WARWICK
THE UNIVERSITY OF WARWICK

Sub-acromial spacer for Tears Affecting Rotator cuff Tendons: a Randomised, Efficient, Adaptive Clinical Trial in Surgery (START:REACTS).

STATISTICAL ANALYSIS PLAN

ISRCTN:	**ISRCTN17825590**
Funding Body:	**NIHR EME**
Ethics Approval:	**West Midlands – Coventry and Warwickshire Research Ethics Committee; 13 Feb 18**
SAP version number:	**2**
Date:	**21 May 2021**
Stage:	**Final**
Protocol version:	**v6.0 31 July 2020**

Author signature: _____ **date:** _____

Senior statistician signature: _____ **date:** _____

Chief Investigator signature: _____ **date:** _____

WARWICK
CLINICAL TRIALS UNIT

FUNDED BY
NIHR | National Institute for Health Research

Contents

1 Administrative information

Title: Sub-acromial spacer for Tears Affecting Rotator cuff Tendons: a Randomised, Efficient, Adaptive Clinical Trial in Surgery (START:REACTS).

Trial registration number: ISRCTN17825590

1.1 SAP amendments

SAP version	Protocol version	Section(s) changed	Details of SAP changes	Date of update
0.1-0.2	3	-	Working drafts	-
0.3	4	-	Working draft updated to reflect v4 of protocol	24/09/19
0.4	4	-	Working draft to circulate to TMG, inclusion of suggestions from Sept 19 DMC meeting. Change of junior statistician	31/10/19
1	4	6; 8	Addition of unbiased treatment estimated to planned analysis. Clarification of WORC scoring. Update to dummy tables. Updated NIHR logo.	20/02/20
1.1	5	2.3; 3.3; 6	Response to covid-19 epidemic: change of primary outcome and added sensitivity analyses	02/07/20
2		2.3, 3.3, 6.5, 7	Analysis finalised prior to the final primary outcome data collected. Edits to text for impact of covid-19 and clarity. References updated	21/05/21

1.2 Supporting documents

This Statistical Analysis Plan (SAP) should be read in conjunction with the study protocol and WCTU Standard Operating Procedures:

- SOP 8: Statistical Considerations
- SOP 9: Randomisation and Blinding
- SOP 15: Information Handling
- SOP 21: Statistical Analysis Plan

Details on the adaptive design element of the study are included in the START:REACTS adaptive charter.

The Trial Master File, including the Data management Plan can be found in the START Trial Manager's office.

1.3 Study oversight

As described in the protocol, the procedures in place for oversight of this study include both a Trial Steering Committee (TSC) and Data Monitoring Committee (DMC). The DMC is advisory to the TSC and write to the TSC and recommend any alterations to the study to ensure the safety of participants and the integrity of the data.

1.4 Key roles and responsibilities

Details of all other START:REACTS co-applicants can be found in the protocol.

Role	Name, address, telephone, email
Chief Investigator	Mr Andrew Metcalfe Warwick Clinical Trials Unit Tel: 02476 150925 Email: a.metcalfe@warwick.ac.uk
Senior Statistician (Trial statistician)	Dr Helen Parsons Senior Research Fellow, Warwick Clinical Trials Unit Tel: 02476 572665 Email: H.Parsons@warwick.ac.uk
Junior Statistician	Aminul Haque; Research Assistant Warwick Clinical Trials Unit Tel: 024 761 50404 Aminul.Haque.1@warwick.ac.uk
Methodological experts	Dr Nicholas Parsons (Adaptive methodology) Principal Research Fellow, Warwick Medical School Email: nick.parsons@warwick.ac.uk Professor Nigel Stallard (Adaptive methodology) Professor of Medical Statistics & Epidemiology, Warwick Medical School Email: N.Stallard@warwick.ac.uk
Administrative contact	START REACTS Trial Manager Warwick Clinical Trials Unit, Clinical Sciences Building, Clinical Sciences Research Laboratories, University Hospitals Coventry and Warwickshire, Clifford Bridge Road, Coventry, CV2 2DX Tel: 02476 968629 Email: start@warwick.ac.uk
Data Monitoring Committee	*Omitted for WRAP upload*

1.5 Signatures

Signatures may also be found on QPulse

Role	Name	Date	Signature
Author of SAP			
Senior statistician (if different to author)			
Chief Investigator			

2 Introduction and summary of protocol

2.1 Trial background and rationale

Shoulder pain is a common and disabling problem, with around 70-80% of problems due to rotator cuff disease. Patients with a symptomatic rotator cuff tear present with pain, restricted movement, loss of strength and disability, and the disease is associated with substantial expense to society through both costs of treatment and sick leave. Rotator cuff repair is a widely accepted treatment for symptomatic rotator cuff tears. However, some tears cannot be repaired ("irreparable tears") and the management of these patients can be very difficult. Arthroscopic debridement is commonly used and benefit has been demonstrated in case-series, but it remains a controversial option, with little or no benefit observed in randomised trials.

In 2013**,** the InSpace subacromial device (Stryker) was introduced into UK orthopaedic practice as a potential treatment option for people with irreparable tears of the rotator cuff. The InSpace device is a saline-filled, balloon made of biodegradable (dissolvable) synthetic material. It is inserted above the main joint of the shoulder at the end of an arthroscopic debridement after an irreparable tear has been identified. It is simple to deploy and adds less than 10 minutes to the operation. In May 2016, an interventional procedure guidance document was published by NICE, five years following its use in clinical practice, demonstrating very limited evidence for its use. Therefore, the device was limited to use in the context of research only and a research recommendation was made to assess its effectiveness.

More details on the background to the trial can be found in the protocol.

2.2 Interventions

A brief description of each trial intervention is provided below, full descriptions can be found in the protocol.

2.2.1 Control group: arthroscopic debridement alone

After being established as eligible at surgery (see section 5.2), the control group will receive arthroscopic debridement (surgery) only.

2.2.2 Intervention group: arthroscopic debridement with InSpace device

The intervention group will receive arthroscopic debridement as in the control group, with the additional insertion of the InSpace device during the surgery.

2.3 Trial aims and objectives

The overarching aim is to implement a novel, efficient adaptive clinical trial design for new surgical interventions. This design will be used to assess the clinical effectiveness and safety of a sub-acromial spacer device for patients with symptomatic irreparable tears of the rotator cuff.

2.3.1 Protocol changes due to the covid-19 pandemic in 2020

The trial was planned and powered using the Constant shoulder score (CS). However, the covid-19 pandemic, which reached the UK in March 2020, required the study management team to make a number of urgent changes to the study protocol. This ensured that the trial could continue to recruit participants and collect follow-up data safely as study population includes many who are categorised as extremely clinically vulnerable to covid-19. Chiefly amongst these changes was the decision to make the Oxford shoulder score (OSS) at twelve months after surgery the primary study outcome. The OSS can be completed remotely (e.g. by telephone), in contrast to the CS which requires in person contact with patients. The decision to change the primary outcome was made by the trial management group, and agreed by the independent data monitoring committee and the trial steering committee on the 25th of March and the 7th of July 2020 respectively.

2.3.2 Primary objective

To quantify and draw inferences on observed differences between arthroscopic debridement of the subacromial space and arthroscopic debridement with insertion of the InSpace device twelve months after surgery, using the Oxford Shoulder Score (OSS) score.

2.3.3 Secondary objectives

The following secondary objectives are covered in this SAP. Other secondary objectives are outlined in the protocol.

1) To quantify and draw inferences on observed differences between the allocation groups at each follow up time point for the following outcomes:
    i. The Oxford Shoulder Score (OSS) at all other time points other than at 12 months
    ii. The Constant score and the Western Ontario Rotator Cuff index (WORC)
    iii. Shoulder pain free movement and strength
    iv. EQ5D
    v. Patient global impression of change (PGIC)
2) To assess the proposed mechanism of action of the device when it is still inflated and to determine if the effect persists when it has deflated (MRI substudy).

3 Study methods

3.1 Trial design

This study is a UK multi-centre, adaptive randomised controlled trial of two parallel treatment arms with an allocation ratio of 1:1.

The trial uses a novel adaptive study design. Further information on the adaptive element of the trial and details on how this will be carried out is described in the adaptive charter.

3.2 Randomisation

Participants will be randomly allocated (1:1) to the two treatment groups via a central computer-based randomisation system provided by the Warwick Clinical Trials Unit (WCTU, independent of the study team). This will be performed by minimisation with a random factor, with a 70% weighting towards balance across the whole study, using site, gender, age group (<70 years and ≥70 years, based on age distribution of previous studies) and cuff tear size (≥3cm or <3cm, measured intra-operatively) as strata.

3.3 Sample size

Initially, the study was designed using the Constant score. For the Constant score at twelve months, a clinically important difference of 10 units was selected along with a standard deviation of 20 and a standardised mean difference of 0.5. A magnitude of 0.5 is expected for the correlations between the Constant scores at three, six and twelve months.

For the Oxford Shoulder Score, anchor-based studies have estimated the target difference as 6 and a standard deviation of 12 has been observed in multiple studies, [1] therefore, a moderate standardised mean difference of 0.5 remains appropriate. An assumed correlation of 0.5 between time-points remains appropriate based on data from previous studies in our unit in a similar population. [2]

Hence, for a study without early stopping, at 5% significance and 90% power would require 170 participants without loss any loss to follow up.

The adaptive design used for this study allows stopping for either efficacy (arthroscopic debridement with the InSpace device performs better than arthroscopic debridement alone) or futility (the device performs no better than arthroscopic debridement alone) at a number of pre-specified interim analyses (early looks). [3] The timing of the early looks will be determined by the information accrued during participant follow-up. The information is dependent on the variance of the 3, 6 and 12 month OSS, and the correlation between the scores, in addition to the numbers and pattern of data accrual (i.e. the relative and absolute numbers of 3, 6 and 12 month scores). Using the best available knowledge of the pattern and timings of data accrual, with estimates of the variances and correlations from the research literature, simulations were undertaken to assess the likely sample size required to detect a clinically important difference in 12 month OSS.

The simulations tested a range of options for the number of interim looks and the type I error rate spent; for examples of simulations see Parsons *et al*. [3] In order to preserve the integrity of the trial the details of the selected design will remain confidential and known to the relevant members of the methodology TMG and DMC only; full details are in the adaptive charter. For the selected design, simulations indicate that a sample size of 188 participants would be required for a power of 90% at the 5% significance level. A total sample size of **221** participants is then needed to allow for a loss to follow up rate of 15%.

3.4 Framework

The objectives in this trial SAP will be tested using a frequentist superiority hypothesis testing framework.

3.5 Blinding

The participants and assessors in the trial will be blinded to the allocation group. With the exception of the trial statisticians, blinding of the TMG members will be maintained where possible

3.6 Interim analyses and stopping guidance

Sequential stopping boundaries will be constructed that allow stopping for futility or stopping to reject the null hypothesis (efficacy), with interim analyses predefined and agreed with the TSC and DMC. Further information on the stopping rules and how they will be implemented is given in the adaptive charter.

The DMC and TSC retain their rights to stop the study for other reasons at any time.

3.7 Timing of final analysis

The final analysis will be carried out once the last 12 month outcome data have been collected.

3.8 Timing of outcome assessments

Visit		**1**	**2**	**Sub-study 1**	**3**	**4 and Sub-study 2**	**5**	**6**
Visit Window (No. weeks ± weeks (w) /months (m))	**Screening**	**Baseline**	**Surgery**	**8 weeks (-2 w/ +4 w) after V2**	**3m (-2 w/+6w) after V2**	**≥6 m (±6 w) After V2**	**12 m (± 3m) After V2**	**24m (±3m) After V2**
Check eligibility and provide PIS	✓							
Check inclusion and exclusion criteria		✓	✓					
Consent (and sub-study consent)		✓						
Baseline assessments		✓						
Randomisation			✓					
Intervention			✓					
Constant Score		✓			✓	✓	✓*	
PROMs		✓			✓	✓	✓	✓
Resource use					✓	✓	✓	✓
Adverse Events					✓	✓	✓	✓
Sub-study MRI				✓		✓		
End of trial								✓

*primary outcome time point

4 Statistical principles

4.1 Confidence intervals and P values

All data will be analysed and reported in accordance with the CONSORT statement and its extension for adaptive designs. [4, 5] Treatment effects will be presented, with appropriate 95% confidence intervals, for both the unadjusted and adjusted analyses. Tests will be two-sided and considered to provide evidence for a significant difference if p-values are less than 0.05 (5% significance level) unless otherwise stated.

4.2 Descriptive statistics

Standard descriptive summaries will be provided for the primary outcome measure (OSS) and all secondary outcome measures. Baseline data will be summarised to check comparability between treatment arms.

4.3 Adherence and protocol deviations

The delivery of the trial interventions will be recorded on a surgical case report form (CRF) which will be used to determine if the participant adhered to the treatment for analysis purposes.

Protocol deviations will be categorised into groups as appropriate after consultation with the Trial Management Group (TMG).

4.4 Analysis populations

All analyses will be conducted as intention to treat unless otherwise specified (e.g. any per protocol analyses as sensitivity analyses).

5 Trial populations

5.1 Screening data

Screening data will be checked to highlight the proportion of patients approached who agreed to participate in the study and observe the reasons why potential participants were not included participation in the trial.

5.2 Eligibility

5.2.1 Inclusion criteria

1. Rotator cuff tear deemed by the treating clinician to be technically irreparable (to be confirmed intra-operatively)
2. Intrusive symptoms (pain and loss of function) which in the opinion of the treating clinician warrants surgery.
3. Non-operative management has been unsuccessful.

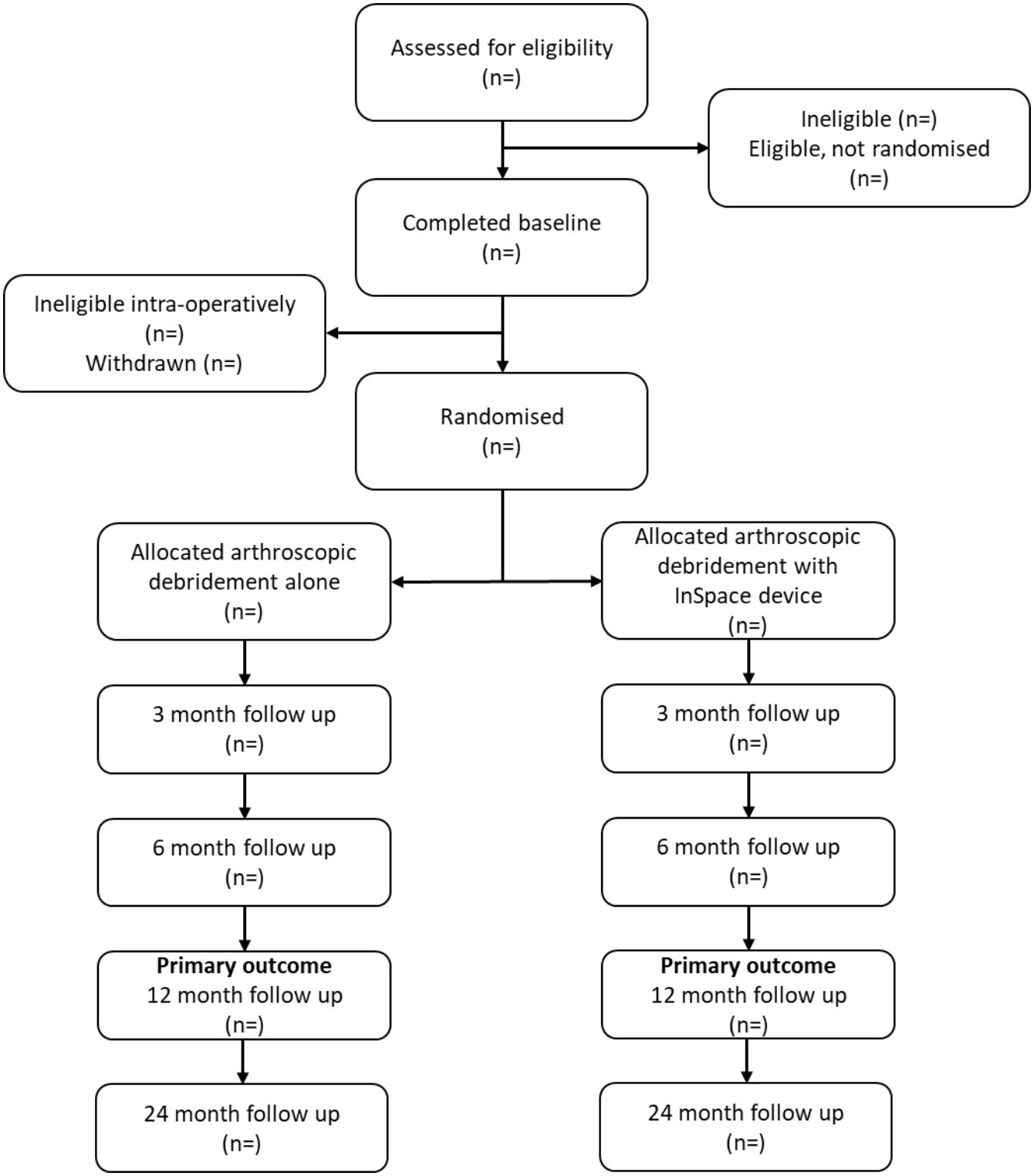
5.2.2 Exclusion criteria

4. Advanced gleno-humeral osteoarthritis on pre-operative imaging.
5. Subscapularis deficiency, defined as a tear involving more than the superior 1cm (approximately) of the subscapularis if repaired, or any tear that is not repaired.
6. The treating clinician determines that interposition grafting or tendon transfers are indicated.
7. Pseudoparalysis as determined by the treating clinician.
8. Unrelated, symptomatic ipsilateral shoulder disorder that would interfere with strength measurement or ability to perform rehabilitation

9. Other neurological or muscular condition that would interfere with strength measurement or ability to perform rehabilitation.
10. Previous proximal humerus fracture that could influence shoulder function.
11. Previous entry into the present trial (i.e. other shoulder).
12. Unable to complete trial procedures.
13. Age under 18
14. Unable to consent to the trial.
15. Unfit for surgery.

5.3 Recruitment

Figure 1: CONSORT chart



5.4 Withdrawal and loss to follow up

The following levels of withdrawal are possible in the study:

- Withdrawal from registration: these participants are not formally considered part of the study as they have withdrawn prior to being allocated a treatment
- Partial withdrawal from follow up: the participant withdraws from in-person follow up. However, the participant will be retained in the study and asked to complete postal questionnaires containing secondary outcomes
- Total withdrawal from the study: the participant withdraws from all follow up (in person and via post)

Due to the nature of the study design, withdrawal from the intervention is not possible. Data from withdrawn participants will be retained and used in analyses unless otherwise requested by the participant.

5.5 Baseline patient characteristics

Baseline data including age, sex, BMI, site and cuff tear size will be summarised to check comparability between treatment arms. Further details can be found in the Table 6, Table 7 and Table 8.

6 Analysis

6.1 Patient reported outcome definitions

The patient reported outcomes used in the trial are defined below:

Oxford Shoulder Score (OSS): A patient reported outcome measure consisting of 12 items designed to assess the outcomes of shoulder surgery. The OSS is measured on a scale of 0-48 where 48 is the best possible outcome. The OSS will be assessed as the primary clinical outcome at 12 months, along with three and six month data that will also be collected. The score will be calculated as described by the OSS scoring system. [6]

Constant-Murley score (Constant score): An outcome measure designed to indicate the overall functionality of patients with a shoulder disorder, measured on a scale of 0-100 with higher numbers representing better functionality. The Constant score consists of four parts, pain, activities of daily living, movement and strength and will be calculated as described in the Ban *et al* paper [7]. This outcome will be collected as a secondary outcome.

Western Ontario Rotator Cuff index (WORC): A patient reported quality of life outcome measure designed for patients with rotator cuff disease and consisting of 21 questions. Each question in the WORC is answered in the format of an 11 point numerical rating scale (NRS), which is converted into a total percentage score, with higher percentages indicating worse symptoms. Subscales are described in the WORC user manual. [8]

EQ-5D: Is a validated, generic health-related quality of life measure consisting of five items each with five possible responses. These are then converted into a health utility score using the UK value set as recommended by the Health Economic Team (e.g. [9]). Both the statistical and health economic analyses will use the same utility values to ensure compatibility.

6.2 Analysis of efficacy

Standard statistical summaries (e.g. medians and ranges or means and variances, dependent on the distribution of the outcome) and graphical plots showing correlations will be presented for the primary outcome measure and all secondary outcome measures. Outcome measure data, such as the OSS and Constant scores, will be assumed to be normally distributed during modelling, but subsidiary analyses may also be undertaken after appropriate variance-stabilising transformation if assumptions of normality prove to be unsustainable.

6.2.1 Primary clinical analysis

The primary analysis will follow the methods and use the test statistic described by Parsons *et al*. [3] The test statistic at the end of the study (after follow-up is complete) is given by the expression $B/sd(B)$, where B is a measure of the treatment effect of the device arm versus the arthroscopic debridement alone arm; expressions for estimating B and $sd(B)$ (the standard deviation of the effect estimate) are given by equations (1) and (2) of Parsons et al. [3] . Unbiased estimates of the treatment effect (*B*) will be determined using the methods of Liu and Hall [10], with confidence intervals estimated using the approach of Todd, Whitehead and Facey [11]. These unbiased estimates will be reported as the primary analysis and will be used for inferences on clinical significance.

If the study recruits to target, without stopping at an interim analysis, then testing at the final analysis (at the end of 12 month follow-up) will use the pre-specified adjusted boundaries from the adaptive charter. If the test statistic is greater than the upper boundary, then the null hypothesis will be rejected at the 5% level. If the study is stopped at an interim analysis, then the final 'overrunning' analysis will use data from all the participants recruited into the study prior to stopping. In this setting, testing will proceed using boundaries calculated by the deletion method of Whitehead [12], with inferences as per the unstopped analysis (i.e. rejection of the null hypothesis at the 5% level if the test statistic is greater than the upper boundary). If recruitment proves to be problematic, or for some unforeseen reason one or more than one (or all) interim analyses do not take place, then testing will proceed in a similar manner using boundaries adjusted to reflect the changes to the design.

In addition to reporting the statistical significance of the tests, treatment effect estimates (arthroscopic debridement with InSpace device versus arthroscopic debridement alone) from the primary analysis will be presented with 95% confidence intervals, and a statement of the clinical significance of reported differences.

6.2.2 Secondary clinical analysis

Summary plots will be used to show both the distributions (e.g. box-and-whisker plots) and temporal trends in intervention arm means for all outcomes (Constant, OSS, EQ-5D and WORC) at 3, 6, 12 and 24 months. Treatment effect estimates (arthroscopic debridement with InSpace device versus arthroscopic debridement alone) will be presented with 95% confidence intervals.

The results of the primary clinical analysis will be augmented with the results of fitting a mixed-effects model to the primary OSS and secondary outcomes (Constant score , EQ-5D and WORC) at 12 months, adjusting for the fixed-effects of baseline score, age group (70 years and ≥70 years), gender and cuff tear size (≥3cm or <3cm), with the inclusion of a random recruiting centre effect. Variables found to be imbalanced at baseline may also be included. Since individual clinicians will treat only a small number of patients enrolled in the trial, we do not expect clinician specific effects to be

important in this study and hence these will not be modelled. Inferences, regarding the effects of the arthroscopic debridement with InSpace device versus the arthroscopic debridement alone, will be made using the conventional (fixed design) boundaries, with significance assessed at the 5% level.

If the study is stopped early at an interim analysis (or otherwise), collinearity between key variables may be observed. For example, sites with a small number of randomisations may not have sufficient participants to use all strata groups. If this occurs, variables may be omitted or transformed where possible (e.g. replace age group with age at baseline). If study site is the variable causing the difficulties, a fixed effect model may be used instead. Similarly, if any pre-planned model variables (e.g. baseline OSS) are poorly reported such that it reduces the available sample size for the analysis, then they will not be included in the definitive model.

Complications will be summarised with between groups comparisons evaluated using chi-squared or Fisher's exact tests. If a large number of participants are observed not to confirm with the protocol, a per-protocol sensitivity analysis will also be constructed. If more extensive analysis of complications is deemed appropriate (i.e. they are sufficiently common to make more complex analysis useful), then mixed-effects logistic regression models (analogous to those described above) will also be fitted.

Other outcomes will be summarised and compared between groups using appropriate tests for the outcome (e.g. proportions and chi-squared tests for binary outcomes, means and t-tests for continuous data).

6.3 Missing data

It seems likely that some data may not be available due to voluntary withdrawal of patients, lack of completion of individual data items or general loss to follow-up. Where possible the reasons for data 'missingness' will be ascertained and reported. Reasons for ineligibility, non-compliance, withdrawal or other protocol violations will be stated and any patterns summarised.

The nature and pattern of the missingness will be carefully considered, including whether data can be treated as missing completely at random (MCAR). Variables that will be checked for their impact on missingness rates will include: site, gender, age group and cuff tear size. If judged appropriate, missing data will be imputed using multiple imputation. Any imputed analyses will be considered as secondary analyses and will be reported along with the primary analysis.

If imputation is undertaken, the resulting imputed datasets will be reported, together with appropriate sensitivity analyses. Any imputation methods used for scores and other derived variables will be carefully considered and justified. In particular, the model used for the multiple imputation will be assessed along with the plausibility of any imputed values.

6.4 Additional Analyses

6.4.1 Sensitivity analyses

Due to the nature of the study, it is not anticipated that there will be many participant cross-overs between allocation groups and/or protocol violations. However, if a sizable number of cross-overs are observed, a per-protocol analysis will be conducted as a sensitivity analysis. This analysis will follow the method set out above (section 6.2.1), but grouped as treatment received rather than as treatment allocated.

6.4.2 Subgroups

Pre-specified sub-group analyses will be undertaken to assess whether there is evidence that the intervention effect differs between:

- The size of the rotator cuff tear as measured at the start of surgery, defined as large or massive cuff tear (≥3cm) or moderate to small (<3cm).
- Gender
- Age (>70 or <70)

The subgroup analyses will follow the methods described for the secondary clinical analysis, with additional interaction terms incorporated into the mixed-effects regression model to assess the level of support for these hypotheses.

The study is not powered to formally test these hypotheses, so they will be reported as analyses only, and as subsidiary to the analysis reporting the main effects of the intervention in the full study population.

6.4.3 Repairable cuff tears

Due to the larger than anticipated number of potential participants excluded intra-operatively due to their cuff tears being found to be repairable, an exploratory analysis comparing the baseline outcome scores of the reparable and irreparable (randomised) patients will be conducted.

Descriptive statistics of the two tear types will be conducted of the baseline information. Direct comparisons between tear types will be assessed using a t-test for the baseline OSS. If deemed informative, as this population is not randomised, regression analyses to adjust for any imbalance between patient age group (<70 years and ≥70) and patient gender will also be carried out.

6.4.4 Effect of covid-19

In March 2020, the UK went into “lockdown” as a response to the covid-19 epidemic. The NHS reduced or stopped routine appointments and elective surgery. As the use of the device is thought to help patients engage with rehabilitation more quickly and successfully, the cancellation or reduction of post-surgery physiotherapy may impact on the recovery of participants during this time.

Hence, outcome data collected before and after the lockdown started will be compared. The number of missing primary outcome scores will be reported along with the number of patient reported physiotherapy contacts. Recovery trajectories will be presented graphically to help aid interpretation.

6.5 MRI sub-study

The aim of the sub-study is to assess the mechanism of action of the InSpace device. Measurements will be taken at two time points: an “early” time point when the devices are likely to be still inflated approximately 8 weeks post-randomisation (when acute post-operative pain has subsided); and a “later” time point, when the devices are likely to have fully deflated at least six months post-randomisation to see if the proposed mechanism for ongoing improvement is maintained.

6.5.1 Sub-study sample size

Based on Gumina's study [13], the minimum acromio-humeral distance (AHD) has a standard deviation of 1.72mm, so to observe a minimum important difference of 1.5mm (above the minimum detectable change of 1.3mm) 80% power at the 5% level, **44 participants** (two groups of 22 participants) are required. Assuming a conservative loss to follow-up rate at six months of 20%, **56 participants** are required for this sub-study.

6.5.2 Sub-study outcomes

The primary outcome will be the minimum acromio-humeral distance (AHD, as defined by Gumina) on the 'deltoid-active' coronal sequences at six months.

Secondary measures will be AHD at the first MRI; AHD on passive and sagittal images, and the change in AHD between active and passive images. The position of the device will be assessed on both sequences (with particular focus on the sagittal images) to check for migration and consistency of placement relative to the acromion.

6.5.3 Sub-study analyses

The primary end-point will be the between group difference on the later MRI, as that is the better indicator of long-term function and will determine whether the early effect of the device is likely to be maintained. The between-group differences on the earlier scan will be a secondary outcome. The primary analysis for this sub-study will use a similar method to the secondary clinical analysis of the main study (see section 6.2.1), and compare the 'deltoid active' AHD on coronal images between the debridement with InSpace device versus debridement alone groups using a t-test, and if possible, a generalised linear regression model adjusting for age, sex, tear size, and recruitment site.

6.5.4 Effects of covid-19 on the sub-study

As well as potentially having an impacting on participant recovery, the effects of the covid-19 pandemic have limited recruitment and data collection for the sub-study. In particular, many MRI scans scheduled six months after randomisation have been delayed or cancelled due to limited staff capacity or participant safety considerations. Hence, the window to collect MRI scans to investigate longer term effects of the InSpace device has been extended. It is not anticipated that sufficient data will be collected to formally investigate temporal effects, however, plots of outcomes over time will be created to aid interpretation.

6.6 Harms

Safety monitoring will be conducted primarily through participant self-reporting. At each follow-up occasion, participants will be asked if they have had any adverse events and how these were managed. The relatedness, expectedness and severity of any serious adverse events (SAEs) will be summarised as displayed in Table 12 of the dummy tables.

Complications deemed serious will be reported separately as SAEs. The number and nature will be reported and assessed by intervention arm, as shown in Table 12 of the dummy tables.

6.7 Statistical software

The routine statistical analysis will mainly be carried out using R [14] or Stata (e.g. StataCorp. 2019. *Stata Statistical Software:* College Station, TX: StataCorp LP).

7 References

[1] D. Christiansen, P. Frost, D. Falla, J. Haahr, L. Frich and S. Svendsen, "Responsiveness and Minimal Clinically Important Change: A Comparison Between 2 Shoulder Outcome Measures.," *J Orthop Sports Phys Ther,* vol. 45, no. 8, pp. 620-5, 2015.

[2] S. Karthikeyan, H. Kwong, P. Upadhyay, N. Parsons, S. Drew and D. Griffin, "A double-blind randomised controlled study comparing subacromial injection of tenoxicam or methylprednisolone in patients with subacromial impingement," *J Bone Joint Surg Br,* vol. 92, no. 1, pp. 77-82, 2010.

[3] N. Parsons, N. Stallard, H. Parsons, P. Wells, M. Underwood, J. Mason and A. Metcalfe, "An adaptive two-arm clinical trial using early endpoints to inform decision making: design for a study of sub-acromial spacers for repair of rotator cuff tendon tear," *Trials,* vol. 20, no. 694 , 2019.

[4] K. F. Schulz, D. G. Altman and D. Moher, "CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials," *BMJ,* p. 340:c332, 2010.

[5] M. Dimairo , E. Coates, P. Pallmann, S. Todd, S. A. Julious, T. Jaki, J. Wason, A. P. Mander, C. J. Weir, F. Koenig, M. Walton, K. Biggs, J. Nicholl, T. Hamasaki, M. Proschan, J. Scott, Y. Ando, D. Hind and D. Altman, "Development process of a consensus driven CONSORT extension for randomised trials using an adaptive design," *BMC Medicine,* vol. 16, p. 210, 2018.

[6] Isis Innovation Limited, *Oxford Shoulder Score (OSS),* 1998.

[7] I. Ban, A. Troelsen, D. Høyrup Christiansen, S. Wulff Svendsen and M. Tange Kristensen, "Standardised test protocol (Constant Score)," *Danish medical journal,* vol. 60, no. 4, pp. 1-8, 2013.

[8] A. Kirkley, S. Griffin and C. Alvarez, *Western Ontario Rotator Cuff Index (WORC),* 1998.

[9] B. van Hout, M. Janssen , Y. Feng , T. Kohlmann, J. Busschbach, D. Golicki, A. Lloyd, L. Scalone, P. Kind and A. Pickard, "Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets," *Value Health,* vol. 15, no. 5, pp. 708-15, 2012.

[10] A. Lui and W. J. Hall, "Unbiased estimation following a group sequential test.," *Biometrika,* vol. 86, no. 1, pp. 71-78, 1999.

[11] S. Todd, J. Whitehead and K. M. Facey, "Point and interval estimation following a sequential clinical trial.," *Biometrika,* vol. 83, no. 2, pp. 453-461, 1996.

[12] J. Whitehead, "Overrunning and underrunning in sequential clinical trials.," *Controlled Clinical Trials,* vol. 13, pp. 106-121, 1992.

[13] S. Gumina, V. Arceri, C. Fagnani, T. Venditto, C. Catalano, V. Candela and L. Nisticò, "Subacromial Space Width: Does Overuse or Genetics Play a Greater Role in Determining It? An MRI Study on Elderly Twins.," *J Bone Joint Surg Am.,* vol. 97, no. 20, pp. 1647-52, 2015.

[14] R Core Team, “R: A language and environment for statistical computing,” R Foundation for Statistical Computing, Vienna, Austria, [Online]. Available: https://www.R-project.org/.

[15] C. Gamble, A. Krishan, D. Stocken and et al., “Guidelines for the Content of Statistical Analysis Plans in Clinical Trials,” vol. 318, no. 23, pp. 2337-2343, 2017.

[16] I. R. White, P. Royston and A. M. Wood, Multiple imputation using chained equations:Issues and guidance for practice, Statistics in medicine, 2011.

[17] N. Devlin, K. Shah, Y. Feng, B. Mulhern and B. van Hout, “Valuing Health-Related Quality of Life: An EQ-5D-5L Value Set for England.,” *Office of Health Economics,* p. Research Paper 16/01., 2016.

[18] StataCorp, “Stata Statistical Software: Release 15,” StataCorp, College Station, TX, 2017.

8 Dummy tables

The following tables fill form the basis for the final statistical report. For brevity, some tables which will be reported at multiple time points are shown once; with variables noted when reported where necessary. Note also that variable level missingness is not reported here, but will be marked in the final report as appropriate.

Table 1: Participant flow from screening data

Reason not recruited		n
Ineligible	Advanced Gleno-humeral OA	
	Subscalpulus deficiency to be confirmed at surgery	
	Interposition grafting or tendon transfers	
	Pseudoparalysis	
	Unrelated symptomatic ipsilateral shoulder pathology	
	Unrelated neurological or muscular condition that would interfere with strength measurement or rehabilitation	
	Previous proximal humerus fracture	
	Previously entered trial	
	Unable to complete trial procedures	
	Age under 18	
	Unable to consent to trial	
	Unfit for surgery	
Eligible, patient unwilling	Prefers conservative care/Does not want treatment	
	Does not want to take part in research	
	Does not want standard operation`	
	Does not want InSpace device	
	Does not like randomisation	
	No reason given	
Other	Other reason given	

Table 2: Participant flow from consented to randomisation. Values reported are numbers and percentages of number consented unless stated otherwise

Reason not randomised		n consented
Excluded intra-operatively: Yes (n,%)		
Excluded intra-operatively: reason (n,%)	Subscapularis deficiency	
	Advanced gleno-humeral osteoarthritis	
	Repairable cuff tear	
	Other	
Withdrawn before surgery: Yes (n,%)		
Withdrawn before surgery: reason (n,%)	No longer receiving surgery	
	No longer wants to be part of the study	
	Other	
	No reason provided	

Table 3: Randomisation by site

Site name	Arthroscopy only (n)	Arthroscopy with device (n)	Randomised (n)
University Hospitals Coventry and Warwickshire			
North Bristol – Southmead			
North Tees			
Cambridge – Addenbrookes			
London North West			
Prince Philip			
Royal National Orthopaedic – Stanmore			
Royal Gwent Hospital			
Cardiff			
Royal Devon and Exeter			
Bournemouth			
Leicester			
Yeovil			
Guy's and St Thomas'			
Salisbury			
West Suffolk			
Southampton			
Doncaster			
Wrexham			
Kingston			
Peterborough			
…			
Total			

Table 4: Summary of baseline data for randomised participants and participants excluded intra-operative. Values reported are means and standard deviations unless otherwise stated

Baseline outcome / characteristic		Randomised (all participants)	Excluded Intra-operatively
Constant score			
Shoulder free range of motion			
Forward flexion			
Abduction			
Age (n,%)	<70 years		
	≥70 years		
Sex (n,%)	Male		
	Female		

Table 5: Summary of CRF completion (randomised participants)

Trial status	Status	Arthroscopy only (n)	Arthroscopy with device (n)	Total (n)
Baseline	Complete			
	Constant missing but data received			
	CRFs missing			
6 weeks	Complete			
	Constant missing but data received			
	Loss to follow up			
	Withdrawn			
3 months	Complete			
	Constant missing but data received			
	Lost to follow up			
	Withdrawn			
6 months	Complete			
	Constant missing but data received			
	Lost to follow up			
	Withdrawn			
12 months	Complete			
	Constant missing but data received			
	Lost to follow up			
	Withdrawn			
24 months	Complete			
	Constant missing but CRFs returned			
	Lost to follow up			
	Withdrawn			

Table 6: Baseline data. Values reported are means and standard deviations unless otherwise stated.

		Arthroscopy only (n=)	Arthroscopy with device (n=)	Total (n=)
Age (years)				
Age group 70 years or older (n,%)				
Sex: Male (n,%)				
Shoulder entered into study: Left (n,%)				
Study shoulder participants' dominant shoulder: Yes (n, %)				
Consented to MRI sub-study: Yes (n,%)				
Symptom duration (years/months)				
Left or right handed: Left (n,%)				
Current smoker: Yes (n,%)				
Diabetic: Yes (n,%)				
Diabetic - Type 1 or 2: Type 1 (n,%)				
Diabetic: Treatment (n,%)	Insulin			
	Medication			
	Diet			
Other medical conditions: Yes (n,%)				

Table 7: Summary statistics of shoulder history. Values reported are means and standard deviations unless otherwise stated.

	Arthroscopy only (n=)	Arthroscopy with device (n=)	Total (n=)
History of previous dislocation: Yes (n,%)			
History of dislocations: number of dislocations on same shoulder (n,%)			
History of dislocations: Time since last dislocation			
Unilateral or bilateral pain and symptoms: Unilateral (n,%)			
Previous physiotherapy: Yes (n,%)			
Steroid injection: Yes (n,%)			
Previous surgery: Yes (n,%)			
Other treatment: Yes (n,%)			

Table 8: Surgery details. Values reported are means and standard deviations unless otherwise stated.

		Arthroscopy only (n=)	Arthroscopy with device (n=)	Total (n=)
Antereo-Posterior size (cm)				
Medio-lateral retraction from GT attachment (cm)				
Biceps tendon intact: Yes (n,%)				
Subscapularis torn: Yes (n,%)				
Subscapularis torn	Size of tear (cm)			
	Repaired: Yes (n,%)			
Limited acromioplasty: Yes (n,%)				
CA ligament retained: Yes (n,%)				
Inspace device used: Yes (n, %)				
Size of device used (n, %)	Small			
	Medium			
	Large			
Device was stable: Yes (n, %)				
Had difficulties using InSpace device: Yes (n, %)				

Table 9: Patient reported outcome measures. Values reported are unadjusted means and standard deviations unless otherwise stated

Outcome	Time point	Arthroscopy only (n=)	Arthroscopy with device (n=)	Total (n=)
Constant-Murley score	Baseline			
	3 months			
	6 months			
	12 months			
	Baseline			

Shoulder free range of motion	3 months			
	6 months			
	12 months			
Forward flexion	Baseline			
	3 months			
	6 months			
	12 months			
Abduction	Baseline			
	3 months			
	6 months			
	12 months			
Oxford Shoulder Score (OSS)	Baseline			
	3 months			
	6 months			
	12 months*			
	24 months			
Western Ontario Rotator Cuff index (WORC)	Baseline			
	3 months			
	6 months			
	12 months			
	24 months			
EQ-5D	Baseline			
	3 months			
	6 months			
	12 months			
	24 months			
Global impression of change	3 months			
	6 months			
	12 months			
	24 months			

*Primary outcome

Table 10: MRI summary. Values reported are means and standard deviations unless otherwise stated

		Arthroscopy only (n=)	**Arthroscopy with device (n=)**	**Total (n=)**
Abduction distance (cm)	First MRI			
	Final MRI			
Pain or discomfort: Yes (n,%)	First MRI			
	Final MRI			
Scan completed as per-protocol: Yes (n,%)	First MRI			
	Final MRI			

Table 11: Summary of AEs and complications reported. Figures denote no. and % of group unless stated

AE		**Arthroscopy only (n=)**	**Arthroscopy with device (n=)**	**Total (n=)**
All reported AEs (n)				
AEs per participant	1			
	2			
	3			
	4			
	…			
Participant experienced any AE				
Additional surgery to study shoulder				
Additional injuries to study shoulder				
Injury to teeth mouth or throat during anaesthetic				
Chest infection				
Myocardial infection				
Nerve or vessel injury due to local anaesthetic				
Exacerbation/persistence of shoulder pain or restrictive range of motion				
Injection into the shoulder region				
Adhesive capsulitis				
Deep infection of the shoulder joint or implant				
Wound healing problems				
Thrombosis (DTV or PE)				
Damage to nerves or vessels in the surgical area				
Mis-placement of the device or its subsequent migration				
Device defect/failure				
Persistent muscle soreness or muscle injury				
Bruising				
Other				

Table 12: Unexpected, related Serious Adverse Events (SAEs)

SAE category (randomised participants)		**Arthroscopy only(n=)**	**Arthroscopy with device (n=)**	**Total (n=)**
All reported SAEs (n)				
SAEs per participant (n)	1			
	2			
	…			
Life-threatening (n,% of SAEs in group)				
Hospitalisation or prolongation of existing hospitalisation (n, % of SAEs in group)				
Persistent or significant disability/incapacity (n,% of SAEs in group)				
Other reason (n, % of SAEs in group)				
Relatedness to intervention* (n, % of SAEs in group)	Related			
	Unrelated			
Potential relatedness to InSpace device only* (n, % of related SAEs in group)	Related			
	Unrelated			
If related, was the SAE expected (n, % of related SAEs in group)	Expected			
	Unexpected			

*Unblinded assessment

Table 13: Impression of change and blinding. Figures denote no. and % of group unless stated

		Arthroscopy only (n=)	**Arthroscopy with device (n=)**	**Total (n=)**
How is your shoulder now?	Substantially Better			
	Moderately Better			
	No Different			
	Moderately Worse			
	Substantially Worse			
Global impression of change	No change or worse			
	Almost the same			
	A little better			
	Somewhat better			
	Moderately better			
	Better			
	A great deal better			
Patient reported treatment group	Arthroscopy only			
	Arthroscopy with device			
	Not sure			

Table 14: Analgesia usage. Figures denote no. and % of group unless stated

		Arthroscopy only (n=)	**Arthroscopy with device (n=)**	**Total (n=)**
Using any analgesia for shoulder: Yes				
Paracetamol	Daily			
	Up to weekly			
	Up to monthly			
	None			
Ibuprofen	Daily			
	Up to weekly			
	Up to monthly			
	None			
Codeine	Daily			
	Up to weekly			
	Up to monthly			
	None			
Co-codamol	Daily			
	Up to weekly			
	Up to monthly			
	None			
Tramadol	Daily			
	Up to weekly			
	Up to monthly			
	None			
…	Daily			
	Up to weekly			
	Up to monthly			
	None			

9 List of abbreviations

Abbreviation	Explanation
AE	Adverse Event
AHD	Acromio-humeral distance
CA	Coraco-Acromial (a small ligament in the shoulder)
Constant	Constant Murley score
CONSORT	*Con*solidated *S*tandards *o*f *R*eporting *T*rials
CRF	Case Report Form
CTU	Clinical Trials Unit
DMC	Data Monitoring Committee
ISRCTN	International Standard Randomised Controlled Trial Number
MCID	Minimal Clinically Important Difference
MRI	Magnetic Resonance Imaging
NICE	The National Institute for Health and Care Excellence
NIHR	The National Institute for Health Research
OSS	Oxford Shoulder Score
PPI	Patient & Public Involvement
PROMs	Patient Reported Outcome Measures
QoL	Quality of Life
QALY	Quality Adjusted Life Year
RCT	Randomised Controlled Trial
RDS	Research Design Service
REACTS	Randomised, Efficient, Adaptive Clinical Trial in Surgery
R&D	Research and Development
SAE	Serious Adverse Event
SOP	Standard Operating Procedure
START	Subacromial spacer for Tears Affecting Rotator cuff Tendons (Study title)
TMG	Trial Management Group
TSC	Trial Steering Committee
VAS	Visual Analogue Scale
WCTU	Warwick Clinical Trials Unit
WORC	Western Ontario Rotator Cuff