 assess for stiffness. Finally, in clinical practice it could be a complex discussion between 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 have not been possible. A general thanks also to all those people whose commitment and efforts in the design and conduct of the UK FROST trial 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 on the study at the University of York, Mr Joe Millar who was the R&D 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 thanks 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 on the Trial Management Group and also to the shoulder research users group at James Cook University Hospital.

We thanks the British Elbow and Shoulder Society who 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 Health Technology Assessment (HTA) programme (project number 13/26/01).

This report presents independent research commissioned 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 Health Technology Assessment programme or the Department of Health. The views and opinions expressed by the interviewees in this publication are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, MRC, CCF, NETSCC, the Health Technology Assessment programme or the Department of Health.

Contribution of authors

Dr Stephen Brealey (Research Fellow, Trial Manager) was a co-applicant and trial manager who 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.

Dr Matthew Northgraves (Research Fellow, Trial Co-ordinator) contributed to the day-today 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 (T&O 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 (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 duration of the study, conducted the statistical analysis and contributed to and commented on drafts of the report.

Mrs Belen Corbacho (Research Fellow, Health Economist) lead the write up of the chapter on the economic evaluation and commented on drafts of the report.

Mrs Lorna Goodchild (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 (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 (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 (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 (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 (Consultant Orthopaedic Surgeon) was a co-applicant and consultant Orthopaedic Surgeon who was the Principal Investigator for 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 (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 Scottish centres, and commented on drafts of the report.

Professor Andrew Carr (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 (Operational and Trials Development Lead) was a trial coordinator who contributed to the conduct of the trial and management of participating sites, development and implementation of the tissue sub-study and contributed to and commented on all drafts of the report.

Professor Joseph Dias (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 (Trial Co-ordinator) made a substantial contribution to the acquisition of data as well as commenting on drafts of the report.

Professor Catherine Hewitt (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 Sallie Lamb (Professor and Director of OCTRU) was 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 (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 (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 (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 (Orthopaedic Research Manager) was a trial co-ordinator who contributed to the study implementation at multiple sites throughout Scotland.

Professor Sally Spencer (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 (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 (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 (Professor of Orthopaedic Surgery) was the lead applicant and Chief Investigator who 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.

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 Asim 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 (The Royal Liverpool and Broadgreen University Hospitals NHS Trust)
Mr Kapil Kumar (Aberdeen Woodend Hospital 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 (Bedford Hospital NHS Trust)
Mr Tim Matthews (University Hospital of Wales, Cardiff & Vale University Health Board)
Mr Damian McClelland (University Hospitals of North Midlands)
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 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)

Prof 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 Park 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, 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, KUMAR, 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, 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 TSURO, 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, M KAAN

Trial Steering Committee independent members

Professor Chris Rogers (Chair – 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 CCG).

Data Monitoring Committee independent members

Professor John Norrie (Chair – Chair of Medical Statistics and Trial Methodology, University of Edinburgh)

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

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

Data sharing agreement

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.
2. Hand GC, Athanasou NA, Matthews T, Carr AJ. The pathology of frozen shoulder. *J Bone Joint Surg Br* 2007;**89**:928-32. http://dx.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.
4. Hand C, Clipsham K, Rees JL, Carr AJ. Long-term outcome of frozen shoulder. *J Shoulder Elbow Surg* 2008;**17**:231-6.
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. http://dx.doi.org/10.1136/bjsm.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.
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.
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. http://dx.doi.org/10.1002/art.20535
9. Bunker TD. Time for a new name for 'frozen shoulder'. *Br Med J (Clin Res Ed)* 1985;**290**:1233-4.
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. http://dx.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 (Bristol, Avon)* 2004;**19**:810-8. http://dx.doi.org/10.1016/j.clinbiomech.2004.05.006
12. Codman E. *Rupture of the supraspinatus tendon and other lesion sin or about the subacromial bursa*. Boston: 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 Technology Assessment* 2012;**16**:1-264.
14. Walmsley S, Osmotherly P, Rivett D. Movement and pain patterns in early stage primary/idiopathic adhesive capsulitis: a factor analysis. *Physiotherapy* 2014;**100**:336-43.
15. Rundquist P, Anderson D, Guanche C, Ludewig P. Shoulder kinematics with frozen shoulder. *Arch Phys Med Rehabil* 2003;**84**:1473-9.
16. Rundquist P, Ludewig PM. Patterns of motion loss in subjects with idiopathic loss of shoulder range of motion. *Clinical Biomechanics* 2004;**19**:810-8.
17. Wolf EM, Cox WK. The external rotation test in the diagnosis of adhesive capsulitis. *Orthopedics* 2010;**33**. http://dx.doi.org/10.3928/01477447-20100329-11
18. Noboa E, López-Graña G, Barco R, S A. 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.
19. 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.
20. Linsell L, Dawson J, Zondervan K, Rose P, Randall T, Fitzpatrick R*, et al.* Prevalence and incidence of adults consulting for shoulder conditions in UK primary care; patterns of diagnosis and referral. *Rheumatology (Oxford)* 2006;**45**:215-21. http://dx.doi.org/10.1093/rheumatology/kei139

21. 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. http://dx.doi.org/10.1186/1471-2474-6-4
22. Rangan A, Hanchard N, McDaid C. What is the most effective treatment for frozen shoulder? *Bmj* 2016;**354**:i4162. http://dx.doi.org/10.1136/bmj.i4162
23. Dennis L BS, Rangan A, Rookmoneea M, Watson J. Managing idiopathic frozen shoulder: a survey of health professionals' current practice and research priorities. *Shoulder and Elbow* 2010;**2**:294-300.
24. 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. http://dx.doi.org/10.1302/0301-620x.89b9.18863
25. Kivimaki J, Pohjolainen T, Malmivaara A, Kannisto M, Guillaume J, Seitsalo S*, et al.* 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. http://dx.doi.org/10.1016/j.jse.2007.02.125
26. 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. 2018;**27**:1401-6. http://dx.doi.org/https://dx.doi.org/10.1016/j.jse.2018.04.002
27. 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. http://dx.doi.org/10.1016/j.jse.2013.01.010
28. 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.
29. Rangan A, Goodchild L, Gibson J, Brownson P, Thomas M, Rees J*, et al.* BESS/BOA patient care pathways. *Shoulder & Elbow* 2015;**7**:299-307.
30. New Zealand Guidelines Group. *The diagnosis and management of soft tissue shoulder injuries and related disorders*. Wellington: New Zealand Guidelines Group; 2004.
31. 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.
32. 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 Technology Assessment* 2015;**19**:1-218.
33. 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.
34. 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.
35. 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.
36. 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* 2019;**107**:150-60.
37. Babcock HM, Matava MJ, Fraser V. Postarthroscopy surgical site infections: review of the literature. *Clin Infect Dis* 2002;**34**:65-71.

38. 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:doi: 10.1093/rheumatology/kez007.
39. 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 placebocontrolled trial. *Arthritis Rheum* 2003;**48**:829-38.
40. 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 (Oxford)* 2005;**44**:529-35.
41. Bloom JE, Rischin A, Johnston RV, Buchbinder R. Image-guided versus blind glucocorticoid injection for shoulder pain. *Cochrane Database Systematic Reviews* 2012;**8**.
42. Dawson J, Fitzpatrick R, Carr AJ. Questionnaire on the perceptions of patients about shoulder surgery. *J Bone Joint Surg (Br)* 1996;**78**:593-600.
43. Dawson J, Rogers K, Fitzpatrick R, Carr AJ. The Oxford Shoulder Score revisited. *Arch Orthop Trauma Surg* 2009;**129**:119-23.
44. Constant CR, Murley AH. A clinical method of functional assessment of the shoulder. *Clin Orthop Relat Res* 1987;**214**:160-4.
45. Dawson J, Hill G, Fitzpatrick R, Carr AJ. The benefits of using patient-based methods of assessment. *J Bone Joint Surg (Br)* 2001;**83**:877-82.
46. Kennedy CA, Beaton DE, Solway S, McConnell S, Bombardier C. *The DASH outcome measure user's manual*. Toronto: Institute for Work and Health; 2011.
47. Beaton DE, Wright JG, Katz JN. Development of the QuickDASH: comparison of three item-reduction approaches. *J Bone Joint Surg Am* 2005;**87**:1038-46.
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**:100-1018.
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.
50. Brooks R. EuroQol: the currrent state ofplay. *Health Policy* 1996;**37**:53-72.
51. Group EQ. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy* 1990;**16**:119-208.
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.
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 the EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;**20**:1727-36.
54. Angst F, Schwyzer H-K, Aeschlimann A, Simmen BR, Goldhahn J. Measures of adult shoulder function. *Arthritis Care Res* 2011;**63**:S174-88.
55. Holmgren T, H. BH, Oberg B, Adolfsson L, Johansson K. Effect of specific exercise strategy on need for surgery in patients with subacromial impingement syndrome: randomised controlled study. *British Medical Journal* 2012;**344**:e787.
56. Farrar JT, Young JJP, 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.
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.
58. National Institute for Health and Care Excellence. *Surgical site infection: prevention and treatment of surgical site infection: CG74*. London: National Institute of Clinical Excellence; 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.
60. Schulz KF, Grimes DA. Multiplicity in randomised trials I: endpoints and treatments. *Lancet* 2005;**365**:1591-5.
61. Hewitt CE, Torgerson DJ, Miles JN. Individual allocation had an advantage over cluster randomization in statistical efficiency in some circumstances. *Journal of Clinical Epidemiology* 2008;**61**:1004-8.
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-A**:1398-407.
63. StatCorp. *Stata Statistical Software: Release 15*. College Station, Tx: StataCorp LLC; 2017.
64. Dawson J, Harris KK, Doll H, Fitzpatrick R, Carr AJ. 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.
65. Dennis L, Brealey S, Rangan A, Rookmoneea M, Watson J. Managing idiopathic frozen shoulder: a survey of health professionals' current practice and research priorities. *Shoulder & Elbow* 2010;**2**:294-300.
66. Ashby R, Turner G, Cross B, Mitchell N, Torgerson D. A randomized trial of electronic reminders showed a reduction in time to respond to postal questionnaires. *J Clin Epidemiol* 2011;**64**:208-12.
67. Mitchell N, Hewitt CE, Lenaghan E, Platt E, Shepstone L, Torgerson DJ*, et al.* Prior notification of trial participants by newsletter increased response rates: a randomized controlled trial. *J Clin Epidemiol* 2012;**65**:1348-52.
68. Brooks R, Group E. EuroQol: the current state of play. *Health policy* 1996;**37**:53-72.
69. Van Hout B, Janssen M, Feng Y-S, 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 in health* 2012;**15**:708-15.
70. National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal*. London: National Institute of Clinical Excellence; 2013.
71. Manca A, Palmer S. Handling missing data in patient-level cost-effectiveness analysis alongside randomised clinical trials. *Applied health economics and health policy* 2005;**4**:65-75.
72. Curtis L. Unit costs of health and social care 2012. Canterbury: Personal Social Services Research Unit, University of Kent; 2013. In; 2015.
73. Joint Formulary Committee. *British national formulary (BNF) 66*: Pharmaceutical Press; 2013.
74. Department of Health NRC. 2015.
75. Office for National Statistics. *EARN01: Average Weekly Earnings. London: Office for National Statistics*; 2019

76. 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.
77. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health economics* 2005;**14**:487-96.
78. Billingham L, Abrams KR, Jones DR. Methods for the analysis of quality-of-life and survival data in health technology assessment. 1999.
79. 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.
80. 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.

81. 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.
82. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Statistics in medicine* 2011;**30**:377-99.
83. Rubin DB. *Multiple imputation for nonresponse in surveys*: John Wiley & Sons; 2004.
84. Johannesson M, Weinstein MC. On the decision rules of cost-effectiveness analysis. *Journal of health economics* 1993;**12**:459-67.
85. Manca A, Lambert PC, Sculpher M, Rice N. Cost-effectiveness analysis using data from multinational trials: the use of bivariate hierarchical modeling. *Medical Decision Making* 2007;**27**:471-90.
86. Gomes M, Grieve R, Nixon R, Ng ESW, Carpenter J, Thompson SG. Methods for covariate adjustment in cost-effectiveness analysis that use cluster randomised trials. *Health economics* 2012;**21**:1101-18.
87. 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, England: University of York* 2013.
88. 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 in Health* 2018;**21**:266-75.
89. 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. http://dx.doi.org/10.1186/s13063-018-2983-y
90. 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**. http://dx.doi.org/10.1136/bmjopen-2013-002889
91. Holmberg C, Karner JJ, Rappenecker J, Witt CM. Clinical trial participants' experiences of completing questionnaires: a qualitative study. *BMJ Open* 2014;**4**:e004363. http://dx.doi.org/10.1136/bmjopen-2013-004363
92. 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. http://dx.doi.org/10.1016/j.physio.2017.09.002
93. Prothero L, Sturt J, de Souza S, Lempp H. Intensive management for moderate rheumatoid arthritis: a qualitative study of patients' and practitioners' views. *BMC Rheumatol* 2019;**3**:12. http://dx.doi.org/10.1186/s41927-019-0057-8
94. Carter B. Clients' experiences of frozen shoulder and its treatment with Bowen technique. *Complement Ther Nurs Midwifery* 2002;**8**:204-10. http://dx.doi.org/10.1054/ctnm.2002.0645
95. Baker SE, Edwards, Rosalind. *How many qualitative interviews is enough? Expert voices and early career reflections on sampling and cases in qualitative research*. National Centre for Research Methods Review Paper: Expert voices and early career reflections on sampling and cases in qualitative research,. GB: (National Centre for Research Methods Reviews) Southampton, GB.; 2012. http://dx.doi.org/http://eprints.ncrm.ac.uk/2273/.
96. Charmaz K. *Constructing Grounded Theory – A Practical Guide through Qualitative Analysis* SAGE; 2006.
97. Rindflesch AB. A grounded-theory investigation of patient education in physical therapy practice. *Physiother Theory Pract* 2009;**25**:193-202. http://dx.doi.org/10.1080/09593980902776613
98. Software Nqda. QSR International Pty Ltd. In. 11 ed; 2015.
99. WHO. *Towards a common language for functioning, disability, and health: ICF. The international classification of functioning, disability and health*; 2002.
100. 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.

101. 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. http://dx.doi.org/10.2340/16501977-1159
102. 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.
103. Toprak M, Erden M. Sleep quality, pain, anxiety, depression and quality of life in patients with frozen shoulder1. *J Back Musculoskelet Rehabil* 2019;**32**:287-91. http://dx.doi.org/10.3233/bmr-171010
104. 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 (Oxford)* 2019; 10.1093/rheumatology/kez046. http://dx.doi.org/10.1093/rheumatology/kez046
105. 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. http://dx.doi.org/10.1007/s11136-014-0708-6
106. 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 (Hoboken)* 2018;**70**:252-9. http://dx.doi.org/10.1002/acr.23254
107. 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. http://dx.doi.org/10.1016/j.jclinepi.2015.06.006
108. 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 (Baltimore)* 2016;**95**:e3469. http://dx.doi.org/10.1097/md.0000000000003469
109. Page MJ, Green S, Kramer S, Johnston RV, McBain B, Chau M*, et al.* Manual therapy and exercise for adhesive capsulitis (frozen shoulder). *Cochrane Database Syst Rev* 2014; 10.1002/14651858.cd011275:Cd011275. http://dx.doi.org/10.1002/14651858.cd011275
110. 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* 2018; 10.1136/bjsports-2017-098742. http://dx.doi.org/10.1136/bjsports-2017-098742
111. 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. http://dx.doi.org/10.1136/bjsports-2016-096084
112. Picorelli AM, Pereira LS, Pereira DS, Felicio 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. http://dx.doi.org/10.1016/j.jphys.2014.06.012
113. Teixeira PJ, Carraca 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. http://dx.doi.org/10.1186/1479-5868-9-78
114. 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. http://dx.doi.org/10.2522/ptj.20150229
115. Patients' preferences within randomised trials: systematic review and patient level meta-analysis. *BMJ* 2008;**337**:a1864. http://dx.doi.org/10.1136/bmj.a1864
116. 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. http://dx.doi.org/10.1186/s12874-017-0304-x

117. 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. http://dx.doi.org/10.1179/106698111x12899036752014
118. Ghogawala Z BF, Carter BS. . Clinical equipoise and the surgical randomized controlled trial. . *Journal* 2008. http://dx.doi.org/https://doi.org/10.1227/01.neu.0000333288.21621.a9
119. 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.
120. 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. http://dx.doi.org/10.1177/1758573218815002
121. Barbosa F, Swamy G, Salem H, Creswell T, Espag M, Tambe A*, et al.* Chronic adhesive capsulitis (Frozen shoulder): Comparative outcomes of treatment in patients with diabetes and obesity. *J Clin Orthop Trauma* 2019;**10**:265-8. http://dx.doi.org/10.1016/j.jcot.2018.02.015
122. Hsu CL, Sheu WH. Diabetes and shoulder disorders. *J Diabetes Investig* 2016;**7**:649-51. http://dx.doi.org/10.1111/jdi.12491
123. 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. http://dx.doi.org/10.1177/1457496918772367
124. 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. http://dx.doi.org/10.1177/1758573216687273
125. Heire P BR, Mubashar M, Bhatti W. . Ultrasound-guided hydrodilatation for adhesive capsulitis–a step-by-step guide. Paper presented at: European Congress of Radiology
126. Guyver PM, Bruce DJ, Rees JL. Frozen shoulder - A stiff problem that requires a flexible approach. *Maturitas* 2014;**78**:11-6. http://dx.doi.org/10.1016/j.maturitas.2014.02.009
127. Arce G. Primary Frozen Shoulder Syndrome: Arthroscopic Capsular Release. *Arthrosc Tech* 2015;**4**:e717-20. http://dx.doi.org/10.1016/j.eats.2015.06.004
128. 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. http://dx.doi.org/10.1177/1758573216676786
129. 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. http://dx.doi.org/10.5604/15093492.1143530
130. Kanbe K. Clinical outcome of arthroscopic capsular release for frozen shoulder: essential technical points in 255 patients. *J Orthop Surg Res* 2018;**13**:56. http://dx.doi.org/10.1186/s13018-018-0758-5
131. Smitherman JA, Struk AM, Cricchio M, McFadden G, Dell RB, Horodyski M*, et al.* Arthroscopy and manipulation versus home therapy program in treatment of adhesive capsulitis of the shoulder: a prospective randomized study. 2015;**24**:69-74.
132. De Carli A, Vadala A, Perugia D, Frate L, Iorio C, Fabbri M*, et al.* Shoulder adhesive capsulitis: manipulation and arthroscopic arthrolysis or intra-articular steroid injections? *International Orthoapedics* 2012;**36**:101-6. http://dx.doi.org/https://dx.doi.org/10.1007/s00264-011-1330-7
133. 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. http://dx.doi.org/10.1016/j.jse.2009.02.002
134. Mukherjee RN, Pandey RM, Nag HL, Mittal R. Frozen shoulder - A prospective randomized clinical trial. 2017;**8**:394-9. http://dx.doi.org/https://dx.doi.org/10.5312/wjo.v8.i5.394
135. 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. 2016;**25**:1937-43. http://dx.doi.org/https://dx.doi.org/10.1016/j.jse.2016.09.021
136. Vastamaki H, Kettunen J, Vastamaki M. The Natural History of Idiopathic Frozen Shoulder - A 2- to 27-year Followup Study. *Clin Orthop Relat Res* 2012;**470**:1133–43
137. Diereks 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.
138. Reeves B. The natural history of the frozen shoulder syndrome *Scand J Rheum* 1975;**4**:193–6. .
139. 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.

Appendices

Appendix 1 Participating Trusts

Basildon and Thurrock University Hospitals NHS Foundation Trust

Bedford Hospital NHS 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 Park 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

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

The Royal Liverpool and Broadgreen University Hospitals NHS 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

Appendix 2 Table of amendments

Type (Non-substantial or Substantial)	Approved date	Documents amended	Brief Description of Amendment
Substantial Amendment 1	12/01/15	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)	1. Clarification in the trial Participant Information Sheet about the possible need for further treatment after all three treatments. Also clarification on the £5 unconditional payment at 12 months is for all treatment groups. 2. Addition of Home Exercise Leaflet. 3. Change to trial protocol to be explicit that will adjust for age, gender and diabetes and that an exploratory sub group analysis will be done for the presence of diabetes.
Substantial Amendment 2	24/05/16	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)	1. Change to allow trial poster to be available in a public part of the hospital. 2. Permission to for hospitals to publicise the trial through initiatives such as the 'OK to ask campaign'. 3. Amended the protocol with additional sub group analysis as proposed by DMEC. 4. Amended the protocol to update sites on what treatment the participant will have whilst awaiting surgery. 5. Amended protocol for Pre-treatment form in the ESP group to complete either on first day of physiotherapy or before steroid injection, whichever is first. 6. Inclusion of text messaging SWAT at three 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	17/08/16	N/A	1. Addition of new participating sites.

Type (Non-substantial or Substantial)	**Approved date**	**Documents amended**	**Brief Description of Amendment**
Substantial Amendment 3	21/12/16	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 provided, to allow us to undertake a nested shoulder capsular tissue and blood study within the host trial.
Non-substantial Amendment 2	13/10/17	N/A	1. Change in PI at Forth Valley and Basildon sites

Appendix 3 Recruitment

Table 36: Recruitment by Site

Site	Screened n	Randomised n (% of screened)	Withdrawn n (% of randomised)
Site1	58	12 (21%)	0 (0%)
Site2	9	6 (67%)	0 (0%)
Site3	12	7 (58%)	0 (0%)
Site4	8	6 (75%)	0 (0%)
Site5	45	11 (24%)	0 (0%)
Site6	8	5 (63%)	0 (0%)
Site7	49	17 (35%)	1 (6%)
Site8	8	4 (50%)	0 (0%)
Site9	11	9 (82%)	2 (22%)
Site10	12	12 (100%)	1 (8%)
Site11	20	18 (90%)	1 (6%)
Site12	4	0 (0%)	n/a
Site13	79	34 (43%)	1 (3%)
Site14	17	11 (65%)	4 (36%)
Site15	7	7 (100%)	0 (0%)
Site16	15	1 (7%)	0 (0%)
Site17	48	45 (94%)	1 (2%)
Site18	15	7 (47%)	0 (0%)
Site19	48	22 (46%)	0 (0%)
Site20	2	1 (50%)	0 (0%)
Site21	16	14 (88%)	0 (0%)
Site22	18	11 (61%)	0 (0%)
Site23	10	3 (30%)	0 (0%)
Site24	58	26 (45%)	0 (0%)
Site25	12	5 (42%)	0 (0%)
Site26	26	18 (69%)	0 (0%)
Site27	3	0 (0%)	n/a
Site28	10	9 (90%)	0 (0%)
Site29	69	49 (71%)	2 (4%)
Site30	13	5 (38%)	0 (0%)
Site31	52	27 (52%)	2 (7%)
Site32	35	16 (46%)	1 (6%)
Site33	11	5 (45%)	0 (0%)
Site34	32	23 (72%)	1 (4%)
Site35	34	32 (94%)	0 (0%)
Site36	32	22 (69%)	2 (9%)
Site37	8	3 (38%)	0 (0%)
Total	914	503 (55%)	19 (4%)

Appendix 4 Practitioner Characteristics

Table 37: Practitioner characteristics

	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 or more	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)

			ESP N = 80, Average Number of Sessions = 8.7		MUA N = 158, Average Number of Sessions = 7.9		ACR N = 156, Average Number of Sessions = 8.3	
			Predominant Pain[a] Average Sessions = 4.9	Predominant Stiffness Average Sessions = 3.8	Predominant Pain Average Sessions = 4.0	Predominant Stiffness Average Sessions = 3.8	Predominant Pain Average Sessions = 5.1	Predominant Stiffness Average Sessions = 3.2
	Patients Recording Problem at Least Once		72	63	140	134	146	120
Treatment Given	Advice and Education	No. of Patients (%) Mean No. of Sessions (SD) Median(Min, Max) % of tot. sessions[b]	71 (99%) 5.3 (3.4) 5 (1,12) 96% (378/395)	63 (100%) 4.6 (2.4) 4 (1,12) 96 % (288/300)	139 (99%) 4.5 (3.5) 4 (1,17) 97% (621/639)	134 (100%) 4.2 (3.3) 3 (1,17) 94% (568/602)	146 (100%) 5.3 (3.5) 4 (1,15) 97% (778/802)	118 (98%) 4.0 (2.8) 3 (1,12) 95% (468/492)
	Home Exercises (instruction/ review)	No. of Patients (%) Mean No. of Sessions (SD) Median(Min, Max) % of tot. sessions	64 (89%) 4.6 (3.4) 4 (1,12) 75% (295/395)	55 (87%) 4.3 (2.8) 4 (1,12) 80% (239/300)	139 (99%) 4.4 (3.5) 3 (1,17) 97% (617/639)	134 (100%) 4.3 (3.3) 3 (1,17) 95% (572/602)	146 (100%) 5.3 (3.5) 4 (1,15) 96% (772/802)	119 (99%) 4.0 (2.7) 4 (1,12) 96% (473/492)
	Supervised Exercises (gentle active/ self-assisted)	No. of Patients (%) Mean No. of Sessions (SD) Median(Min, Max) % of tot. sessions	71 (99%) 5.3 (3.4) 5 (1,12) 95% (374/395)	61 (97%) 4.6 (2.5) 4 (1,12) 93% (280/300)	132 (94%) 4.2 (3.4) 3 (1,17) 86% (552/639)	123 (92%) 3.8 (3.1) 3 (1,17) 78% (471/602)	142 (97%) 5.0 (3.6) 4(1,15) 89% (711/802)	108 (90%) 3.8 (2.6) 3 (1,12) 83% (406/492)
	Supervised Exercises (function based)	No. of Patients (%) Mean No. of Sessions (SD) Median(Min, Max) % of tot. sessions	12 (17%) 3.2 (2.8) 2 (1,9) 10% (38/395)	60 (95%) 3.5 (2.3) 3 (1,11) 70% (211/300)	64 (46%) 3.3 (3.2) 2 (1,13) 33% (214/639)	109 (81%) 3.5 (3.1) 2 (1,17) 64% (386/602)	77 (53%) 3.3 (2.8) 2 (1,13) 31% (251/802)	99 (83%) 3.3 (2.6) 3 (1,12) 66% (326/492)
	Manual Shoulder Mobilisation	No. of Patients (%) Mean No. of Sessions (SD) Median(Min, Max) % of tot. sessions	8 (11%) 1.9 (1.8) 1 (1, 6) 4% (15/395)	17 (27%) 2.1 (1.1) 2 (1,4) 12% (36/300)	82 (59%) 3.2 (2.1) 3 (1,10) 41% (264/639)	82 (61%) 3.4 (2.9) 2 (1,15) 47% (280/602)	90 (62%) 3.3 (2.6) 2 (1,11) 37% (293/802)	71 (59%) 2.9 (2,2) 2 (1, 11) 41% (204/492)
	Posture Correction	No. of Patients (%) Mean No. of Sessions (SD) Median(Min, Max) % of tot. sessions	32 (44%) 3.2 (2.5) 2 (1,10) 26% (101/395)	23 (27%) 2.3 (1.5) 2 (1,6) 18% (54/300)	57 (41%) 2.6 (2.1) 2 (1,9) 23% (149/639)	47 (35%) 2.9 (2.6) 2 (1,11) 22% (135/602)	68 (47%) 3.1 (2.4) 2 (1,13) 26% (208/802)	51 (43%) 2.5 (2.2) 2 (1,11) 26% (128/492)
	Other	No. of Patients (%) Mean No. of Sessions (SD) Median (Min, Max) % of tot. sessions	27 (38%) 2.6 (2.7) 1 (1,12) 17% (69/395)	22 (35%) 2.9 (2.5) 2 (1,10) 21% (64/300)	61 (44%) 2.6 (2.5) 2(1,12) 25% (158/639)	34 (25%) 2.6 (2.2) 2 (1,9) 15% (89/602)	50 (34%) 3.2 (3.4) 1.5 (1,15) 20% (158/802)	29 (24%) 2.8 (2.6) 2 (1,11) 16% (81/492)

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

[b] Percentage out of the total sessions for predominant pain (or stiffness), including sessions during which the particular treatment was not given

Table 39: Physiotherapy elements received (pain specific treatments)

			ESP	MUA	ACR
			Predominant Pain[a] Average Sessions = 4.9	**Predominant Pain Average Sessions = 4.0**	**Predominant Pain Average Sessions = 5.1**
Treatment Given	Patients Recording Problem at Least Once		72	140	146
	Hydrotherapy	No. of Patients (%) Mean No. of Sessions (SD) Median (Min, Max) % of tot. sessions[b]	5 (7%) 3.6 (2.6) 4 (1,7) 5% (18/395)	8 (6%) 4.0 (3.1) 3.5 (1,8) 5% (32/639)	8 (5%) 4.8 (2.1) 5 (1,7) 5% (38/802)
	Relaxation Techniques	No. of Patients (%) Mean No. of Sessions (SD) Median (Min, Max) % of tot. sessions	32 (44%) 2.3 (1.8) 1 (1,8) 18% (72/395)	17 (12%) 2.1 (1) 2 (1,4) 5% (35/639)	25 (17%) 1.8 (1.4) 1 (1,6) 6% (45/802)
	Superficial Cold	No. of Patients (%) Mean No. of Sessions (SD) Median (Min, Max) % of tot. sessions	15 (21%) 2.3 (1.4) 2 (1,6) 9% (35/395)	7 (5%) 1.6 (1.1) 1 (1,4) 2% (11/639)	6 (4%) 2.8 (2.6) 1.5 (1,7) 2% (17/802)
	TENS	No. of Patients (%) Mean No. of Sessions (SD) Median (Min, Max) % of tot. sessions	9 (13%) 2 (1) 2 (1,4) 5% (18/395)	4 (3%) 2.5 (1.7) 2 (1,5) 2% (10/639)	2 (1%) 1.5 (0.7) 1.5 (1,2) 0.4% (3/802)
	Trigger Point Therapy	No. of Patients (%) Mean No. of Sessions (SD) Median (Min, Max) % of tot. sessions	21 (29%) 3.2 (2.5) 2 (1,10) 17% (67/395)	16 (11%) 2.1 (1.5) 2 (1,6) 5% (34/639)	19 (13%) 1.7(1.1) 1 (1,4) 4% (33/802)

[a] Columns relating to predominant pain also include sessions for which patients indicated pain and stiffness equally, as 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)

			ESP	MUA	ACR
			Predominant Stiffness Average Sessions = 3.8	**Predominant Stiffness Average Sessions = 3.8**	**Predominant Stiffness Average Sessions = 3.2**
Treatment Given	Patients Recording Problem at Least Once		63	134	120
	Supervised Exercises (Stretching)	No. of Patients (%) Mean No. of Sessions (SD) Median (Min, Max) % of tot. sessions[a]	1 (2%) 4.0 (-) 4 (4,4) 1% (4/300)	75 (56%) 3.0 (2.7) 2 (1,15) 37% (223/602)	67 (56%) 2.7(2.4) 2 (1,11) 36% (178/492)
	Supervised Exercises (Strengthening)	No. of Patients (%) Mean No. of Sessions (SD) Median (Min, Max) % of tot. sessions	20 (32%) 2.0 (1.3) 1.5 (1,5) 13% (40/300)	59 (44%) 3.1 (2.6) 2 (1,11) 30% (182/602)	52 (43%) 2.4 (2) 2 (1,8) 26% (126/492)
	Soft Tissue Techniques	No. of Patients (%) Mean No. of Sessions (SD) Median (Min, Max) % of tot. sessions	9 (14%) 2.6 (1.4) 3 (1,5) 8% (23/300)	36 (27%) 2.6 (2.1) 2 (1,8) 15% (92/602)	34 (28%) 2.3 (1.8) 2(1,8) 16% (77/492)
	PNF	No. of Patients (%) Mean No. of Sessions (SD) Median (Min, Max) % of tot. sessions	22 (35%) 2.3 (2) 1 (1,8) 17% (50/300)	16 (12%) 2.6 (1.8) 2 (1,6) 7% (42/602)	22 (18%) 2.4 (1.9) 2 (1,9) 11% (53/492)
	Spinal/ Scapulothoracic Manual Therapy	No. of Patients (%) Mean No. of Sessions (SD) Median (Min, Max) % of tot. sessions	22 (35%) 2.0 (1.3) 2 (1,5) 15% (45/300)	21 (16%) 2.8(1.8) 2 (1,7) 10% (59/602)	16 (13%) 2.1 (1.7) 1 (1,7) 7% (33/492)

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

Appendix 6 OSS subdomains

Table 41: Unadjusted OSS Pain Subdomain by Treatment Arm

	MUA	ACR	ESP	Total
Baseline				
N	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
Min, Max	0, 16	0, 12	0, 11	0, 16
3 Months				
N	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
Min, Max	0, 16	0, 16	0, 16	0, 16
6 Months				
N	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
Min, Max	0, 16	1, 16	0, 16	0, 16
12 Months				
N	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
Min, Max	1, 16	1, 16	1, 16	1, 16

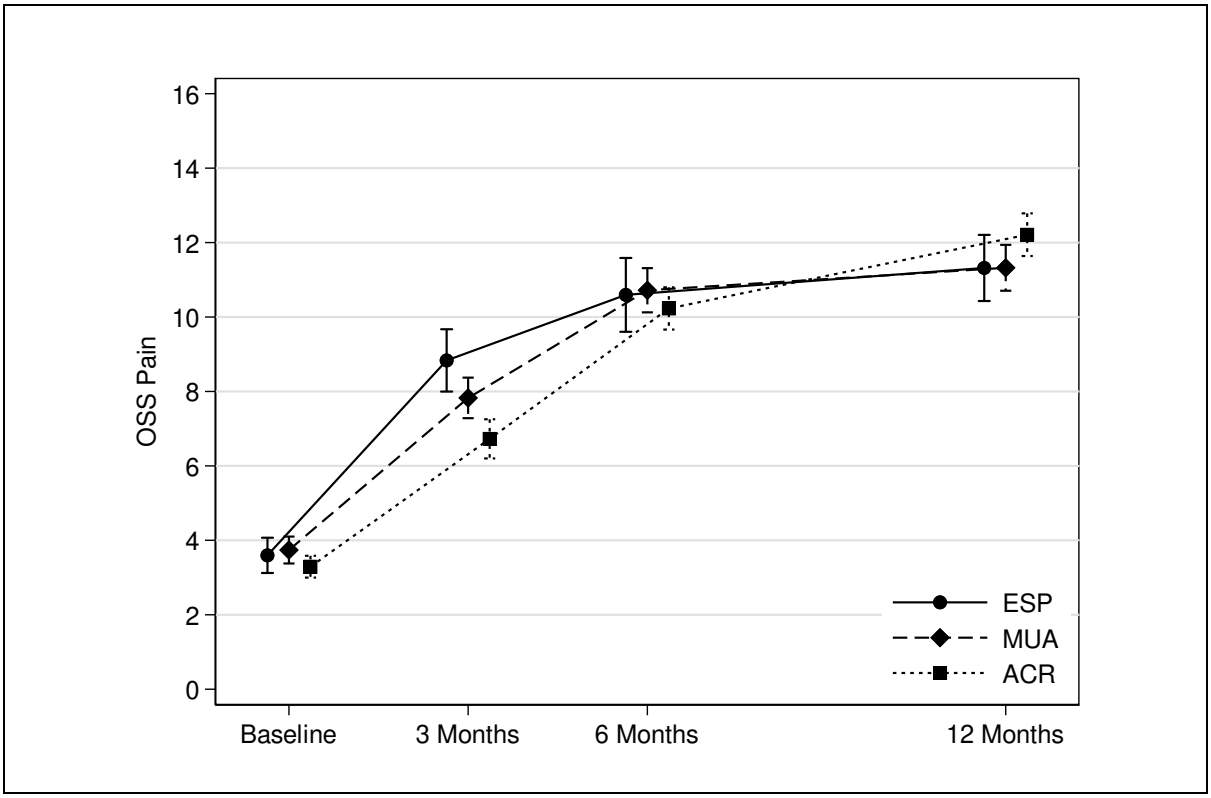


Figure 16: Unadjusted Mean OSS Pain Items and 95% CIs by Treatment Arm

Table 42: Unadjusted OSS Function Subdomain by Treatment Arm

	MUA	ACR	ESP	Total
Baseline				
N	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
Min, Max	2, 32	1, 29	2, 31	1, 32
3 Months				
N	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
Min, Max	4, 32	2, 32	1, 32	1, 32
6 Months				
N	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
Min, Max	2, 32	4, 32	4, 32	2, 32
12 Months				
N	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
Min, Max	3, 32	1, 32	3, 32	1, 32

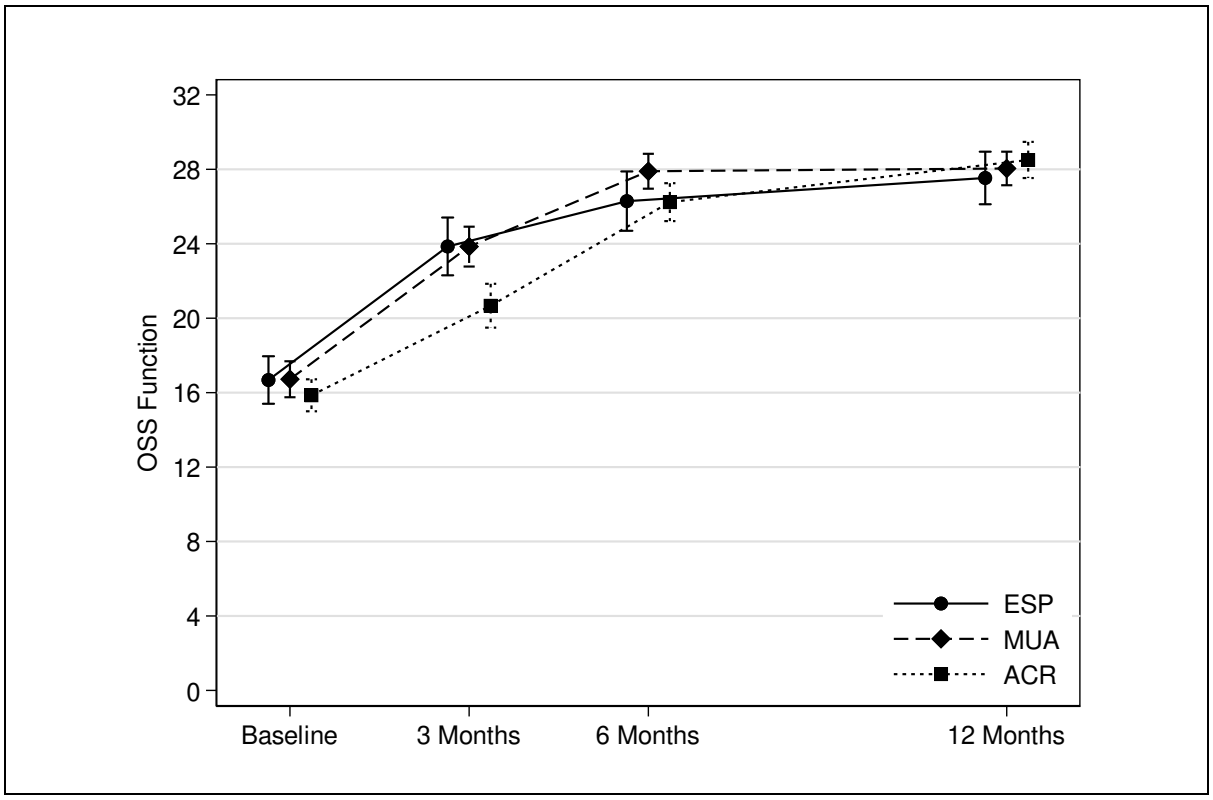


Figure 17: Unadjusted Mean OSS Function Items and 95% CIs by Treatment Arm

Appendix 7 Illustration of Treatment Effects

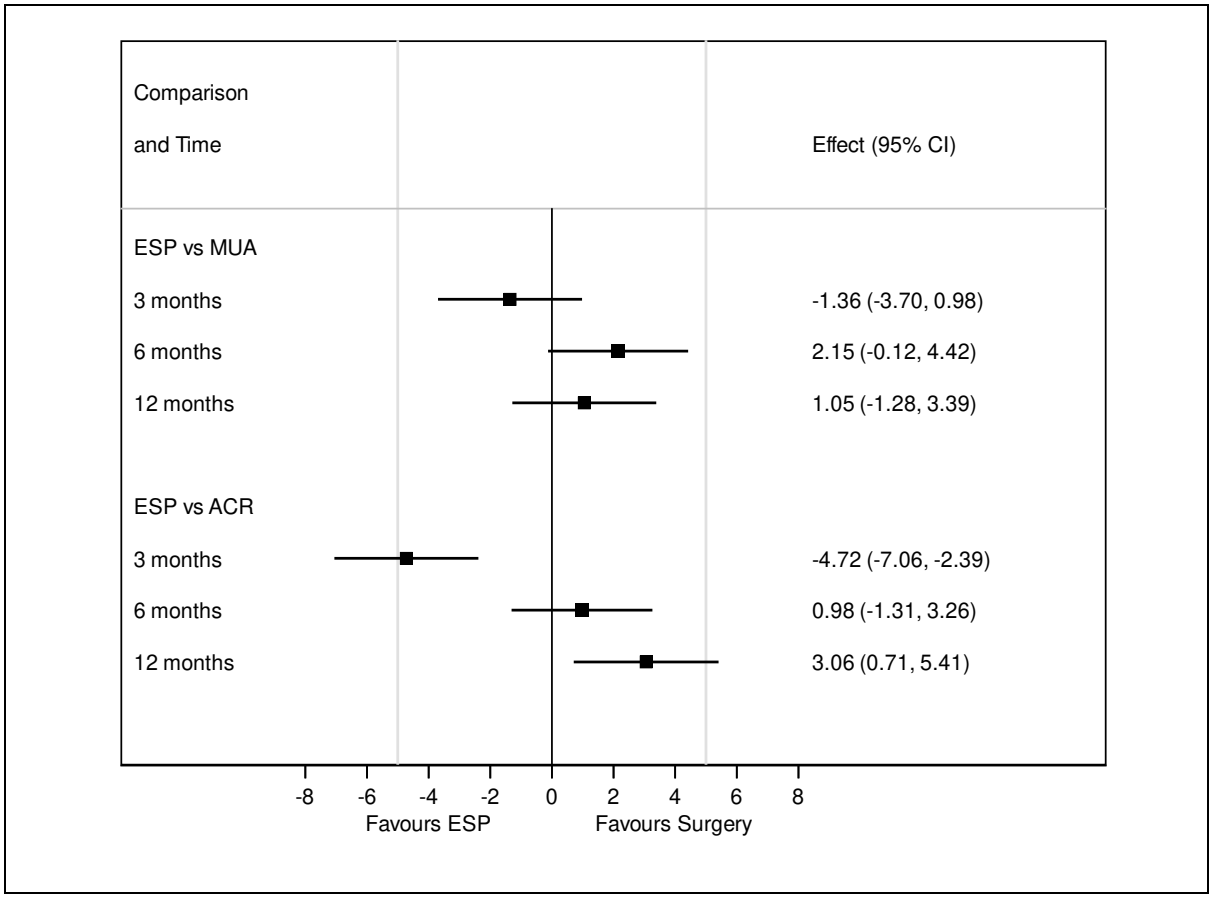


Figure 18: Estimated Mean OSS Differences from Primary Analysis Model by Treatment Arm (ESP vs Surgery) – grey lines indicate sought minimal clinically important difference

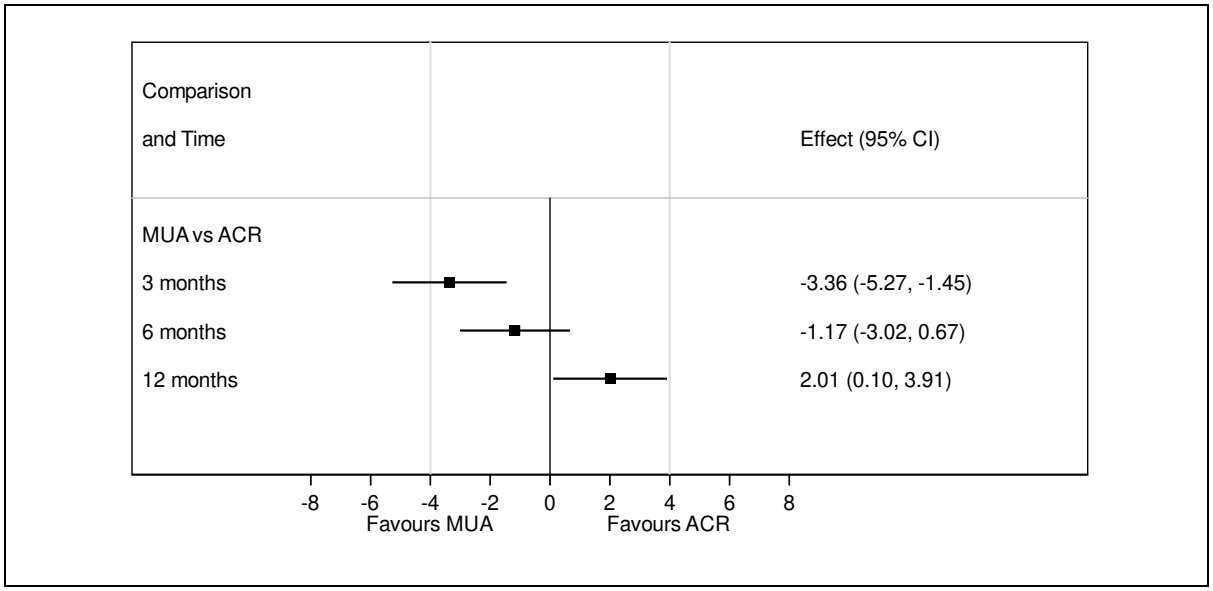


Figure 19: Estimated Mean OSS Differences from Primary Analysis Model by Treatment Arm (MUA vs ACR) – grey lines indicate sought minimal clinically important difference

Appendix 8 OSS 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						
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 (min, max)	20 (2, 48)	20 (3, 36)	19 (1, 37)	19 (4, 35)	20 (2, 42)	18 (4, 34)
12 Months						
N	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 (min, max)	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 six weeks beyond their intended follow-up

As more than 5% of responses were received beyond their 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, 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.

Table 44: Estimated Mean OSS Differences by Treatment Arm (Estimates from analysis excluding data received after 6 weeks)

	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

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), employment status was included as an additional covariate in the analysis model as a sensitivity analysis. Results were similar to those observed in the primary analysis (Table 45).

Table 45: Estimated Mean OSS Differences by Treatment Arm (Estimates from analysis adjusted for employment status)

	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

	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)	p-value
	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 Sub-group descriptive statistics (OSS)

Table 46: Unadjusted OSS by Treatment Arm and Diabetes Status

	Diabetic (n= 150, 30%)			Non-diabetic (n=353, 70%)		
	MUA	**ACR**	**ESP**	**MUA**	**ACR**	**ESP**
Baseline						
N	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
Min, Max	2, 40	4, 37	2, 42	15, 48	6, 35	9, 30
3 Months						
N	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
Min, Max	7, 48	5, 48	7, 48	20, 46	9, 43	14, 48
6 Months						
N	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
Min, Max	5, 48	9, 48	6, 48	15, 48	7, 46	10, 48
12 Months						
N	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
Min, Max	7, 48	8, 48	10, 48	27, 48	3, 48	20, 48

Table 47: Unadjusted OSS by Treatment Arm and Previous Physiotherapy

	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						
N	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
Min, Max	2, 40	2, 37	2, 42	2, 48	1, 37	6, 39
3 Months						
N	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
Min, Max	7, 48	2, 47	4, 46	5, 48	8, 48	8, 48
6 Months						
N	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
Min, Max	5, 48	7, 48	6, 48	3, 48	10, 48	7, 48
12 Months						
N	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
Min, Max	4, 48	2, 48	4, 48	5, 48	6, 48	10, 48

Table 48: Unadjusted OSS by Treatment Arm and Patient Preference

	Randomised to preferred treatment (n=131, 26%)			Randomised to non-preferred treatment (n=105, 21%)			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
Min, Max	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
Min, Max	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
Min, Max	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
Min, Max	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 Length of Symptoms at Baseline

	Duration of symptoms: less than 9 months (n=249, 61%)			Duration of symptoms: less than 9 months (n=245, 49%)		
	MUA	**ACR**	**ESP**	**MUA**	**ACR**	**ESP**
Baseline						
N	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
Min, Max	2, 36	2, 34	2, 39	1, 48	1, 37	4, 42
3 Months						
N	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
Min, Max	8, 46	2, 48	6, 48	5, 48	3, 47	4, 48
6 Months						
N	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
Min, Max	3, 48	7, 48	12, 48	5, 48	10, 48	6, 48
12 Months						
N	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 (38.5 (8.00)
Median	43	45	43.5	43	44	43
Min, Max	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

Characteristic	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						
N	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
Min, Max	2, 48		2, 33		6, 39	4, 34
12 Months						
N	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
Min, Max	4, 48		2, 48		10, 48	10, 46

Appendix 11 Secondary Outcomes Descriptives

Table 51: Unadjusted QuickDASH by Treatment Arm

	MUA	ACR	ESP	Total
Baseline				
N	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
Min, Max	0, 100	14, 100	14, 98	0, 100
3 Months				
N	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
Min, Max	0, 91	0, 93	0, 96	0, 96
6 Months				
N	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
Min, Max	0, 98	0, 91	0, 91	0, 98
12 Months				
N	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
Min, Max	0, 98	0, 93	0, 89	0, 98

Table 52: Unadjusted Pain NRS by Treatment Arm

	MUA	ACR	ESP	Total
Baseline				
N	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
Min, Max	0, 10	0, 10	0, 10	0, 10
3 Months				
N	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
Min, Max	0, 10	0, 10	0, 10	0, 10
6 Months				
N	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
Min, Max	0, 10	0, 9	0, 9	0, 10
12 Months				
N	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
Min, Max	0, 9	0, 9	0, 10	0, 10

Table 53: Unadjusted Extent of Recovery by Treatment Arm

	MUA	ACR	ESP	Total
Baseline				
N	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
Min, Max	0, 100	0, 100	50, 100	0, 100
3 Months				
N	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
Min, Max	0, 100	0, 100	0, 100	0, 100
6 Months				
N	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
Min, Max	0, 100	0, 100	0, 100	0, 100
12 Months				
N	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
Min, Max	0, 100	0, 100	0, 100	0, 100

Table 54: Predominant Shoulder Problem (for patients who received their allocated treatment)

	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 & 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 & Stiffness equally	15 (10%)	29 (19%)	11 (14%)

[a] for patients who attended two or more physio sessions

Appendix 12 Adverse Events

Table 55: Serious adverse events (summary by randomised group)

Number of Events	MUA N=2	ACR N=8	ESP N=0
Type			
Prolonged hospitalisation	0	2	0
Required hospitalisation	0	0	0
Other medically important condition	1	1	0
Not reported through SAE form	1	5	0
Relationship to trial treatments			
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			
Expected	0	2	0
Unexpected	1	2	0
Not reported through 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 adverse events (summary by randomised group)

Number of Events	MUA N=15	ACR N=13	ESP N=5
Relationship to trial treatments			
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 through AE form	3	3	1
Expectedness			
Expected	9	7	3
Unexpected	3	3	1
Not reported through AE form	3	3	1
Severity			
Mild	3	6	2
Moderate	5	2	1
Severe	2	2	0
Missing / not reported through 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), keyhole surgery was the most popular treatment, followed by physiotherapy (see Table 57). Few patients gave MUA as their preferred treatment. While clinicians did not have a preferred treatment for nearly half of these patients (45%), the agreed treatment was often keyhole surgery in line with patient preferences (43%). Average strength of any treatment preference was high for MUA, Keyhole Surgery and Physiotherapy (mean of 9 out of 10), but lower for individuals who wanted surgery but did not mind what surgery might be performed (Table 58).

Table 57: Treatment preferences for non-consenting patients

	Patient Preference (Consent Status CRF) n=281	Detailed Patient Preference (Optional Preferences CRF) N=158	Clinician Advice (Consent Status CRF) n=271	Agreed Treatment (Consent Status CRF) n=270
Any Surgery	-	20 (13%)	-	-
Manipulation under anaesthetic	26 (9%)	11 (7%)	27 (10%)	32 (12%)
Keyhole surgery	116 (41%)	58 (37%)	91 (34%)	117 (43%)
Physiotherapy	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	Manipulation under anaesthetic	Keyhole Surgery	Physiotherapy
N	14	11	54	34
Mean (SD)	5.1 (4.41)	9.0 (1.10)	9.0 (1.69)	9.0 (1.34)
Median (min, max)	3 (1, 10)	9 (7, 10)	10 (1, 10)	10 (6, 10)

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

Table 59: Treatment Expectations

	Manipulation under anaesthetic	Keyhole Surgery	Physiotherapy
Non-consenting Patients			
Very ineffective	4 (3%)	2 (1%)	38 (26%)
Fairly ineffective	12 (8%)	4 (3%)	43 (29%)
Can't 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%)
Can't decide	231 (47%)	208 (43%)	252 (52%)
Fairly effective	121 (25%)	85 (17%)	73 (15%)
Very effective	90 (18%)	151 (31%)	35 (7%)

Baseline preferences among randomised patients are presented in Table 60. Approximately half of the participants had no treatment preference, and for those who did have a preference, this was predominantly for either surgery, or keyhole surgery 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 were allocated to (Table 61). Preference for keyhole surgery was much greater among patients who received their preferred treatment rather than if another treatment had initially been preferred (50% vs 31%), whereas the opposite was true for physiotherapy (13% vs 23%).

Table 60: Treatment Preferences for Randomised Patients at Baseline

	N	Baseline Preference				
		MUA	Keyhole Surgery	Either Surgery	Physiotherapy	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 prior physio	306	25 (8%)	48 (16%)	57 (19%)	20 (7%)	156 (51%)
Did not have prior physio	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 months follow-up				
		MUA	Keyhole Surgery	Either Surgery	Physiotherapy	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%)
Keyhole Surgery	63	16 (25%)	31 (49%)	5 (8%)	7 (11%)	4 (6%)
Either Surgery	70	17 (24%)	33 (47%)	6 (9%)	9 (13%)	5 (7%)
Physiotherapy	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 preference[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 with no baseline preference

Appendix 14 OSS Change Scores

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

Table 62: Anchoring of OSS Change Scores

Difference in OSS Score between Baseline and 12 months	How is your shoulder compared to a year ago?				
	Much better	Slightly better	About the same	Slightly worse	Much worse
N	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
Min, Max	-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		*Quick*DASH	
	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				
N	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
Min, Max	1, 48	3, 36	0, 100	23, 100
3 Months				
N	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
Min, Max	2, 48	5, 47	0, 96	0, 91
6 Months				
N	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
Min, Max	3, 48	6, 48	0, 98	0, 89
12 Months				
N	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
Min, Max	2, 48	7, 48	0, 98	0, 93

Appendix 16 Health related quality of life and QALYs

Table 64 HRQoL: number of questionnaires returned and completed EQ-5D scores

	Baseline n (%)		**Month 3 n (%)**		**Month 6 n (%)**		**Month 12 n (%)**	
Group	**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 HRQoL: Proportion reporting EQ-5D-5L levels 1 to 5 by dimension, group and time point (complete cases)

		Baseline						3 months						6 months						12 months					
		MUA		ACR		ESP		MUA		ACR		ESP		MUA		ACR		ESP		MUA		ACR		ESP	
		n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%
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	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 prob.		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	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 prob.		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
Usual Act.	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 prob.		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/D*	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 prob.		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 Deprs.**	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	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 prob.		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

		Missing values (%)						
		Total	ESP	MUA	ACR	Range	Mean	SD
BASELINE VARIABLES								
age	Age at trial entry	0	0	0	0	30 to 70	54.25	7.72
gender	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% No Dia.	
alloc	Treatment allocation	0	0	0	0	1,3		
OUTCOME VARIABLES FOR HEALTH RELATED QUALITY OF LIFE								
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 ^	0	0	0	0	59.8 to 768.4	279.46	148.8
Cost_MUA	Costs of MUA ^	0	0	0	0	259.2 to 972.0	424.81	115.5
Cost_ACR	Costs of ACR ^	0	0	0	0	877.3 to 3,082.3	2,170.46	431.1
Cost_PPP	Costs of physiotherapy ~	0	0	0	0	0 to 975.2	209.65	152.9
Cost_add	Additional treatments [a]	0	0	0	0	0 to 167.97	2.83	21.0
Cost_further	Further treatments [b]	0	0	0	0	0 to 1,521.87	41.41	204.2
Cost_other	Other treatments [c]	0	0	0	0	0 to 668	7.18	49.42
Cost_crossovers	Treat. after crossover [d]	0	0	0	0	0 to 125.01	0.50	7.87
Cost_Hosp_INP	Inp costs re complications[e]	0	0	0	0	0 to 4,926.24	32.85	312.1
Cost_Hosp_OUP	Out costs re complications[f]	0	0	0	0	0 to 875.07	19.37	82.71
Cost_GP_pr	Costs of GP visits (surgery)	33.0	37.4	31.8	32.0	0 to 822.8	57.26	110.6
Cost GP_phone	Costs of GP visits (phone)	34.2	38.3	32.3	34.0	0 to 197.6	6.33	23.01
Cost Nurse_pr	Costs of Practice Nurse	36.4	40.4	34.3	36.4	0 to 75.95	2.10	6.54
Cost_Nure_dis	Costs of District Nurse	33.8	37.4	32.8	33.0	0 to 380	1.94	21.69
Cost_Physio_c	Costs of District Physio	33.4	35.3	32.8	33.0	0 to 1,214.4	56.27	183.1
Cost_OT_c	Costs Occupational Therapist	16.9	16.2	16.4	17.7	0 to 282	0.67	13.79
OUTCOMES FOR COSTS 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	40.5	44.4	38.8	40.4	0 to 5,732.54	1,372.36	1,095.99

^For those who had ESP/surgery (MUA/ARCR).

~ Costs of Post Procedure Physiotherapy for those who had surgery (MUA/ARCR).

a Any treatments received before/during receiving randomised treatment.

b Any treatments received after completing randomised treatment.

c Any non-trial treatments the patient had if they did not start/complete their randomised treatment.

d Cost of further treatments following crossover.

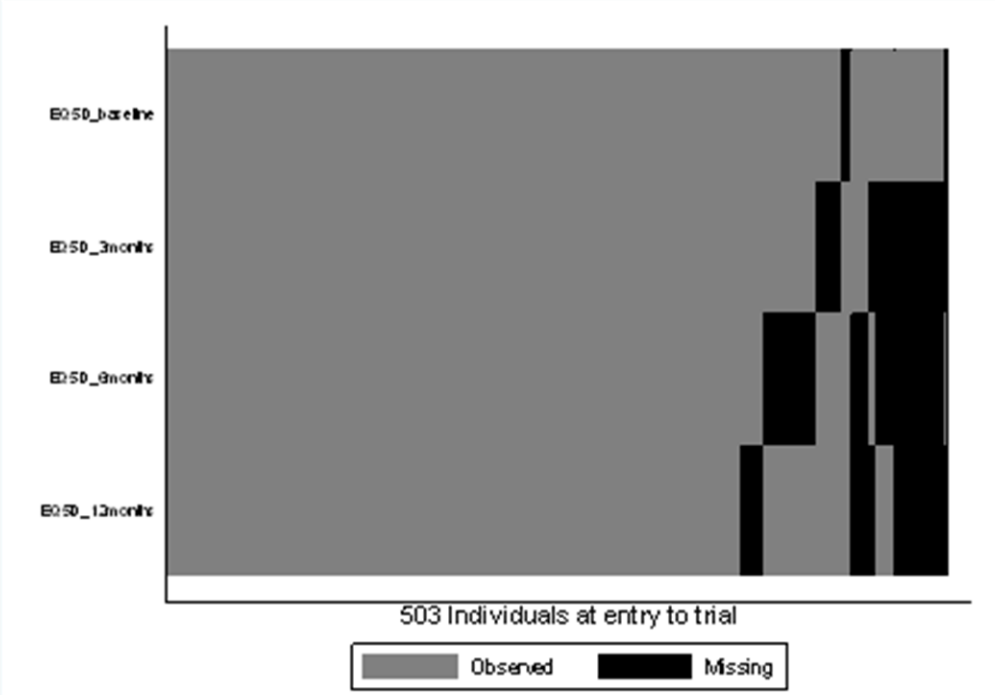
e Hospital inpatient stay costs related to complications.

f Outpatient hospital costs related to complications.

g Costs of Adverse event

Table 67 Number and proportions of patients with complete data by treatment arm

Complete at	ESP (N=99)	MUA (N=201)	ACR (N=203)
COMPLETE- HEALTH RELATED QUALITY OF LIFE			
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 HEALTH RELATED QUALITY OF LIFE 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%)



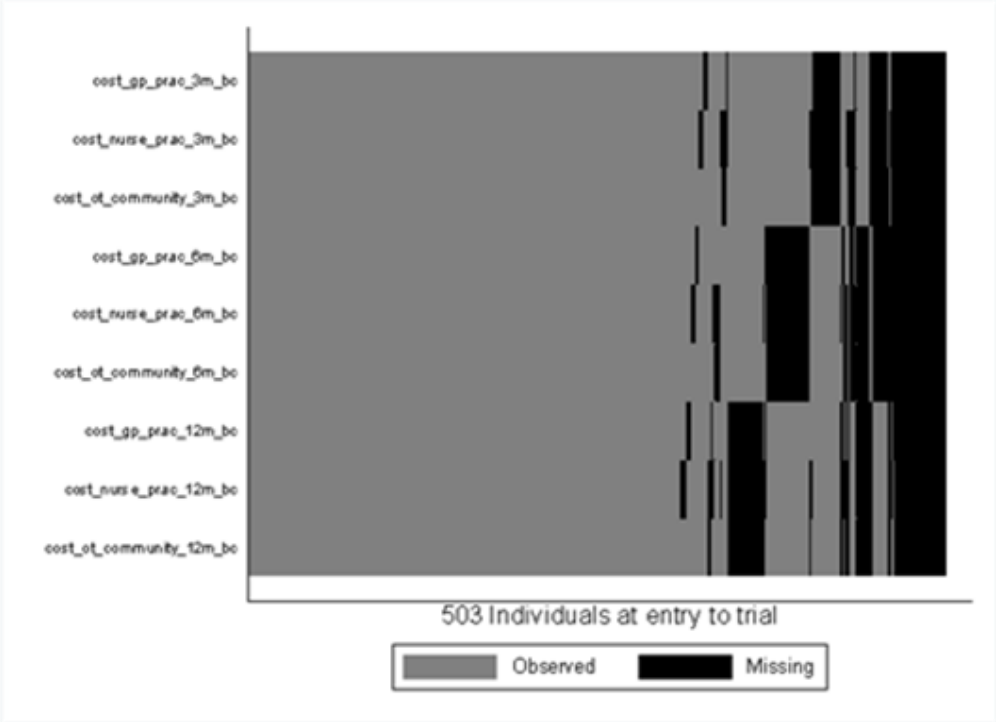


Figure 20: Pattern of missing data in UKFROST dataset

Table 68: Logistic regression for (i) missingness of costs and QALYs on baseline variables; and (ii) for 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 – 1.32)	0.60 (0.34 – 1.05)
Treatment allocation (ACR vs ESP)	0.85 (0.52 – 1.41)	0.71 (0.41 – 1.23)

Gender	1.26 (0.85 - 1.88)	0.87 (0.55 – 1.37)
Age	0.99 (0.97- 1.01)	0.95 (0.93 – 0.98)**
Diabetes	1.11 (0.89 – 1.38)	1.06 (0.82 – 1.35)
EQ-5D at baseline	0.28 (0.14 – 0.57)**	0.31 (0.14 – 0.67)**
QALYs at 3 months	0.003 (0.00 to 0.09)**	0.00 (0.00 to 0.50)**
QALYs at 6 months	0.007 (0.00 to 0.306)**	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)

** Statistically significant at 5% level

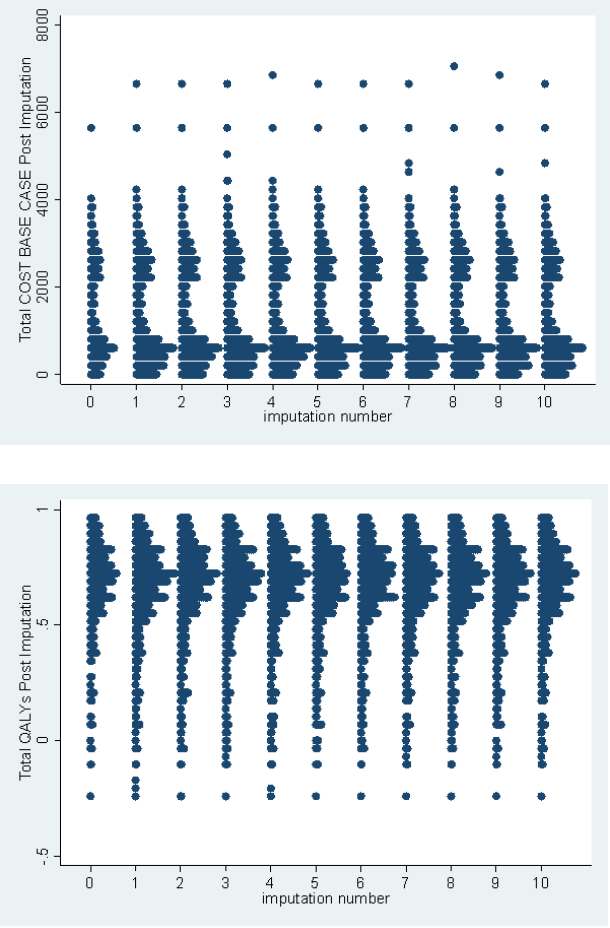


Figure 21: Comparison of the distribution of imputed values (imputation 1 to 10) with the observed data (imputation number 0) for costs and QALYs. Individual values are represented by dots; the width of a row of dots represents the frequency of values in the distribution

Table 69: Sensitivity analysis (Scenario 6): summary for incremental analysis (ITT), cost-effectiveness results and uncertainty of different methods to handle 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) 228.605 (0.94 to 456.27)	0.0396 (-0.0008 to 0.0800) 0.0339 (-0.0138 to 0.0816)	6,984 6,750	88% 81%
Same MNAR parameters in MUA and ESP				
M1: -10% QoL in both arms	228.605 (0.94 to 456.27)	0.0414 (-0.0041 to 0.0868)	5,227	89%
M2: +10% cost in both arms	234.7271 (-6.91 to 476.36)	0.0339 (-0.0138 to 0.0816)	6,935	80%
M3: -50% QoL in both arms	228.605 (0.94 to 456.27)	0.0713 (0.0221 to 0.1206)	3,204	99%
M4: +50% cost in both arms	259.2152 (-52.66 to 571.09)	0.0339 (-0.0138 to 0.0816)	7,665	78%
M5: -10% QoL and +10% costs in both arms	234.7271 (-6.91 to 476.36)	.0413277 (-0.004 to 0.087)	5,680	88%
M6: -50% QoL and +50% costs in both arms	259.2152 (-52.66 to 571.09)	0.0710225 (0.0217 to 0.1203)	3,650	98%
Different MNAR parameters in MUA and ESP				
M7: -10% QoL in ESP	228.605 (0.94 to 456.27)	0.0559849 (0.010 to 0.102)	4,083	96%
M8: -10% QoL in MUA	228.605 (0.94 to 456.27)	0.0192851 (-0.0281 to 0.0667)	11,854	62%
M9: +10% cost in ESP	199.748 (-32.80 to 432.29)	0.0338503 (-0.0139 to 0.0816)	5,901	82%
M10: +10% cost in MUA	261.540 (28.02 to 495.06)	0.0338673 (-0.0138 to 0.0816)	7,722	79%
M11: -50% QoL in ESP	228.605 (0.94 to 456.27)	0.144459 (0.101 to 0.188)	1,582	99%
M12: -50% QoL in MUA	228.605 (0.94 to 456.27)	-0.0390401 (-0.0895 to 0.0114)	-5,856	3%
M13: +50% cost in ESP	84.318 (-171.7 to 340.42)	0.0337907 (-0.0139 to 0.0815)	2,495	87%
M14: +50% cost in MUA	393.28 (130.9 to 655.60)	0.0338787 (-0.014 to 0.082)	11,608	71%