Subacromial balloon spacer for irreparable rotator cuff tears of the shoulder (START:REACTS): a group-sequential, double-blind, multicentre randomised controlled trial

Andrew Metcalfe, Helen Parsons, Nicholas Parsons, Jaclyn Brown, Josephine Fox, Elke Gemperlé Mannion, Aminul Haque, Charles Hutchinson, Rebecca Kearney, Iftekhar Khan, Tom Lawrence, James Mason, Nigel Stallard, Martin Underwood, Stephen Drew*, on behalf of the START:REACTS team*

Summary

Background New surgical procedures can expose patients to harm and should be carefully evaluated before widespread use. The InSpace balloon (Stryker, USA) is an innovative surgical device used to treat people with rotator cuff tears that cannot be repaired. We aimed to determine the effectiveness of the InSpace balloon for people with irreparable rotator cuff tears.

Methods We conducted a double-blind, group-sequential, adaptive randomised controlled trial in 24 hospitals in the UK, comparing arthroscopic debridement of the subacromial space with biceps tenotomy (debridement only group) with the same procedure but including insertion of the InSpace balloon (debridement with device group). Participants had an irreparable rotator cuff tear, which had not resolved with conservative treatment, and they had symptoms warranting surgery. Eligibility was confirmed intraoperatively before randomly assigning (1:1) participants to a treatment group using a remote computer system. Participants and assessors were masked to group assignment. Masking was achieved by using identical incisions for both procedures, blinding the operation note, and a consistent rehabilitation programme was offered regardless of group allocation. The primary outcome was the Oxford Shoulder Score at 12 months. Pre-trial simulations using data from early and late timepoints informed stopping boundaries for two interim analyses. The primary analysis was on a modified intention-to-treat basis, adjusted for the planned interim analysis. The trial was registered with ISRCTN, ISRCTN17825590.

Findings Between June 1, 2018, and July 30, 2020, we assessed 385 people for eligibility, of which 317 were eligible. 249 (79%) people consented for inclusion in the study. 117 participants were randomly allocated to a treatment group, 61 participants to the debridement only group and 56 to the debridement with device group. A predefined stopping boundary was met at the first interim analysis and recruitment stopped with 117 participants randomised. 43% of participants were female, 57% were male. We obtained primary outcome data for 114 (97%) participants. The mean Oxford Shoulder Score at 12 months was 34·3 (SD 11·1) in the debridement only group and 30·3 (10·9) in the debridement with device group (mean difference adjusted for adaptive design –4·2 [95% CI –8·2 to –0·26];p=0·037) favouring control. There was no difference in adverse events between the two groups.

Interpretation In an efficient, adaptive trial design, our results favoured the debridement only group. We do not recommend the InSpace balloon for the treatment of irreparable rotator cuff tears.

Funding Efficacy and Mechanism Evaluation Programme, a Medical Research Council and National Institute for Health and Care Research partnership

Published **Online** April 21, 2022 https://doi.org/10.1016/S0140-6736(22)00652-3

See **Online/Comment** https://doi.org/10.1016/S0140-6736(22)00733-4

*Contributed equally

University Hospitals Coventry and Warwickshire NHS Trust, Coventry, UK (A Metcalfe PhD, Prof C Hutchinson MD, T Lawrence FRCS, Prof M Underwood MD, S Drew FRCS, Prof R Kearney PhD); **Warwick Medical School, University of Warwick, Coventry, UK** (A Metcalfe, H Parsons PhD, N Parsons PhD, J Brown BSc, E Gemperlé Mannion BA, A Haque MSc, Prof C Hutchinson, Prof R Kearney, I Khan PhD, Prof J Mason PhD, Prof N Stallard PhD, Prof M Underwood); **Houghton le Spring, UK** (J Fox)

Correspondence to: Mr Andrew Metcalfe, Warwick Medical School, University of Warwick, Coventry CV4 7HL, UK **a.metcalfe@warwick.ac.uk**

Introduction

Tears of the rotator cuff tendons surrounding the shoulder joint are a common cause of shoulder pain and disability.[1–3] People with rotator cuff tears often have severe pain that wakes them from sleep and restricts even simple activities of daily living.[2–4] Rotator cuff tears are an increasingly common presentation to health-care services and result in substantial expense to society through treatment costs and loss of ability to work.[5,6] Approximately half of those who present with a tear of the rotator cuff are treated with surgery.[5] Surgical repair of the torn tendon is often performed but around a third of tears cannot be repaired.[5] Tears can become irreparable as the tendon becomes scarred and retracted or the muscle atrophies, such that the torn tissue cannot be repaired to its original site of attachment;[7] these types of tear are typically large tears and are more common in older people. People with irreparable tears have more severe pain and disability, worse outcomes from surgery, and fewer treatment options compared with those who have had a repair.[7–9] Consequently, new surgical techniques have been introduced to improve care, including the InSpace subacromial balloon spacer (Stryker, USA).[7–9] The InSpace device is a saline-filled biodegradable balloon that is inserted surgically in the space between the

Research in context

Evidence before this study

The InSpace device received Conformité Européenne marking in 2010, and had been used in 29 000 procedures until July, 2021, when it received Food and Drug Administration approval in the USA. The National Institute for Health and Care Excellence (NICE) reviewed the available data for InSpace devices in May, 2016, and found one published case series (n=20) and two conference abstracts (n=82). These showed overall significant improvements in shoulder scores but some cases of synovitis and early pain in individuals receiving the balloon. They recommended it only be used for research, including a randomised controlled trial. We searched Medline, Embase, the Cochrane Central Register of Controlled Trials, and CINAHL from inception to Sept 14, 2016, for randomised trials. We used keywords for rotator cuff, supraspinatus, infraspinatus, subscapularis, or teres minor; and tear, tears, torn, or rupture. We found 57 trials (n=4542), mostly of repair, but no trials using the InSpace device. A 2020 systematic review found 19 studies of the InSpace device, all case series, including 337 patients. To our knowledge, except for our trial, there has been one other randomised controlled trial, a company-funded study performed in the USA of partial rotator cuff repair compared with the use of the device. Pre-publication data were posted to a trial registry in 2021, but had not been published in a peer-review journal before April, 2022. Partial cuff repair is a different comparator to the one used for our study.

Added value of this study

To our knowledge, this is the only randomised trial to compare the InSpace device to the same treatment without the device. It was a multicentre study using eligibility criteria that aligned with the accepted indications for the device. A high proportion of people who were approached consented for the trial; therefore, the findings are likely to be generalisable to the majority of people who would receive the device. Patients and assessors were blinded to trial group allocation, with intraoperative randomisation ensuring allocation concealment to ensure a low risk of bias. Early stopping rules were established prospectively using trial simulations and the study stopped at the first preplanned interim analysis. We found that the InSpace balloon is unlikely to provide benefit or be cost effective, and it might be harmful, especially in females. This is the first published randomised trial on the balloon and the first study to clearly demonstrate an absence of benefit for the device. We used a novel adaptive design that utilised the correlation between early and late timepoints to increase the efficiency of the interim analyses, this methodology could be of benefit to future researchers, especially in evaluating surgical technologies.

Implications of all the available evidence

Rotator cuff tears are a common cause of pain and disability. Although many tears can be repaired, some cannot. Irreparable rotator cuff tears are a difficult problem to treat. A range of surgical procedures are available from minor keyhole (arthroscopic) procedures to more major joint replacement procedures. The InSpace balloon is one option, and it can be inserted as part of a relatively simple arthroscopic procedure. It has been used for the past decade in Europe and was recently approved for use in the USA. This trial has delivered high-quality evidence that the InSpace device is not an effective treatment, could be harmful, and is unlikely to be cost-effective. We do not recommend its use. Additional randomised controlled trials of other treatments are needed, as there is a low level of evidence for all available treatment options for irreparable rotator cuff tears.

humerus and the acromion.[9–11] By maintaining the gap between the acromion and the humerus, and potentially reducing friction, the device aims to improve the mechanics of the affected shoulder and aid rehabilitation.[9–11] The device received a Conformité Européenne mark in 2010, and was introduced to the UK in 2013. In 2016, the National Institute for Health and Care Excellence released a research recommendation, which included the need for a clinical trial.[12] It received Food and Drug Administration (FDA) clearance in the USA in July, 2021, with approximately 29 000 devices having been implanted outside the USA before this.[13] Early small case series documented encouraging clinical results, but some studies have reported poor results or cases of inflammation and pain; therefore, to determine if the device is effective, randomised data are needed.[9–12]

New surgical techniques and devices inherently expose patients to risk.[14,15] They are often introduced into clinical practice on the basis of early case series data, and trials might follow. Surgical trials typically take many years to complete, during which time patients can suffer the consequences of ineffective or harmful treatments. Adaptive designs can reduce the time needed to perform trials and expose fewer people to risk.[16] Adaptive trials can be challenging in surgery as longer-term outcomes are often used, typically 12 months or more. By using the correlation between earlier outcomes and later primary outcomes, adaptive designs could be extended to many more settings, including trials of new surgical techniques.[17–19]

We report on a novel, efficient, adaptive clinical trial to assess the clinical effectiveness of a subacromial spacer balloon for people with symptomatic irreparable tears of the rotator cuff.[18,20]

Methods

Study design and participants

We performed a participant-assessor double-blind, multicentre, superiority randomised controlled trial (IDEAL stage 3) across 24 hospitals in the UK using a group sequential adaptive design with two preplanned interim analyses. The study was approved by the Health

Research Authority and the West Midlands – Coventry and Warwickshire Research Ethics Committee (18/WM/0025) in February, 2018. The adaptive design methods, study protocol, and statistical analysis plan are publicly available and links are provided in the appendix (p 1).[18,20]

Participants were recruited in outpatient clinics or from the surgical waiting list. We included adults with a rotator cuff tear and intrusive symptoms (pain and loss of function) for whom conservative management had been unsuccessful, and for whom the treating clinician considered that surgery was warranted and the tear technically irreparable. Exclusion criteria were: advanced shoulder osteoarthritis on usual care preoperative imaging; subscapularis deficiency; pseudoparalysis (these three criteria are contraindications for the device); cases in which the clinician 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 (ie, for the other shoulder); unable to complete trial procedures; and those unfit for surgery. Full details on exclusion criteria are in the published protocol.[20]

All participants gave written informed consent. Eligibility was assessed before consent, on the morning of surgery and intraoperatively (after assessment of the tear and surrounding structures in the shoulder) immediately before randomisation. If a surgeon found that the rotator cuff tear could be repaired at the time of surgery, the participant was excluded from the study and was not randomised. This decision was at the judgement of the surgeon at the time and is consistent with normal clinical practice for such tears.

Randomisation and masking

After standard shoulder arthroscopy, intraoperative confirmation of eligibility, and measurement of the tear size, participants were randomly assigned 1:1 to a treatment group (debridement only group or debridement with device group) using a central web-based system that was accessed using either a web-based platform or an automated telephone system. To maintain allocation concealment, randomisation could only be accessed after intraoperative eligibility and tear size were confirmed. The randomisation sequence was generated using a minimisation algorithm with a random element (70% weighting), which included factors for site, sex, age (≥70 years or <70 years), and intraoperative cuff tear size (≥3 cm or <3 cm).

Participant masking was maintained using incisions in the same location and size for both groups. The only difference between the incisions needed for the two procedures is that the lateral portal incision for the balloon is typically 0·5 cm larger (1·5 cm instead of 1 cm) than the standard incision. The same 1·5 cm incision was used for both groups to maintain masking, a difference that our patient representatives felt was minimal, but it ensured there was no visible difference between the groups. In the setting of awake surgery, drapes and screens were used with careful intraoperative communication using written notes to communicate the allocation to the surgeon. The clinical operation note was blinded to prevent accidental unblinding by perioperative care or rehabilitation staff. Intervention-related information was recorded by surgeons directly onto a secure database with an unblinding process available to National Health Service staff in case of emergency. Outcome assessments were only performed by staff who were masked to the group allocation, and who had not been involved in the surgery or the randomisation. After the collection of the primary outcome at 12-months, participants were asked which group they thought they were in or if they were unaware of the allocation.

See Online for appendix

Procedures

The control group (debridement-only) underwent arthroscopic debridement of the subacromial space and biceps tenotomy (if not already torn), which was performed by subspeciality trained shoulder surgeons, who followed a technique manual and surgical video. Further details are provided in the protocol paper.[20] The intervention group (debridement with device) underwent the same procedure, followed by insertion of the InSpace balloon (Stryker, USA). The manufacturer's recommended technique was followed for the technique manual and was confirmed with them before distributing it to the surgeons. Surgical training was offered to all surgeons delivering the trial interventions, and a training course was run at the start of the trial. A company representative from OrthoSpace (Stryker after 2019) was invited to attend cases in theatre for technical support. Fidelity was assessed with arthroscopic photos, assessed by a subspecialty trained shoulder surgeon who was masked to treatment allocation, a surgical form, and self-reporting of physiotherapy visits.

All participants in both groups were offered the same rehabilitation, including a home exercise programme and at least three face-to-face physiotherapy sessions. The anaesthetic choice and use of prophylactic antibiotics were chosen according to usual clinical practice.

All primary and secondary outcomes were assessed at 3, 6, and 12 months after randomisation. Face-to-face assessments were performed at these timepoints, although these were severely limited by COVID-19 restrictions. Participant-reported outcome measures (PROMs) were collected at the same timepoints, either by post or in clinic if a face-to-face assessment was performed. Where postal follow-up was unsuccessful twice, we did telephone follow-up.

Outcomes

The primary outcome for the study was the Oxford Shoulder Score, a 12-item participant-reported measure (scored 0–48; 48 is the best score) of

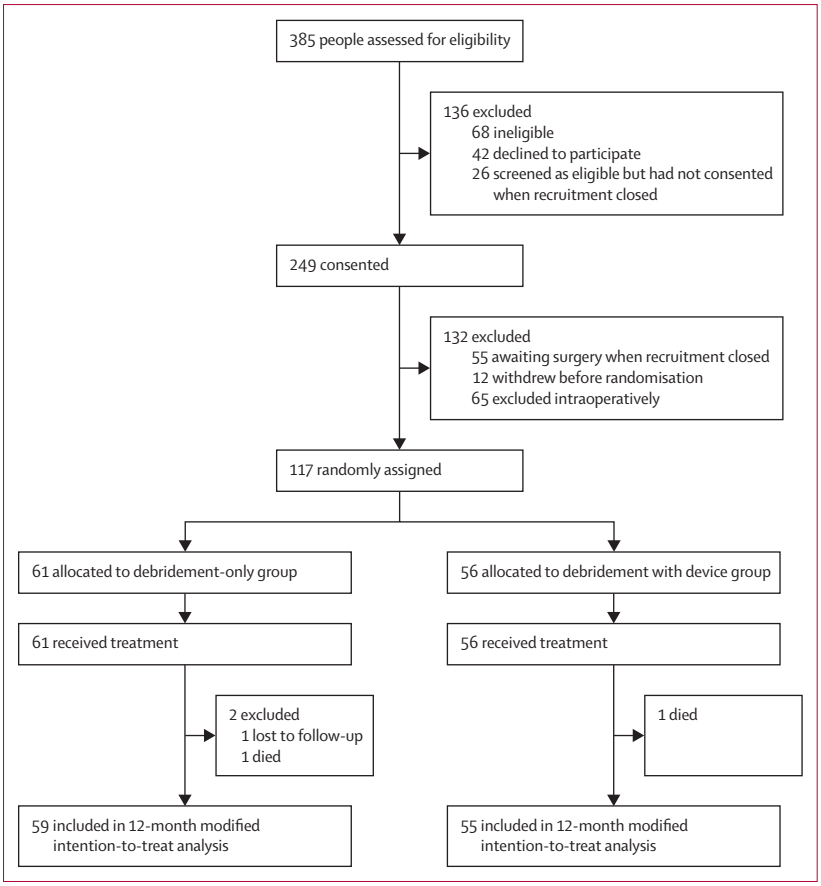

Figure 1: **Trial profile**
A full detailed version is available in the appendix (p 7).

shoulder-related pain and function, 12 months after participant randomisation.[21] In trials of treatments for rotator cuff tears, outcomes plateau by 12 months with a strong correlation between 12-month and 24-month outcomes.[19] Little additional information is provided by waiting for the 24-month score.[19] The study was originally designed around the Constant Score.[20–22] However, this requires face-to-face assessment, which could not be performed during much of the COVID-19 pandemic. In March, 2020 (in the recruitment phase of the trial), before any data analysis and with the approval of the oversight committees and the funder, we changed the primary outcome to the Oxford Shoulder Score.[20–22]

The secondary outcomes were the Constant Score (collected where possible), the range of pain free flexion and abduction, the Western Ontario Rotator Cuff (WORC) index (scored 0–100), EuroQol EQ-5D-5L, change in symptoms, Participant Global Impression of Change, resource use, and adverse events. Secondary outcomes that could not be taken due to missed face-to-face visits were recorded as missing data. We defined adverse events as any shoulder condition or any event related to the anaesthetic or rehabilitation. Serious adverse events were defined according to accepted Good Clinical Practice definitions. A MRI substudy was performed and will be reported separately.

Statistical analysis

Based on a meta-analysis of relevant trials,[19] we set the target difference in the Constant score at 10 units with a standard deviation of 20. The Oxford Shoulder Score requires the same sample size to show a target difference of six (based on its published minimal clinically important difference), with a standard deviation of 12, giving the same effect size as required for the Constant score; therefore, no sample size changes were required when the primary outcome was changed.[19,23–25] Correlations (Pearsons) of r=0·5 were expected between the Oxford Shoulder Score scores for participants at different timepoints.[19] Using previously described methods, simulations indicated that a maximum sample size of 188 (n=94 in each group) would provide at least 90% power to detect the target difference at the 5% level.[18] Allowing for 15% loss to follow-up, the maximum sample size required for the study was 221 participants.

Trial simulations were performed at the start of the study and were used to determine a set of predefined interim stopping boundaries for a group sequential adaptive design.[17,18] All of the available Oxford Shoulder Score data at 3, 6, and 12 months, and their respective correlations were used at each timepoint to increase the efficiency of the adaptive design (appendix p 4).[18] Two interim analyses were planned with binding rules for futility at the first interim analysis and futility and efficacy for the second interim analysis. The rules were prospectively agreed with the independent Data Monitoring Committee. The timing of the two analyses was determined by monitoring the information obtained from the observed correlations and variances of the Oxford Shoulder Score scores at each timepoint, which was performed monthly once the first 12-month data were received.

The primary analysis was on a modified intention-to-treat basis, adjusted for the planned interim analyses. We did not impute missing data points as the level of missing data was very low and considered likely to be missing at random. As such, the intention-to-treat analysis is described as modified, although we otherwise used a full analysis set with no modifications. The treatment-effect estimate for the primary analysis has been described in previous methodological work, which exploited correlations between the early and 12-month outcomes for participants, in which the latter were not available, with adjustment for potential bias due to the interim analyses made using Todd's approach.[18,26] If the study stopped at the first interim analysis, then no adjustment for bias would be necessary. Differences in favour of the intervention are expressed as positive values throughout. A secondary mixed-effect model was fitted to adjust for the baseline

scores and the predefined subgroups (age group, sex, and tear size), and the recruiting centre fitted as a random-effect variable. We did predefined exploratory subgroup analyses to assess any interaction effects with the interventions, these should be interpreted cautiously given the sample size. Secondary outcomes were analysed with mixed-effect models. Model assumptions were assessed visually for example using qq-plots and histograms. Analgesia usage, adverse events, and serious adverse events were analysed by Fisher's test, and change in symptoms and the Patient Global Impression of Change (PGIC) were analysed using adjusted proportional odds ordered regression models. Multiple imputation was not used as the level of missing data was very low. Statistical analyses were performed in R (version 4.0.3). A health economic evaluation was performed in parallel and a substudy using MRI was performed on a small number of participants, these will be reported separately. Study oversight was provided by an independent Data Monitoring Committee and a Trial Steering Committee. The study is registered with ISRCTN, ISRCTN17825590.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. Stryker had the opportunity to review the paper for intellectual property infringements or technical inaccuracies related to the company before submission. In the study design phase, the surgical and physiotherapy manuals were reviewed by Ortho-Space to confirm that the guidance followed their recommended techniques, and company representatives were invited to cases to ensure surgeons had full technical support.

Results

Between June 1, 2018, and July 30, 2020, we assessed 385 people for eligibility, of which 317 were eligible. 249 (79%) people were consented for inclusion in the study. 117 participants were randomly allocated to a treatment group, 61 participants to the debridement only group and 56 to the debridement with device group (figure 1). On July 30, 2020, recruitment and randomisation were stopped after the futility boundary had been crossed at the first interim analysis (appendix p 4). Baseline variables were well balanced (table 1). Five (8%) participants in the debridement only group and one (2%) participant in the debridement plus device group had rotator cuff tears of less than 3 cm. The mean tear sizes were similar between the groups (debridement only, 4·3 cm [SD 1·3] *vs* debridement with device, 4·2 cm [1·3]; table 1).

12-month primary outcome data were obtained from 114 (97%) of the 117 participants. Of the three participants with missing data, two participants had died (neither trial related) and one participant could not be contacted (table 2). The Oxford Shoulder Score improved in both groups compared with the baseline data (figure 2).

	Debridement only group (n = 61)	Debridement with device (n=56)	Total (n=117)
Age	67·3 (9·0)	66·4 (7·6)	66·9 (8·3)
Age group			
70 years and older	28 (46%)	20 (36%)	48 (41%)
Younger than 70 years	33 (54%)	36 (64%)	69 (59%)
Sex			
Female	28 (46%)	22 (39%)	50 (43%)
Male	33 (54%)	34 (61%)	67 (57%)
Rotator cuff tear size			
Large (≥3 cm)	56 (92%)	55 (98%)	111 (95%)
Medium or Small (<3 cm)	5 (8%)	1 (2%)	6 (5%)
Right shoulder affected	42 (69%)	37 (66%)	79 (68%)
Baseline PROM			
OSS	21·7 (9·4)	23·1 (8·5)	22·4 (9·0)
Constant Murley	33·6 (13·0)	29·9 (13·4)	31·9 (13·2)
WORC	34·4 (14·2)	33·7 (13·1)	34·1 (13·6)
EQ-5D-5L	0·501 (0·258)	0·486 (0·247)	0·494 (0·251)
Shoulder function			
Abduction angle,°	76·3 (32·8);n=58	63·9 (22·2); n=51	70·5 (28·9); n=109
Flexion angle,°	74·1 (25·1); n=58	67·8 (29·8); n=51	71·1 (27·4); n=109
Abduction Strength, kg	1·9 (1·7)	1·5 (2·7)	1·7 (1·9)
Acromiohumeral distance on baseline x-ray	6·8 (2·7); n=47	6·8 (2·1); n=46	6·8 (2·4); n=93
Symptom duration, years	4·3 (6·2)	5·5 (7·1)	4·9 (6·7)
Other medical conditions	53 (87%)	45 (80%)	98 (84%)
Current smoker	4 (7%)	5 (9%)	9 (8%)
Type 2 diabetes	9 (15%)	9 (16%)	18 (15%)
Unilateral symptoms	43 (70%)	39 (70%)	82 (70%)
Bilateral symptoms	18 (30%)	16 (29%)	34 (29%)
Previous received physiotherapy treatment	42 (69%)	44 (79%)	86 (74%)
Previously received steroid injection	34 (56%)	36 (64%)	70 (60%)
Number of steroid injections taken	2 (1–6)	2 (1–10)	2 (1–10)
Previously had surgery on shoulder	16 (26%)	9 (16%)	25 (21%)
Anterior-posterior tear size, cm	4·3 (1·3)	4·2 (1·3)	4·2 (1·3)
Medio-lateral retraction from greater tuberosity attachment, cm	4·3 (1·0)	4 (1·0)	4·1 (1·0)
Biceps tendon intact	38 (62%)	39 (70%)	77 (66%)
Subscapularis torn	12 (20%)	14 (25%)	26 (22%)
Subscapularis tear size, cm	0·7 (0·3)	0·8 (0·4)	0·8 (0·4)
Subscapularis repaired	2 (3%)	2 (4%)	4 (3%)

Data are n (%), mean (SD), or median (IQR) unless otherwise specified. Baseline data are complete for all participants except where numbers are given. PROM=patient-reported outcome measures. OSS=Oxford Shoulder Score. WORC=Western Ontario Rotator Cuff Index.

Table 1: **Baseline characteristics and operative findings.**

The mean Oxford Shoulder Score at 12 months was 34·3 (SD 11·1) in the debridement only group and 30·3 (10·9) in the debridement with device group. In the primary (adjusted for adaptive design only) analysis, the mean difference was –4·2 (95% CI –8·2 to –0·26; p=0·037) favouring debridement only. Using a prespecified secondary adjusted model to account for the baseline Oxford Shoulder Score, sex, tear size, and age group, a similar mean difference was observed (–4·2 [95% CI

	Debridement only (n = 61)	Debridement with device (n = 56)	Adjusted mean difference* (95% CI)
Primary outcome			
Oxford Shoulder Score, 0–48			
3 months	30·4 (11·2); n=59	25 (10·4); n=54	−5·4 (−8·5 to −2·4)
6 months	33·3 (10·4); n=58	28·5 (11); n=54	−5·6 (−9·0 to −2·1)†
12 months‡	34·3 (11·1); n=59	30·3 (10·9); n=55	−4·2 (−7·8 to −0·6)
Secondary outcomes			
Constant score, 0–100§			
3 months	46 (15·7); n=45	36·7 (21); n=41	−5·8 (−12·5 to 0·7)
6 months	49 (18·6); n=29	45·2 (19·9); n=26	−2·0 (−12·3 to 8·3)†
12 months	63·6 (11·2); n=11	47·5 (13·2); n=11	−13·8 (−24·0 to −3·6)†
Abduction angle§			
3 months	88·8 (36·6); n=45	69·7 (39·7); n=41	−14·3 (−30·9 to 2·1) †
6 months	97·5 (34·5); n=28	87·8 (41·9); n=26	−7·2 (−28·2 to 13·8)†
12 months	124·1 (37); n=12	87·1 (32·1); n=11	−34·1 (−77·1 to 8·8)†
Flexion angle§			
3 months	96·6 (36·1); n=45	84·2 (44·7); n=41	−4·5 (−20·5 to 11·5)†
6 months	103·9 (30·4); n=28	100·3 (46·4); n=26	−4·0 (−24·0 to 16·0)†
12 months	139·1 (26·4); n=12	98·8 (40·1); n=11	−56 ·8 (−91·1 to −22·5)†
Abduction Strength§, kg			
3 months	2·1 (1·9); n=45	1·8 (2·7); n=41	−0·3 (−1·3 to 0·7)¶
6 months	2·1 (1·3); n=30	1·8 (2·0); n=27	−0·4 (−1·5 to 0·8)¶
12 months	3·8 (2·0); n=11	1·5 (1·3); n=11	−2·3 (−3·8 to −0·8)¶
Western Ontario Rotator Cuff Index, 0–100			
3 months	54·8 (24·5); n=56	40·2 (19·2); n=52	−12·0 (−19·5 to −4·8)
6 months	60·2 (25·7); n=53	49·1 (22·6); n=52	−11 ·0 (−19·8 to −2·1)
12 months	61·6 (25·7); n=56	51·7 (23·5); n=51	−8·4 (−16·8 to −0·1)
EQ-5D-5L, −0·224–1			
3 months	0·632 (0·237); n=59	0·556 (0·275); n=55	−0·061 (−0·145 to 0·022)†
6 months	0·666 (0·253); n=58	0·592 (0·254); n=54	−0·064 (−0·144 to 0·015)
12 months	0·667 (0·287); n=58	0·590 (0·286); n=55	−0·056 (−0·150 to 0·035)
Overall change since operation at 12 months			0·6 (0·3 to 1·2)‖
Substantially better	24 (41%)	16 (29%)	..
Moderately better	17 (29%)	19 (34%)	..
No difference	12 (20%)	7 (13%)	..
Moderately worse	2 (3%)	6 (11%)	..
Substantially worse	4 (7%)	7 (13%)	..
Participant Global Impression of Change, since the operation at 12 months			0·5 (0·3 to 1·1)‖
Almost the same	5 (9%)	6 (11%)	..
A little better, no noticeable change	6 (10%)	3 (5%)	..
Somewhat better, change has not made a difference	4 (7%)	6 (11%)	..
Moderately better, slight but noticeable change	7 (12%)	6 (11%)	..
Better, definite improvement with a difference	17 (29%)	13 (24%)	..
Considerable improvement making a huge difference	13 (22%)	7 (13%)	..
Analgesia (number of participants taking any pain medications)			
Baseline	51/59 (84%)	41/53 (73%)	1·9 (0·7 to 5·1)**
3 months	38/58 (62%)	37/56 (66%)	0·8 (0·4 to 1·9)**
6 months	38/57 (62%)	34/54 (61%)	1·1 (0·5 to 2·4)**
12 months	30/59 (49%)	30/54 (49%)	0·8 (0·4 to 1·8)**

Data are n/N (%) or mean (SD) unless indicated otherwise. Adjusted model mean differences and 95% CIs are presented unless specified otherwise. Unadjusted data is shown in the text. OR<1 favours the debridement group only. OR=odds ratio. *Negative values favour the debridement-only group. †Fixed-effects model only (mixed-effects model did not converge). ‡Primary outcome. §Data collection limited by COVID-19 restrictions. ¶Unadjusted model results. ‖OR calculated via adjusted proportional ordered regression. **OR with 95% CI with Fishers exact test.

Table 2: **Patient reported outcomes at follow-up**

–7·8 to –0·6; p=0·026; table 2). The planned per-protocol analysis was not undertaken because all participants received their allocated (ie, randomised) intervention.

The Constant Score, range of flexion and abduction, and WORC index results were consistent with the primary analysis (table 2; The Constant Score, range of motion, and strength measures had a high amount of missing data due to COVID-19 restrictions). The mean difference in WORC index (–8·4 [95% CI –16·8 to –0·1]; p=0·055) and EQ-5D-5L at 12 months were not significant (–0·056 [95% CI –0·150 to 0·035]; p=0·24), and the direction of change favoured debridement only for both.

The overall change (adjusted odds ratio 0·6 [95% CI 0·3 to 1·2]; p=0·21) and Patient Global Impression of Change scores (adjusted odds ratio 0·5 [95% CI 0·3 to 1·1]; p=0·08) did not demonstrate a significant difference, although the direction of change favoured debridement only. There was no difference observed in analgesia use at 12 months (table 2).

There were no clear differences in safety events between the two groups. 11 (20%) participants had an adverse event in the debridement with device group, and nine (15%) participants had an adverse event in the debridement only group. Some people had multiple adverse events recorded, there were 17 adverse events in the debridement with device group and 11 in the debridement only group. There were six serious adverse events, four (7%) in the debridement with device group and two (3%) in the debridement only group. Three serious adverse events were considered unrelated to the surgery, two of these were for persistent shoulder pain after a fall, one in each group, and a humerus fracture after a fall in the debridement with device group. Three serious adverse events were considered related to the surgery: two, one in each group, for persistent pain or disability at 12 months (defined as requiring ongoing secondary care review), and one for further surgery, a reverse shoulder replacement in the debridement with device group.

In the prespecified subgroup analyses, sex was found to have a significant interaction with the 12-month Oxford Shoulder Score (adjusted model interaction term –9·5 [95% CI –16·5 to –2·6]; p=0·0099; appendix p 8). For men, the unadjusted mean difference was 0·7 (95% CI –4·7 to 6·1). For women, the unadjusted mean difference was –10·9 (–16·7 to –5·1). Interaction terms for age group (–5·3 [95% CI –12·8 to 2·3]; p=0·18; appendix p 9) and tear size (6·8 [–14·9 to 28·8]; p=0·55; appendix p 9) were not significant. Due to the limited number of less than 3 cm tear size counts within the two intervention groups, we did an additional post-hoc subgroup analysis with the anteroposterior tear size as a continuous variable. No interaction was demonstrated (1·2 points [95% CI –1·7 to 4·0]; p=0·40). Arthroscopic pictures were judged adequate in 111 (95%) of 117 cases, images were unavailable for three participants and inadequate for three, no technical errors were identified. At 12 months, 73 of 114 (64%) participants did not know which group they were in, one (1%) did not answer, 40 (35%) believed they knew the allocation. Of these, 26 (65%) were correct and 14 (35%) were incorrect (appendix p 16).

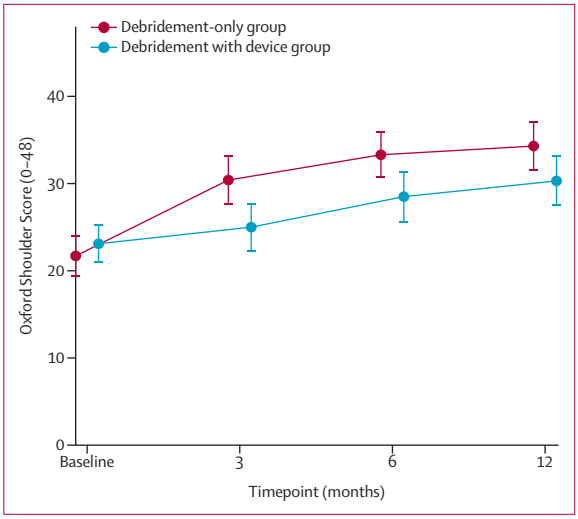

Figure 2: **Oxford Shoulder Score means and 95% confidence intervals for each timepoint**

	Debridement only (n=61)	Debridement with device (n=56)	Total (n=117)
Participants with any adverse event	9 (15%)	11 (20%)	20 (17%)
Adverse event per participant			
1	7	9	16
≥2	2	2	4
Total number of reported adverse events*	11	17	28
Exacerbation or persistence of shoulder pain or restrictive range of motion	5	6	11
Injection into the shoulder region	1	3	4
Adhesive capsulitis	0	2	2
Persistent muscle soreness or muscle injury	0	1	1
Other	5	4	9
Participants with any serious adverse event*	2 (3%)	4 (7%)	6 (5%)

Data are n (%). *Related serious adverse events included persistent pain or diability and further surgery; unrelated serious adverse events included persistent shoulder pain and humerus fracture.

Table 3: **Adverse events**

Discussion

We used a blinded randomised controlled trial design with predefined stopping boundaries to test whether the InSpace device was of benefit for people with irreparable rotator cuff tears. The study stopped at more than half the maximum potential sample size of 221, allowing us to report the findings early. In the primary analysis, arthroscopic debridement only was found to be superior to arthroscopic debridement with the InSpace device for

people with an irreparable rotator cuff tear of the shoulder, based on the Oxford Shoulder Score 12 months after surgery. Secondary outcomes were in agreement with the primary outcome, effectively excluding the possibility of any meaningful benefit for the InSpace device. The treatments differed only in the use of the device and blinding was largely maintained; therefore, any difference was most likely due to the device.

The effect of the intervention was moderated by sex, with worse results for the device in females, the reason for this is unknown. Possible explanations include mechanical factors such as size or strength, biological factors such as the host response to the balloon material, or it could be a chance observation. Even in men, it remains unlikely that the device is of benefit.

Pre-publication data were recently reported in a trial registry from a company-funded trial of the InSpace device (NCT02493660). This reported non-inferiority of the balloon device compared with partial cuff repair using a composite primary outcome. The primary composite outcome required participants to reach a certain level of recovery by 6 weeks and maintain it to 12 months, although many people undergoing partial cuff repair would still be in the early phases of recovery by 6 weeks.[27,28] Non-composite outcomes did not appear to show benefit over partial cuff repair. Partial cuff repair is more commonly used in North America than in Europe, and evidence supporting its use is limited to case series data, with variable outcomes reported.[7,27,28] Also, the study did not test the effect of the balloon itself over an otherwise identical procedure; therefore, this company-funded trial cannot directly answer the question whether the balloon is effective.

The InSpace device was given Conformité Européenne marking, entered clinical practice, and has been widely used on the basis of basic science and small case-series data.[8,10]The first pre-publication trial data were only uploaded in 2021. FDA approval was granted more recently, although the trial data has not yet been published in peer-review format (NCT02493660).

We have not identified the underlying mechanism responsible for our findings, although there was uncertainty in the literature about how the device would provide long-term benefit, particularly as the balloon deflates over time, typically a few months after implantation.[10,11] Case-series data have mostly demonstrated improvements in outcome from baseline to final follow-up, although a small number of case series have demonstrated less satisfactory results.[10] Both treatments delivered in our trial demonstrated improvement compared with baseline, which could be incorrectly interpreted as beneficial in a case-series alone.[19] Our previous meta-analysis demonstrated improvements in outcome compared with baseline for all types of treatment for rotator cuff tears; therefore, case series data is not able to provide meaningful information on the benefit of a new treatment, such information can only be achieved in a randomised trial.[19]

These findings demonstrate the crucial importance of early randomised trials in evaluating new technologies, and raises questions about the early introduction of technologies without robust randomised trial evidence. It demonstrates the importance of high-quality evidence, for which robust studies are needed to provide clear evidence of benefit where financial resource is limited.

People with symptomatic tears of the rotator cuff that are irreparable have few proven treatment options, despite the pain and loss of function that they often have. There is little evidence for surgical treatment, and many untested technologies are in use.[7,8] In the absence of effective solutions, people with symptomatic tears of the rotator cuff might resort to reverse shoulder replacement, a major procedure with risks of disabling complications.[7] Trials are needed for potential solutions, including arthroscopic debridement and more complex procedures, to ensure patients can be offered effective and proven treatments for this challenging condition.[8]

The adaptive design methodology used in this trial provides a template that can be used more widely and would be valuable in small populations or where recruitment is challenging, both of which are common problems when testing new surgical procedures. A prespecified adaptive design, using the correlations between data from early and late timepoints, increased the efficiency of the interim analyses, which allowed the study to report much sooner.[17,18] At our pre-pandemic rate of recruitment, the study would have taken an additional 11 months to recruit the full (maximum) sample size of 221. This would have been much longer due to COVID-19 restrictions, exposing many more people to ineffective treatment or potential harm.

By setting interim futility and efficacy stopping rules, the study was designed to report early either for convincing evidence of a lack of benefit (futility) or for strong clinical benefit (efficacy). Although the sample size appears small for a definitive trial, the study was designed to have 90% power and the adaptive design functioned well, providing a robust answer to the clinical question. Traditional sample size estimates are often based on assumptions about the data that might not be observed when the study is conducted, and although adaptive designs are becoming widely accepted in other fields, they remain uncommon in surgery. There is little benefit in continuing to recruit to achieve greater precision if the study is likely to be conclusive with a smaller sample size. This exposes fewer people to risk of harm in the trial itself, as well as preventing harm or high costs for people having the treatment in the wider community.[16] Having both futility and efficacy stopping rules allow a study to report early if an intervention is ineffective, but it potentially also allows early implementation of effective interventions, similarly benefitting patients.[16]

Our study has limitations. The study was not powered for subgroup analyses, either at the achieved or the maximum sample size. The findings of the subgroup

analysis should therefore be interpreted with considerable caution. We based our exclusion criteria on current clinical use and manufacturer's recommendations. The device could be of benefit in a different population, although we are not aware of it being widely used for other purposes. Our predefined primary outcome was at 12-months, and while data collection is ongoing for 24-month outcomes, our previous systematic review of randomised trials of rotator cuff tears demonstrated high correlation between study findings at 12 and 24 months.[19] As a result of COVID-19 restrictions, we were not able to complete data collection for objective measures, although the objective measures we did take in the study correlate well with the Oxford Shoulder Score (appendix p 13).[25]

We are not able to draw direct conclusions about arthroscopic debridement for people with an irreparable rotator cuff tear. It is a commonly used treatment in UK practice that allowed us to isolate the effects of the InSpace device in a blinded surgical trial. There remains uncertainty among experts about benefit compared with non-surgical care and this could be a focus for future trials.[7]

In this study, we implemented an efficient, adaptive, blinded multicentre randomised trial, which found that arthroscopic debridement was superior to the same procedure performed with the InSpace device for people with irreparable rotator cuff tears. The InSpace device is unlikely to be of benefit and might be harmful; therefore, we do not recommend use of the device in this population.

Contributors

AM led the study and was involved in conception, design, management, analysis, and prepared the final draft. HP was lead trial statistician and was involved in conception, design, management, analysis, and wrote sections of the final draft. NP led the adaptive design aspects and was involved in conception, design, management, analysis, and edited the final draft. JF was the patient lead and was involved in design, management, analysis, and edited the final draft. JB was the Senior Project Manager and was involved in design, management, analysis, and edited the final draft. EGM was the trial manager and was involved in design, management, analysis, and edited the final draft. AH was a study statistician and was involved in the management, analysis, and edited the final draft. CH led the radiological interpretation and was involved in design, management, analysis, and edited the final draft. RK was physiotherapy lead and was involved in conception, design, management, analysis, and edited the final draft. IK was a health economist and was involved in design, management, analysis, and edited the final draft. TL was a shoulder surgeon and led fidelity assessments, and he was involved in conception, design, management, analysis, and edited the final draft. JM was health economic lead, he was involved in conception, design, management, analysis, and edited the final draft. NS was a statistician and provided expertise and oversight for the adaptive design, he was involved in conception, design, management, analysis, and edited the final draft. MU was the senior trials methodologist he was involved in conception, design, management, analysis, and edited the final draft. SD was the senior shoulder specialist, he was involved in conception, design, management, analysis, and edited the final draft. AH and HP directly accessed and verified the underlying data reported in the manuscript. The decision to submit was made by the chief investigator (AM) and senior authors (MU and SD) in conjunction with all of the authors. All authors had full access to the data.

Declaration of interests

Outside of this study, the authors report no personal financial conflict of interest with Stryker or any other related commercial organisation. AM, HP, EGM, CH, JM, and MU are coinvestigators on two other National Institute for Health and Care Research (NIHR)-funded trials (RACER-Knee and RACER-Hip, AM 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. AM, HP, NP, EGM, AH, CH, RK, IK, JM, NS, and MU all work on other NIHR-funded studies. CH, RK, JM and MU are or have been members of funding panels in NIHR, although not on the Efficacy and Mechanism Evaluation programme. RSK is chair of the NIHR Research for Patient Benefit board, a paid position in NIHR but unrelated to the trial. MU was until March, 2021, 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. SD Held an Educational Consultancy Contract with Wright for the past 36 months for lecturing on the Aequalis Shoulder Re-placement System. Wright was acquired by Stryker in 2020.

Data sharing

Deidentified individual participant data that underlie the results reported in this article will be available 6 months after the time of publication, with no end date, to researchers who provide a methodologically sound proposal, agreed between the researchers and Warwick Clinical Trials Unit. Proposals should be directed to WCTUDataAccess@warwick.ac.uk or a.metcalfe@warwick.ac.uk. To gain access, data requestors will need to sign a data access agreement. The study protocol, statistical analysis plan, participant information sheets, surgical and rehabilitation manuals, and constant score manual will be available online.

Acknowledgments

This project (funders reference 16/61/18) is funded by the Efficacy and Mechanism Evaluation Programme, a Medical Research Council (MRC) and National Institute for Health and Care Research (NIHR) partnership. The views expressed in this publication are those of the authors and not necessarily those of the MRC, NIHR, or the Department of Health and Social Care. To reduce treatment costs to trial sites, 23 InSpace devices were provided for free under an agreement with the manufacturers of the balloon (initially Ortho-Space, Israel and later Stryker, USA). Ortho-Space also funded some of the training centre costs of a study-related cadaveric course to train participating surgeons in the use of the balloon at the start of the study. The full independence of the trial team is protected by legal agreements. The trial was co-sponsored by the University Hospitals Coventry and Warwickshire NHS Trust and the University of Warwick.

References

1 Linsell L, Dawson J, Zondervan K, 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.

2 Largacha M, Parsons IM 4th, Campbell B, Titelman RM, Smith KL, Matsen F 3rd. Deficits in shoulder function and general health associated with sixteen common shoulder diagnoses: a study of 2674 patients. *J Shoulder Elbow Surg* 2006; **15:** 30–39.

3 Minns Lowe CJ, Moser J, Barker K. Living with a symptomatic rotator cuff tear 'bad days, bad nights': a qualitative study. *BMC Musculoskelet Disord* 2014; **15:** 228.

4 Carr AJ, Cooper CD, Campbell MK, et al. Clinical effectiveness and cost-effectiveness of open and arthroscopic rotator cuff repair [the UK Rotator Cuff Surgery (UKUFF) randomised trial]. *Health Technol Assess* 2015; **19:** 1–218.

5 Parikh N, Martinez DJ, Winer I, Costa L, Dua D, Trueman P. Direct and indirect economic burden associated with rotator cuff tears and repairs in the US. *Curr Med Res Opin* 2021; **37:** 1199–211.

6 Monrad N, Ganestam A, Kallemose T, Barfod KW. Alarming increase in the registration of degenerative rotator cuff-related lesions a nationwide epidemiological study investigating 244 519 patients. *Knee Surg Sports Traumatol Arthrosc* 2018; **26:** 188–94.

7 Cvetanovich GL, Waterman BR, Verma NN, Romeo AA. Management of the irreparable rotator cuff tear. *J Am Acad Orthop Surg* 2019; **27:** 909–17.

8 Kooistra B, Gurnani N, Weening A, van den Bekerom M, van Deurzen D. Low level of evidence for all treatment modalities for irreparable posterosuperior rotator cuff tears. *Knee Surg Sports Traumatol Arthrosc* 2019; **27:** 4038–48.

9 Savarese E, Romeo R. New solution for massive, irreparable rotator cuff tears: the subacromial "biodegradable spacer". *Arthrosc Tech* 2012; **1:** e69–74.

10 Johns WL, Ailaney N, Lacy K, Golladay GJ, Vanderbeck J, Kalore NV. Implantable subacromial balloon spacers in patients with massive irreparable rotator cuff tears: a systematic review of clinical, biomechanical, and financial implications. *Arthrosc Sports Med Rehabil* 2020; **2:** e855–72.

11 Senekovic V, Poberaj B, Kovacic L, et al. The biodegradable spacer as a novel treatment modality for massive rotator cuff tears: a prospective study with 5-year follow-up. *Arch Orthop Trauma Surg* 2017; **137:** 95–103.

12 National Institute for Health and Care Excellence (NICE). Biodegradable subacromial spacer insertion for rotator cuff tears. 2016. https://www.nice.org.uk/guidance/ipg558 (accessed April 12, 2022).

13 Stryker. Stryker announces the FDA clearance of the first biodegradable subacromial balloon spacer, filling a gap in the shoulder continuum of care. 2021. https://www.stryker.com/us/en/about/news/2021/stryker-announces-the-fda-clearance-of-the-first-biodegradable-s.html#:~:text=Advanced%20Digital%20Healthcare-,Stryker%20announces%20the%20FDA%20clearance%20of%20the%20first%20biodegradable%20subacromial,the%20shoulder%20continuum%20of%20care. (accessed March 25, 2022).

14 Sedrakyan A, Campbell B, Merino JG, Kuntz R, Hirst A, McCulloch P. IDEAL-D: a rational framework for evaluating and regulating the use of medical devices. *BMJ* 2016; **353:** i2372.

15 Barkun JS, Aronson JK, Feldman LS, Maddern GJ, Strasberg SM. Evaluation and stages of surgical innovations. *Lancet* 2009; **374:** 1089–96.

16 Bhatt DL, Mehta C. Adaptive designs for clinical trials. *N Engl J Med* 2016; **375:** 65–74.

17 Stallard N. A confirmatory seamless phase II/III clinical trial design incorporating short-term endpoint information. *Stat Med* 2010; **29:** 959–71.

18 Parsons N, Stallard N, Parsons H, et al. 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 tears. *Trials* 2019; **20:** 694.

19 Khatri C, Ahmed I, Parsons H, et al. The natural history of full-thickness rotator cuff tears in randomized controlled trials: a systematic review and meta-analysis. *Am J Sports Med* 2019; **47:** 1734–43.

20 Metcalfe A, Gemperle Mannion E, Parsons H, et al. Protocol for a randomised controlled trial of subacromial spacer for tears affecting rotator cuff tendons: a randomised, efficient, adaptive clinical trial in surgery (START:REACTS). *BMJ Open* 2020; **10:** e036829.

21 Dawson J, Rogers K, Fitzpatrick R, Carr A. The Oxford shoulder score revisited. *Arch Orthop Trauma Surg* 2009; **129:** 119–23.

22 Constant CR, Gerber C, Emery RJH, Søjbjerg JO, Gohlke F, Boileau P. A review of the Constant score: modifications and guidelines for its use. *J Shoulder Elbow Surg* 2008; **17:** 355–61.

23 Jones IA, Togashi R, Heckmann N, Vangsness CT Jr. Minimal clinically important difference (MCID) for patient-reported shoulder outcomes. *J Shoulder Elbow Surg* 2020; **29:** 1484–92.

24 Beard DJ, Rees JL, Cook JA, et al. Arthroscopic subacromial decompression for subacromial shoulder pain (CSAW): a multicentre, pragmatic, parallel group, placebo-controlled, three-group, randomised surgical trial. *Lancet* 2018; **391**: 329–38.

25 Christiansen DHH, Frost P, Falla D, Haahr JP, Frich LH, Svendsen SW. Responsiveness and minimal clinically important change: a comparison between 2 shoulder outcome measures. *J Orthop Sports Phys Ther* 2015; **45:** 620–25.

26 Todd S, Whitehead J, Facey KM. Point and interval estimation following a sequential clinical trial. *Biometrika* 1996; **83:** 453–61.

27 Shon MS, Koh KH, Lim TK, Kim WJ, Kim KC, Yoo JC. Arthroscopic partial repair of irreparable rotator cuff tears: preoperative factors associated with outcome deterioration over 2 years. *Am J Sports Med* 2015; **43:** 1965–75.

28 Cuff DJ, Pupello DR, Santoni BG. Partial rotator cuff repair and biceps tenotomy for the treatment of patients with massive cuff tears and retained overhead elevation: midterm outcomes with a minimum 5 years of follow-up. *J Shoulder Elbow Surg* 2016; **25:** 1803–09.