Subacromial spacers for adults with symptomatic, irreparable rotator cuff tears: the START:REACTS novel group sequential adaptive RCT

Andrew Metcalfe,[1,2*] Susanne Arnold,[1] Helen Parsons,[1] Nicholas Parsons,[1] Gev Bhabra,[2] Jaclyn Brown,[1] Howard Bush,[2] Michael Diokno,[2] Mark Elliott,[3] Josephine Fox,[4] Simon Gates,[1,5] Elke Gemperlé Mannion,[1] Aminul Haque,[1] Charles Hutchinson,[1,2] Rebecca Kearney,[6] Iftekhar Khan,[1] Tom Lawrence,[2] James Mason,[1] Usama Rahman,[2] Nigel Stallard,[1] Sumayyah Ul-Rahman,[1] Aparna Viswanath,[7] Sarah Wayte,[2] Stephen Drew[2] and Martin Underwood[1,2] on behalf of the START:REACTS team

[1]Warwick Medical School, University of Warwick, Coventry, UK
[2]University Hospitals Coventry and Warwickshire NHS Trust, Coventry, UK
[3]WMG, University of Warwick, Coventry, UK
[4]Patient Representative, Durham, UK
[5]Cancer Research UK Clinical Trials Unit, University of Birmingham, Birmingham, UK
[6]Bristol Medical School, University of Bristol, Bristol, UK
[7]South Tees Hospitals NHS Foundation Trust, Middlesbrough, UK

[*]Corresponding author a.metcalfe@warwick.ac.uk

Disclosure of interests of authors

Full disclosure of interests: Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at https://doi.org/10.3310/TKJY2101.

Primary conflicts of interest: Andrew Metcalfe, Helen Parsons, Elke Gemperlé Mannion, Charles Hutchinson, James Mason, and Martin Underwood are co-investigators on two other NIHR-funded trials: Robotic Arthroplasty: A Clinical and cost Effectiveness Randomised controlled trial (RACER)-Knee and RACER-Hip (Andrew Metcalfe leads RACER-Knee), for which Stryker also fund treatment costs and some imaging costs. As with the presented study, the full independence of the study team is protected by legal agreements.

Andrew Metcalfe, Susanne Arnold, Helen Parsons, Nicholas Parsons, Elke Gemperlé Mannion, Aminul Haque, Charles Hutchinson, Rebecca Kearney, Iftekhar Khan, James Mason, Nigel Stallard and Martin Underwood all work on other NIHR-funded studies. Charles Hutchinson, Rebecca Kearney, James Mason and Martin Underwood are or have been members of funding panels in NIHR, although not on the EME programme. Rebecca Kearney is chair of the NIHR Programme Grants for Applied Research

(PGfAR) committee, a paid position in NIHR but unrelated to the trial. She is also a previous chair of the NIHR Research for Patient Benefit (RfPB) committee and previous member of the Health Technology Assessment (HTA) Clinical Evaluation and Trials Committee and NIHR Integrated Clinical Academic (ICA) doctoral committee. Martin Underwood was a member of the NIHR Journals Library Editors Group and HTA Commissioning Committee. Until March 2021 he was an NIHR Senior Investigator. Until March 2020 he was an editor of the NIHR journal series, and a member of the NIHR Journal Editors Group, for which he received a fee. Simon Gates was a member of the NIHR Clinical Trials Unit Standing Advisory Committee and EME – Funding Committee Members and currently part of the HTA General Committee. Stephen Drew held an educational consultancy with Wright from 1 April 2016 until it was acquired by Stryker in 2021, when it migrated to an educational consultancy with Stryker from 1 April 2022.

Outside of this study, the authors report no personal financial conflict of interest with Stryker or any other related commercial organisation.

Published August 2023
DOI: 10.3310/TKJY2101

Scientific summary

Subacromial spacers for adults with symptomatic, irreparable rotator cuff tears: the START:REACTS novel group sequential adaptive RCT

Efficacy and Mechanism Evaluation 2023; Vol. 10: No. 3
DOI: 10.3310/TKJY2101

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Tears of the rotator cuff tendons of the shoulder are a common cause of shoulder pain and disability. Individuals often present with pain and restricted movement, as well as loss of strength and function affecting even simple activities of daily living, such as hair brushing.

Of those presenting with a rotator cuff tear, around half are treated surgically but roughly one-third of tears cannot be repaired. Compared with those who have a repair, people with irreparable tears have more severe pain, worse outcomes from treatment and more limited management options. New surgical techniques have been introduced to improve care, including the InSpace® (Stryker, Kalamazoo, MI, USA) subacromial balloon spacer.

The InSpace balloon is a saline-filled biodegradable balloon that is inserted surgically in the space between the humerus and acromion at the end of an arthroscopic debridement for people with an irreparable rotator cuff tear. By acting as a cushion between the acromion and the humerus, and potentially reducing friction, the device aims to improve the mechanics of the affected shoulder and aid rehabilitation. It deflates between 3 and 6 months after insertion, by which time it is hoped that the mechanics of the shoulder have improved.

The National Institute for Health and Care Excellence called for a randomised trial, recommending that the device should be used in research only. It received Food and Drug Administration (FDA) clearance in the United States in July 2021. Approximately 29,000 devices had been implanted outside the United States prior to FDA approval. Encouraging clinical results were observed from small early case series but some studies have reported poor results or cases of inflammation and pain.

Novel surgical procedures can expose patients to harm and should be carefully evaluated before widespread use, ideally with clinical trials. Surgical trials can provide high levels of evidence but can take a long time to complete. Adaptive designs can reduce the time needed to perform trials, potentially exposing fewer people to risk than with traditional fixed sample designs. Surgical trials typically use outcomes of 12 months or more, but by using the correlation between early and late outcomes, the advantages of adaptive designs could be extended to trials of new surgical techniques.

Objective

The primary objective was to assess the clinical effectiveness of arthroscopic debridement with the InSpace subacromial spacer balloon compared with arthroscopic debridement alone for people with symptomatic irreparable tears of the rotator cuff, based on the Oxford Shoulder Score (OSS) at 12 months.

Our secondary clinical objectives were:

- to compare the two interventions using both patient-reported and objective outcome scores at 3, 6 and 12 months
- to perform an economic analysis to assess the comparative cost-effectiveness of the two interventions
- to perform a magnetic resonance imaging (MRI) substudy to evaluate the proposed mechanism of the balloon, with scans taken at 8 weeks and at least 6 months after treatment.

The methodological objectives of the study were:

- to implement an efficient adaptive clinical trial design with the potential to stop early either for futility or efficacy, using outcome data available at 3, 6 and 12 months
- to evaluate the novel adaptive trial design using real trial data from a number of previous high-impact orthopaedic trials
- to explore the challenges of supporting adaptive design decision-making health economic analyses
- to compare the use of frequentist and Bayesian design and analysis on the conduct and interpretation of an adaptive surgical clinical trial with reference to decision-making by data monitoring committees (DMCs) during the study.

Trial methods

Design

We performed a participant- and assessor-blinded multicentre superiority randomised controlled trial (RCT; IDEAL stage 3) across 24 centres in the UK using a group sequential adaptive design with two preplanned interim analyses.

Participants

Adults with a rotator cuff tear with intrusive symptoms (pain and loss of function) deemed by the treating clinician to be technically irreparable, for whom conservative management had been unsuccessful.

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

- advanced shoulder osteoarthritis on preoperative imaging; subscapularis deficiency or pseudoparalysis (these three criteria are contraindications for the InSpace balloon)
- the clinician had determined that interposition grafting or tendon transfers were indicated
- an unrelated ipsilateral shoulder disorder
- neurological or muscular conditions that would interfere with strength measurement or rehabilitation
- previous proximal humeral fracture
- previous entry into the trial (i.e. for the other shoulder)
- those unable to complete trial procedures and those unfit for surgery.

All participants gave prior written consent. Eligibility was assessed prior to consent on the morning of surgery and intraoperatively (after assessment of the tear and surrounding structures in the shoulder) immediately prior to randomisation.

Intervention

The control group (debridement only) underwent arthroscopic debridement of the subacromial space and biceps tenotomy (if not already torn). Surgery was performed by subspecialty trained shoulder surgeons, using their normal surgical technique, within the confines described in a trial-specific surgical manual and surgical video.

The intervention group (debridement with InSpace balloon) underwent the same arthroscopic debridement procedure, followed by insertion of the InSpace balloon. The manufacturer's recommended technique was followed as documented in the surgical manual and was confirmed with them before distributing to surgeons. Surgical training was offered and a training course was run at the start of the trial. A company representative was invited to attend cases for technical support.

Participants were offered the same rehabilitation, including a home exercise programme and at least three face-to-face physiotherapy sessions. For both groups, fidelity was assessed with an operative

record form, and arthroscopic photographs. The number of physiotherapy visits for each participant in both groups was also documented.

Outcome measures

The primary outcome was the OSS 12 months after randomisation. The OSS is a 12-item participant-reported measure of shoulder-related pain and function. A higher score (0–48) corresponds with a better outcome.

The study was originally designed with the Constant score as the primary outcome. However, during the COVID-19 outbreak in March 2020 (in the recruitment phase of the trial), it was decided to change this to the OSS. This was because the Constant score requires face-to-face contact to measure and is usually assessed in hospital clinics; this would have exposed participants to unnecessary risk during the height of the pandemic. The OSS correlates well with the Constant score; both outcomes assess pain and shoulder function, are similarly responsive and have comparable worthwhile effect sizes in rotator cuff pathology.

We collected secondary outcomes at baseline and at 3, 6 and 12 months post-randomisation, including (where possible) the Constant score, range of pain-free movement and strength of shoulder abduction and flexion the shoulder, the Western Ontario Rotator Cuff (WORC) index (scored 0–100), health utility assessed using the EuroQol-5 Dimensions, five-level version (EQ-5D-5L), Participant Global Impression of Change scale, health-care resource use, analgesia use and adverse events. We defined adverse events as any condition of the affected shoulder or any event related to the anaesthetic or rehabilitation.

Trial results

Of 317 eligible people, 249 (79%) agreed to join the trial. A predefined futility stopping boundary was met at the first interim analysis and recruitment stopped with 117 participants randomised. A total of 61 (52%) participants were randomised to receive debridement surgery alone and 56 (48%) were randomised to receive debridement with the InSpace balloon.

Twelve-month primary outcome data were obtained from 114 of the 117 participants (97%). Of the three items of missing primary outcome data, two participants had died (neither death was trial related) and one participant could not be contacted. Scores improved compared with baseline in both groups. In the primary intention-to-treat analysis, the mean OSS at 12 months was 34.3 in the debridement-only group (n = 59) and 30.3 in the debridement with InSpace balloon group [n = 55; mean difference −4.2, favouring control; 95% confidence interval (CI) −8.2 to −0.26; p = 0.037]. Using a prespecified secondary adjusted model to account for the baseline OSS, sex, tear size and age group, a similar treatment effect was observed (effect −4.2, 95% CI −7.8 to −0.6; p = 0.026). There was no difference in safety events.

The Constant score, range of flexion and abduction and WORC index results were consistent with the primary analysis (the Constant score and range of motion measures had a high amount of missing data due to COVID-19 restrictions). The differences in WORC index and EQ-5D-5L at 12 months were not statistically significant (WORC index: −8.4, 95% CI −16.8 to −0.1; p = 0.055; EQ-5D-5L: −0.056, 95% CI −0.15 to 0.03; p = 0.24). In both cases the direction of change favoured debridement-only.

In cost-effectiveness analyses, quality-adjusted life-years were higher in the debridement-only group and costs were lower compared with the debridement with the InSpace device, in terms of both direct health-care costs and wider societal costs. As a result, debridement-only dominated with a probability of <1% that the InSpace device is cost-effective.

Magnetic resonance imaging substudy

We developed and refined a technique for dynamic MRI of a shoulder under deltoid load. We undertook a developmental work package in which we applied the technique to participants awaiting rotator cuff repair surgery who underwent electromyography evaluation of deltoid muscle contraction, which allowed us to determine the most appropriate technique to consistently achieve muscle activation in the narrow confines of a MRI scanner. We then piloted the MRI technique using this muscle activation protocol and a fast acquisition sequence for collecting both resting and active images. The main outcome of interest was the acromiohumeral distance, which was used as a marker of humeral migration under load.

We applied this technique in a mechanistic substudy, with scans taken at 8 weeks and at least 6 months after randomisation. Recruitment was severely hampered both by the early adaptive stop and the effects of the pandemic, which prevented continuing data collection. Despite the small sample size, we were able to observe narrowing of the acromiohumeral distance under load, demonstrating that the MRI technique was effective. We did not observe any between-group differences, although numbers were very low for this analysis.

Adaptive designs for surgical trials

We applied the novel adaptive design methodology that was developed for the main study to a number of previous high-impact randomised trials in trauma and orthopaedic surgery. The study assessed whether each of a selected number of RCTs, originally implemented using conventional sample size designs, would have stopped early if a group sequential trial design had been used, and what the final outcome would have been had they done so.

We received anonymised data from seven large multicentre trials: Wound Management of Open Lower Limb Fractures (WOLLF); Distal Radius Acute Fracture Fixation Trial (DRAFFT); UK Fixation of Distal Tibia Fractures (UK FixDT); UK Full Randomised Controlled trial of Arthroscopic Surgery for Hip Impingement versus best Conventional (FASHIoN); the Warwick Arthroplasty Trial (WAT); Can Shoulder Arthroscopy Work? (CSAW); the Total or Partial Knee Arthroplasty Trial (TOPKAT). The temporal sequence of data accumulation was replicated in exactly the manner it was in the original study, using the dates when each outcome measure was taken. We selected planned interim analyses and stopping boundaries and simulated how each study might have progressed using the methodological approach described in this monograph.

The results for five of the studies (WOLLF, FixDT, DRAFFT, FASHIoN and WAT) showed that adaptive design using early outcome data would have been feasible and likely to provide designs that were at least as efficient, and possibly more efficient, than the original fixed sample size designs. For WOLLF and FixDT the simulations showed that it was highly likely that these studies would have (correctly) stopped early for futility, both over one year early, saving potential considerable effort and resources. The two studies that showed modest effect estimates at interim analyses in favour of the test treatment (WAT and DRAFFT) did not stop early, which was consistent with the final results of these studies. The FASHIoN trial showed good evidence in favour of the test surgical intervention in the final analysis but fell short of stopping at the interim analyses. For this study, it would have been possible to select different but sensible boundaries that would have resulted in early stopping for efficacy. TOPKAT and CSAW would not have been suitable for the adaptive design methods in their current form, as they had either longer primary outcomes times or less early time point data, although it is reasonable to think that these issues could have been resolved if the method was applied prospectively.

For all the studies, it was clear that the feasibility and practicality of using the proposed adaptive design methods was determined in large part by: (1) whether the timing of recruitment allows for interim

analyses; (2) the availability of early outcome data and correlations with final outcomes; (3) recruitment and outcome data accrual profiles; and (4) the estimates of correlation and covariance parameters at the design planning stage.

Interim analysis report study

To understand the influence of different approaches to adaptive trial design on the conduct of DMC decisions, we performed an exploratory study of the interpretation of Bayesian and frequentist interim analysis reports by potential committee members. We found that potential DMC members do not always choose to follow the stopping rules that are presented and would benefit from more ancillary information to support their decision-making and understanding of the analysis, regardless of the statistical framework used.

Conclusions

We used a blinded RCT with predefined stopping boundaries to test whether the InSpace balloon was of benefit for people with irreparable rotator cuff tears. The study stopped at just over half the maximum sample size, allowing us to report the findings early. In the primary analysis, arthroscopic debridement was found to be superior to arthroscopic debridement with the InSpace balloon for people with an irreparable cuff tear of the shoulder, based on the OSS 12 months after surgery. Secondary outcomes and cost-effectiveness analysis agreed, and effectively exclude the possibility of any meaningful benefit for the InSpace balloon. This trial has delivered evidence that the InSpace balloon is not an effective treatment, could be harmful, and is very unlikely to be cost-effective.

Randomised trials are needed early in the evaluation of new technologies to prevent harm to patients and cost to society, but also to allow effective treatments to be offered widely. We have demonstrated that using adaptive designs in surgical trials are possible and practical. By delivering efficient, effective trial designs early in the introduction of new procedures and technologies, we will make major cost savings for the health service and deliver better patient outcomes both now and into the future.

Trial registration

This trial is registered as ISRCTN17825590.

Funding

This project (project reference 16/61/18) was funded by the Efficacy and Mechanism Evaluation (EME) programme, a Medical Research Council and National Institute for Health and Care Research (NIHR) partnership. The trial is co-sponsored by the University of Warwick and University Hospitals Coventry and Warwickshire NHS Trust. This study will be published in full in *Efficacy and Mechanism Evaluation*; Vol. 10, No. 3. See the NIHR Journals Library website for further project information.

Efficacy and Mechanism Evaluation

ISSN 2050-4365 (Print)

ISSN 2050-4373 (Online)

Efficacy and Mechanism Evaluation (EME) was launched in 2014 and is indexed by Europe PMC, DOAJ, Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and NCBI Bookshelf.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

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

The full EME archive is freely available to view online at www.journalslibrary.nihr.ac.uk/eme.

Criteria for inclusion in the *Efficacy and Mechanism Evaluation* journal

Reports are published in *Efficacy and Mechanism Evaluation* (EME) if (1) they have resulted from work for the EME programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

EME programme

The Efficacy and Mechanism Evaluation (EME) programme funds ambitious studies evaluating interventions that have the potential to make a step-change in the promotion of health, treatment of disease and improvement of rehabilitation or long-term care. Within these studies, EME supports research to improve the understanding of the mechanisms of both diseases and treatments.

The programme supports translational research into a wide range of new or repurposed interventions. These may include diagnostic or prognostic tests and decision-making tools, therapeutics or psychological treatments, medical devices, and public health initiatives delivered in the NHS.

The EME programme supports clinical trials and studies with other robust designs, which test the efficacy of interventions, and which may use clinical or well-validated surrogate outcomes. It only supports studies in man and where there is adequate proof of concept. The programme encourages hypothesis-driven mechanistic studies, integrated within the efficacy study, that explore the mechanisms of action of the intervention or the disease, the cause of differing responses, or improve the understanding of adverse effects. It funds similar mechanistic studies linked to studies funded by any NIHR programme.

The EME programme is funded by the Medical Research Council (MRC) and the National Institute for Health and Care Research (NIHR), with contributions from the Chief Scientist Office (CSO) in Scotland and National Institute for Social Care and Health Research (NISCHR) in Wales and the Health and Social Care Research and Development (HSC R&D), Public Health Agency in Northern Ireland.

This report

The research reported in this issue of the journal was funded by the EME programme as project number 16/61/18. The contractual start date was in February 2018. The final report began editorial review in April 2022 and was accepted for publication in August 2022. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The EME editors and production house have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research. The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the MRC, the EME programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the EME programme or the Department of Health and Social Care.

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