BMJ Open

Finnish Subacromial Impingement Arthroscopy Controlled Trial (FIMPACT): a protocol for a randomised trial comparing arthroscopic subacromial decompression and diagnostic arthroscopy (placebo control), with an exercise therapy control, in the treatment of shoulder impingement syndrome

Mika Paavola,[1] Antti Malmivaara,[2] Simo Taimela,[1] Kari Kanto,[3] Teppo LN Järvinen,[1] on behalf of the FIMPACT Investigators

To cite: Paavola M, Malmivaara A, Taimela S, *et al*. Finnish Subacromial Impingement Arthroscopy Controlled Trial (FIMPACT): a protocol for a randomised trial comparing arthroscopic subacromial decompression and diagnostic arthroscopy (placebo control), with an exercise therapy control, in the treatment of shoulder impingement syndrome. *BMJ Open* 2017;**7**:e014087. doi:10.1136/bmjopen-2016-014087

► Prepublication history for this paper is available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2016-014087).

Received 30 August 2016
Revised 13 February 2017
Accepted 9 March 2017

For numbered affiliations see end of article.

Correspondence to
Dr. Mika Paavola;
mika.paavola@hus.fi

ABSTRACT

Introduction Arthroscopic subacromial decompression (ASD) is the most commonly performed surgical intervention for shoulder pain, yet evidence on its efficacy is limited. The rationale for the surgery rests on the tenet that symptom relief is achieved through decompression of the rotator cuff tendon passage. The primary objective of this superiority trial is to compare the efficacy of ASD versus diagnostic arthroscopy (DA) in patients with shoulder impingement syndrome (SIS), where DA differs only by the lack of subacromial decompression. A third group of supervised progressive exercise therapy (ET) will allow for pragmatic assessment of the relative benefits of surgical versus non-operative treatment strategies.

Methods and Analysis Finnish Subacromial Impingement Arthroscopy Controlled Trial is an ongoing multicentre, three-group randomised controlled study. We performed two-fold concealed allocation, first by randomising patients to surgical (ASD or DA) or conservative (ET) treatment in 2:1 ratio and then those allocated to surgery further to ASD or DA in 1:1 ratio. Our two primary outcomes are pain at rest and at arm activity, assessed using visual analogue scale (VAS). We will quantify the treatment effect as the difference between the groups in the change in the VAS scales with the associated 95% CI at 24 months. Our secondary outcomes are functional assessment (Constant score and Simple shoulder test), quality of life (15D and SF-36), patient satisfaction, proportions of responders and non-responders, reoperations/treatment conversions, all at 2 years post-randomisation, as well as adverse effects and complications. We recruited a total of 210 patients from three tertiary referral centres. We will conduct the primary analysis on the intention-to-treat basis.

Ethics and Dissemination The study was approved by the Institutional Review Board of the Pirkanmaa Hospital District and duly registered at ClinicalTrials.gov. The findings of this study will be disseminated widely through peer-reviewed publications and conference presentations.

Trial registration number NCT00428870; Pre-results.

Strengths and limitations of this study

- Efficacy design: Strict eligibility criteria
- Placebo-surgery controlled trial: Blinding of both the participants and the outcome assessors in the comparison between index surgery and control (placebo surgery)
- Inclusion of a non-surgical treatment option to allow a pragmatic assessment of the relative benefits of surgical versus non-operative treatment strategies
- Potential confounding due to participants' knowledge of the treatment delivered in our secondary comparison between surgical and non-operative treatments

INTRODUCTION

Subacromial decompression is one of the most frequently performed procedures in orthopaedics.[1 2] It is carried out to treat patients with shoulder pain attributed to 'subacromial impingement syndrome' (SIS). Conventional wisdom dictates that SIS is caused by' impingement' of the rotator cuff (RC) between the humeral head and the

overlying acromion while lifting the arm. The appropriateness of this mechanistic explanation has been challenged lately where the generic label of 'subacromial pain syndrome' is currently advocated.[3 4] The aim of the subacromial decompression procedure, typically carried out arthroscopically, is to decompress the RC tendon passage through the subacromial space through resection and smoothening of the hypertrophied or prominent anterolateral undersurface of the acromion. Management of shoulder pain has been estimated to account for 4.5 million visits annually to physicians in the USA alone,[5] accounting for US$3 billion in costs each year.[6] Since 44%–65% of all shoulder complains are related to SIS, it is estimated that annual direct medical costs of SIS are over $1 billion in the USA.[7 8]

Since the introduction of subacromial decompression surgery in the early 1970s[9], the number of procedures has steadily increased across the entire western world. With the advent of arthroscopy, the number of these surgeries has increased dramatically—fivefold from the 1980s to 2005 in the USA[10] and 700% between 2000 and 2010 in the UK.[11] Remarkably, there is a stark absence of evidence from high-quality controlled trials to support the existing practice of performing subacromial decompression for patients with SIS. Two recent systematic reviews concluded that subacromial decompression provides no superior benefits in terms of pain relief, function or quality of life compared with non-surgical treatment.[12 13] There is even a placebo controlled trial to show the beneficial effect of exercise therapy (ET) over placebo physiotherapy.[14] However, the proponents of the procedure have argued that the evidence is skewed in favour of the therapeutic potential of surgery due to a significant crossover (5%–15%) from conservative treatment to surgery.[14–16] Although such concern is obviously warranted, it should also be recalled that surgeons' own perceptions on the success of any surgery might similarly be biased due to a considerable surgical placebo effect.

The outcome of any medical (surgical) intervention—particularly when treating primarily subjective symptoms—is a cumulative effect of three main elements: placebo effects, critical therapeutic (surgical) element and non-specific effects, most importantly, the normal variation in the course of the disease and the regression-to-the-mean phenomenon.[17 18] Conceding that the act of surgery per se produces a profound placebo response, a 'true' treatment effect is impossible to disentangle from the non-specific (placebo) effects—such as the patients' or researchers' expectations of benefit—without a placebo comparison group.[19] The critical therapeutic element is the component of the surgical procedure that is believed to provide the therapeutic effects (here, subacromial decompression), which are distinct from aspects of the procedures that are diagnostic or required to access the disease being treated (here, shoulder arthroscopy).

To the best of our knowledge, there is only one other ongoing study aiming to assess the true efficacy of subacromial decompression surgery in patients with SIS using a placebo controlled study design. According to the published protocol of this CSAW (Can Shoulder Arthroscopy Work?) trial,[20] the investigators have chosen a very similar approach to that of our Finnish Subacromial Impingement Arthroscopy Controlled Trial (FIMPACT). In brief, the CSAW trial is a three-group pragmatic randomised controlled trial comparing arthroscopic acromioplasty, active monitoring with specialist reassessment and investigational shoulder arthroscopy only. CSAW aims for recruitment of 300 patients with SIS to assess the efficacy of the surgery against no surgery, the need for a specific component of the surgery (acromioplasty) and the quantification of the possible placebo effect. As readily apparent, the two trials (FIMPACT vs CSAW) are very similar in design with the only notable differences being the primary outcome measure (pain at rest and after activity vs Oxford Shoulder Score, a score that assesses both pain and activities of daily living impairment), the primary outcome assessment point (24 months vs 6 months) and the intervention delivered for the third group (ET vs active monitoring with specialist reassessment), respectively.

The primary hypothesis of our FIMPACT trial is that arthroscopic subacromial decompression (ASD) is superior to diagnostic arthroscopy (DA) in patients with SIS. In addition, we will perform a pragmatic comparison of surgical and non-surgical treatment options (ASD vs ET). The relative benefits of ASD and ET will be assessed without a priori hypothesis of the superiority of one or the other.

MATERIALS AND METHODS

Overview of study design

The FIMPACT trial is an ongoing multicentre, three-group randomised controlled superiority study with a primary objective to assess the efficacy of ASD versus DA in patients diagnosed with SIS. Our design also enables the pragmatic comparison of surgical and non-surgical treatment strategies (ASD vs ET) (figure 1). To obtain three balanced study groups (of similar group size), we performed a twofold, sequential randomisation as follows: First, we randomised patients to surgical or conservative treatment in a 2:1 ratio and then randomised those allocated to surgery to ASD or DA in a 1:1 ratio. The initial patient screening for the trial began at one site (Tampere) on 1 February 2005 and was then expanded to two additional tertiary referral centres in March 2006 and December 2006 to improve recruitment and overall generalisability of the results. The recruitment was completed (all 210 required patients enrolled) in August 2013.

Ethical approval

Ethical approval was obtained on 28 December 2004 from the institutional review board of the Pirkanmaa Hospital District (R04200). Local research and development approvals were gained for each recruiting centre.

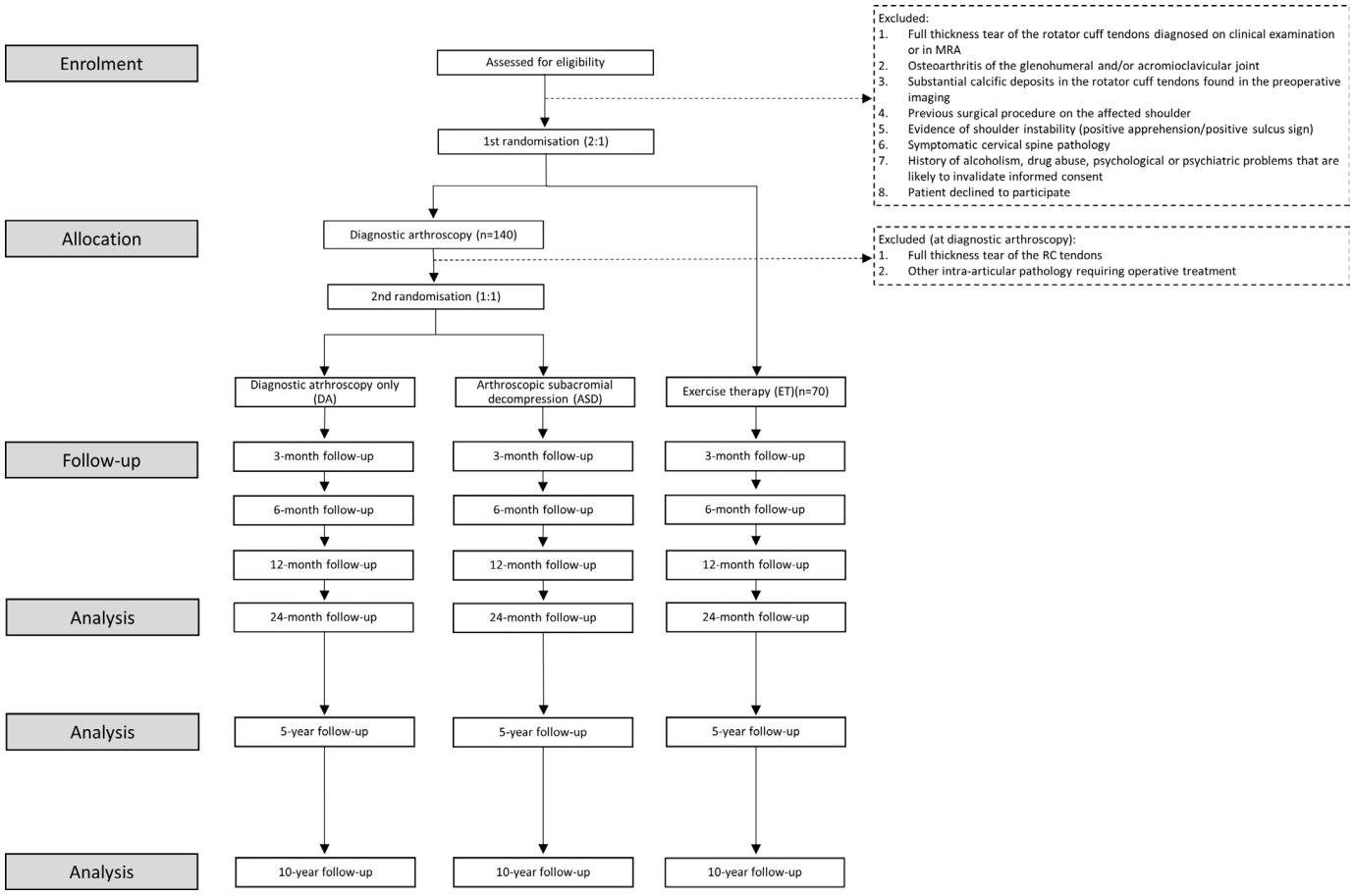


Figure 1 Flow chart of the trial: enrolment, assigned intervention and follow-up scheme. MRA, MRI with intra-articular contrast; RC, rotator cuff.

Participant selection

We assessed for eligibility all patients complaining of subacromial shoulder pain to any of the participating clinics. These participants were screened according to the inclusion and exclusion criteria and a recruitment surgeon confirmed the clinical diagnosis of SIS. To qualify as a recruitment surgeon, all trial surgeons had to have experience of more than 500 shoulder arthroscopies before the start of the trial. Detailed clinical examination of the shoulder was performed on all referred patients to rule out possible instability, clinical signs of RC rupture, frozen shoulder or other causes of symptoms. Standard X-rays and MRI were obtained from all potential participants and assessed by both a musculoskeletal radiologist and an orthopaedic surgeon. For patients found eligible for this study (fulfilling indications for ASD), we obtained written informed consent and randomised them into non-operative or operative groups (1:2) immediately after the baseline appointment. If patient had bilateral symptoms, only one shoulder was included in the study.

Eligibility criteria

We used specific eligibility criteria to ensure that recruited participants were only those with SIS. Accordingly, a standardised clinical examination was first performed, followed by a subacromial injection test. To exclude patients with concomitant pathology, particularly RC rupture, standard X-rays and MRI with intra-articular contrast injection (MRA) were carried out on all potential participants.

Inclusion criteria

1. Adult men or women ages 35 years to 65 years
2. Subacromial pain for greater than 3 months with no relief from non-operative means (physiotherapy, non-steroidal anti-inflammatory medication, corticosteroid injections and rest)
3. Pain provoked by abduction and positive painful arc sign
4. Positive impingement test (temporary relief of pain by subacromial injection of lidocaine)
5. Pain in at least two out of three of isometric tests (abduction 0° and 30° or external rotation)
6. Provision of informed consent from the participant
7. Ability to speak, understand and read in the language of the clinical site

Exclusion criteria

1. Full thickness tear of the RC tendons diagnosed on clinical examination (marked weakness in any of the examined muscles) or MRA

2. Osteoarthritis of the glenohumeral and/or acromioclavicular joint diagnosed on clinical examination and on X-rays
3. Substantial calcific deposits in the RC tendons found in the preoperative imaging
4. Previous surgical procedure on the affected shoulder
5. Evidence of shoulder instability (positive apprehension/positive sulcus sign)
6. Symptomatic cervical spine pathology
7. History of alcoholism, drug abuse, psychological or psychiatric problems that are likely to invalidate informed consent
8. Patient declined to participate

Recruitment process

Consultant orthopaedic surgeons carried out eligibility screening among patients referred to the study centres through standard clinical practice for shoulder pain. Patients meeting the eligibility criteria were introduced to the study. If patients expressed interest in participating, written information about the study was provided and they were asked to opt in. If the interest continued, arrangements were made for obtaining required imaging (X-rays and MRA) and for a separate baseline appointment.

Informed consent

At the first appointment, all participants were introduced to the detailed written information about the study and asked to sign a written informed consent form provided in the online supplementary appendix. At the baseline appointment (arranged within 45 days of initial contact), baseline data were completed and participant's willingness to participate in the study was confirmed. This procedure ensured that all potential participants had a reflection period for consent of at least 48 hours before giving their final consent to participate. Particular attention was paid to ensure that the participants realised that on entering the study they may receive only DA, in which case the subacromial decompression would not be performed. They were also informed that participation in the study is entirely voluntary and any decision they make would not affect their possible future care. In addition, every participant was informed of their right to withdraw from the trial whenever they desire without the need to supply any reason for such decision.

Baseline assessment

Baseline assessment included documentation of gender, birth date, education, employment, hand dominance, time from the onset of symptoms, recreational habits and employment status. We asked participants to assess their general heath and usage of pain medication. Modalities of any prior conservative treatment were also recorded (table 1).

Baseline clinical symptoms

The recruiting surgeon carried out a clinical history and a clinical examination related to shoulder pain. Shoulder complaints other than SIS, such as full-thickness RC tears, frozen shoulder, osteoarthritis of the acromioclavicular joint and instability were ruled out as much as clinical diagnosis allows.

Baseline imaging

Standard X-rays of the shoulder were obtained to assess possible glenohumeral or acromioclavicular osteoarthritis. An MRA was also obtained to rule out any other intra-articular or extra-articular pathologies. A musculoskeletal radiologist and an orthopaedic surgeon assessed all the images.

Randomisation and concealment

We used a two-phase sequential randomisation. In phase I, the participants were randomised into non-surgical or surgical treatment with allocation ratio 1:2. In phase II, those allocated to surgical treatment were further randomised to ASD or DA with 1:1 ratio (figure 1).

An independent statistician with no involvement in the execution of the trial prepared separate randomisation lists for each study centre using a computer-generated algorithm. Randomisation was carried out using sequentially numbered sealed opaque envelopes. The envelopes were kept in a secure, agreed location at each centre. To ensure concealment, block randomisation was applied using blocks varying in size randomly, the block size known only by the statistician.

To initially enter a participant into the study (phase I), an envelope containing the treatment assignment (non-surgical (ET) or surgery (ASD or DA), ratio 1:2) was opened during the baseline appointment. Participants randomised to ET started standardised physiotherapy within 2 weeks of the baseline appointment. Participants allocated to surgical treatment were scheduled for surgery aimed to be completed within 12 weeks of randomisation.

At the day of surgery, a DA was first carried out to confirm the eligibility of the participant (to rule out full-thickness RC tear and other obvious intra-articular pathology). Research/staff nurse then completed the randomisation procedure (phase II) by opening an envelope containing the surgical treatment allocation (ASD or DA, ratio 1:1). The allocation was revealed to the surgeon by showing the paper, but not expressed verbally.

Interventions

Diagnostic arthroscopy

All participants in the two operative groups first underwent arthroscopic examination of the shoulder with the use of standard posterior and lateral portals and a 4 mm arthroscope. To maintain concealment, the surgery was carried out under general anaesthesia. The orthopaedic surgeon evaluated and graded possible intra-articular pathological changes. The RC integrity was also evaluated from the subacromial space without performing routine bursectomy. If the integrity of the RC could not be assessed, bursal tissue was bluntly stretched with troachar or resected on the tendon side to allow visualisation.

Table 1 Baseline characteristics

	ASD	DA	ET
Age (years), mean (SD)			
Gender (female/male), n (%)			
Dominant hand affected, n (%)			
Socioeconomic status/workload			
Heavy manual labour (construction work, etc), n (%)			
Heavy manual labour (variable workload), n (%)			
Mostly manual labour including daily office work, n (%)			
Mostly office work with occasional manual assignments, n (%)			
Full-time office work, n (%)			
Unemployed, n (%)			
Pensioner/disability pensioner, n (%)			
Student, n (%)			
Home maker/housewife/other, n (%)			
Subjective health			
Duration of symptoms (months), mean (SD)			
Ability to work normally regardless of the shoulder symptoms? (yes/no), n (%)			
Recreational ability regardless of the shoulder symptoms? (yes/no), n (%)			
Prior treatments			
Rest, n (%)			
Pain medication, n (%)			
Topical pain medication, n (%)			
Corticosteroid injection, n (%)			
Ultrasound, laser or any other similar therapies, n (%)			
Physiotherapy including ET, n (%)			
Other, n (%)			
Generic health states			
15D			
SF-36			
Pain measurements/shoulder scores			
Pain at rest (0–100 VAS Scale), mean (SD)			
Pain during activity (0–100 VAS Scale), mean (SD)			
Constant-Murley Score (CS), mean (SD)			
The simple shoulder test (SST), mean (SD)			

ASD, arthroscopic subacromial decompression; DA, diagnostic arthroscopy; ET, exercise therapy; SF-36, Short Form 36; VAS, Visual Analogue Scale.

If arthroscopic examination revealed any unexpected pathology (such as capsular pathology, full-thickness RC tear or osteoarthritis), the patient was treated according to current clinical practice guidelines for the given pathology while under the same anaesthesia. In such a case, the participant was excluded from the trial. Patients with partial tears were included in the study, while patients with a full-thickness tear were excluded and RC repair was carried out.

After the arthroscopic examination of the glenohumeral joint and subacromial space, confirming the eligibility of the participant, the participants were randomly assigned to receive either ASD or DA only. If the patient was allocated to the DA group, the operation was terminated. To ensure concealment of the participants and the staff other than those in the operating theatre, the participants were kept in the operating theatre for the required time to perform subacromial decompression.

Arthroscopic subacromial decompression

Debridement of the subacromial bursa was performed with a shaver and/or electrocoagulation, followed by the resection of the bony spurs and projecting anterolateral

undersurface of the acromion by a shaver as described by Ellman.[21]

Postoperative care

In both the ASD and the DA groups, the postoperative rehabilitation was identical. All surgically treated participants received one visit to an independent physiotherapist for guidance and instructions for home exercises. Subsequent rehabilitation was carried out according to the standardised rehabilitation protocols of the participant centres. Since the initial rehabilitation after a surgery needs to be 'tempered' due to joint irritation, the rehabilitation protocol of the operatively treated groups (ASD and DA) was not identical to the ET group.

Exercise therapy

In the ET group, supervised progressive physiotherapy was started within 2 weeks of randomisation using a standardised protocol. The protocol was based on the same principles as the regimen shown effective for the treatment of SIS earlier,[14] but was updated—with the help of the principal investigator of the original study[14]—to conform with the state-of-the-art ET for SIS. The regimen was based on daily home exercises, and included 15 visits to an independent physiotherapist for guidance and monitoring of the progress, carried out approximately once a week. The aim of the supervised exercise treatment was to restore painless, normal mobility of the shoulder girdle, eliminate any capsular tightness and to increase the dynamic stability of the glenohumeral joint and the scapula.

Compliance to treatment allocation and possible crossover

Participants allocated to the ET group were told at the time of giving consent that they would be allowed to consider crossing over to the ASD group if they didn't get adequate relief of symptoms (preferably no sooner than 6 months post randomisation). Similarly, in the two surgical treatment groups, the participants were informed of the possibility of unblinding if debilitating symptoms persisted 6 months or more after operation. If the participant was allocated to the DA group, ASD was offered. If the participant had undergone ASD, he/she was offered extended physiotherapy. No prespecified criteria were used for determining 'inadequate relief of symptoms/debilitating symptoms', rather it was left to the participants and the study physicians to make the clinical judgement together.

Outcome measures

The outcomes used in this study and the timetable for follow-up assessments are summarised in table 2.

Primary outcome measure

Visual Analogue Scale

As the primary outcome measure, we used a Visual Analogue Scale (VAS) to measure the patient's perceived pain intensity at rest and at arm activity during the 24 hours preceding the assessment. Shoulder pain was assessed on a 100 mm scale ranging from 0 (no pain) to 100 (extreme pain). We considered 15 as the minimal clinically important difference (MCID) for VAS.[22]

Secondary outcome measures

Constant-Murley Score

The Constant-Murley Score (CS) is the most commonly used scoring system for evaluation of various disorders of the shoulder.[23] It consists of both objective (range of motion and strength) and subjective measurements (pain assessment, workload and leisure time activities), which are summarised in a score between 0 and 100. A higher score indicates better shoulder function. The minimal detectable change of the Constant Score is 17 for patients with SIS.[24]

In addition, as night pain is considered one of the hallmark symptoms in patients with SIS and our two primary outcome measures (patient's perceived pain intensity at rest and at arm activity in the last 24 hours) do not specifically address this issue, a specific question from the CS (unaffected sleep: 'Yes' or 'No') will be analysed separately.

Simple shoulder test

The simple shoulder test (SST) was developed to assess any impairment of the patient's activities of daily living.[25] The SST consists of 12 questions with yes (1) or no (0) response options. The maximum SST score is 12 indicating normal shoulder function, minimum score of 0 points refers severely diminished shoulder function. The SST has good reliability and responsiveness in patients with RC symptoms.[26] The MCID for the SST in RC disease is 2 points.[27]

15D

The 15D instrument (a health-related quality of life instrument with 15 dimensions) is a generic health-related quality of life (HRQoL) instrument comprising 15 dimensions.[28] For each dimension, the respondent must choose one of the five levels that best describes his/her state of health at that moment (the best level being 1 and the worst level being 5). A set of utility or preference weights is used in an addition aggregate formula to generate a single index number, the utility or 15D score. The maximum 15D score is 1 (no problems on any dimension) and the minimum score is 0 (being dead). The responsiveness, reliability and validity of 15D have been thoroughly established, and this instrument has been used extensively in clinical and healthcare research.[29 30]

Short form 36

The short form or SF-36 is a generic HRQoL instrument to quantify the physical, functional and psychological aspects of HRQoL. It consists of 36 questions in eight subscales that assess physical, functional, social and psychological well-being.[31] The score ranges from 0 to 100, where a higher score is associated with better health. The Physical and Mental Component Summary

Table 2 Outcomes and follow-up time points

Assessment	Screening	Enrolment (baseline)	Surgery	3 months	6 months	12 months	24 months	5 years	10 years
Screening form	X								
Informed consent		X							
Baseline characteristics form		X							
X-ray and MRI	X								X
Randomisation		X (first)	X (second)						
Arthroscopic findings form			X						
Follow-up form*				X		X			
Clinical examination		X			X		X	X	X
Complications/adverse effects form†			(X)	(X)	(X)	(X)	(X)	(X)	(X)
VAS, at rest		X		X	X	X	X	X	X
VAS, at arm activity		X		X	X	X	X	X	X
Constant- Murley Score (CS)		X			X		X	X	X
Simple shoulder test (SST)		X			X		X	X	X
SF-36		X		X	X	X	X	X	X
15D		X		X	X	X	X	X	X
Return to work				X	X	X	X	X	X
Return to previous leisure activities				X	X	X	X	X	X
Responder analysis				X	X	X	X	X	X
Patients' satisfaction to the treatment				X	X	X	X	X	X
Patients' assessment of the treatment allocation				X					
Health resource utilisation				X	X	X	X	X	X

*Letter/telephone interview
†If required
SF-36, Short Form 36; VAS, Visual Analogue Scale.

Scales are then calculated as composites of the related subscales. The SF-36 is one of most widely used measures of HRQoL.[32]

Patient satisfaction and responder analysis

We elicited patients' global assessment of satisfaction to the treatment with this question: 'Are you satisfied with the treatment you have received?' We used a VAS scale ranging from 0 (completely disappointed) to 100 (completely satisfied).

Additionally, we elicited patient satisfaction to the treatment outcome with the following question at each follow-up time point (table 2): 'How satisfied are you with the outcome of your treatment?' on a 5-item scale. The response options for this question are provided in the online supplementary appendix. Participants who reported very satisfied or satisfied will be categorised as 'Responders' and patients who responded very dissatisfied or dissatisfied as 'Non-responders'.

Return to previous leisure activities

Similarly, at each follow-up (table 2), participants were asked to respond to the following question: 'Have you been able to return to your previous leisure activities?' ('yes' or 'no').

Patients' perception of operative treatment-group assignment

At the 3-month follow-up point, the patients in the two operative groups were asked to guess whether they had undergone ASD or DA.

Health resource utilisation and costs

For the cost-effectiveness analysis, at each follow-up visit the participants were asked to fill in a questionnaire inquiring about the use of healthcare resources. The questionnaire contains a list of items of healthcare resources available and the participants were asked to fill in the number of visits per item during the recall period of each follow-up time point. The resource use will be calculated based on the number of visits times unit cost per item and expressed as mean costs by items of resource use, and the mean direct total healthcare resource costs. All costs will be discounted to the 2016 price level.

Time to return to work

Information about return to work was recorded at each follow-up time point (table 2).

Complications and adverse effects

Complications directly related to the interventions were registered. The participants were also encouraged to contact the participating hospitals if any adverse effects (AEs) occurred and contacts to the healthcare system were monitored at every follow-up visit. Potential AEs were categorised to serious adverse effects (SAEs) and minor adverse effects (MAEs) if the participants sought treatment. Death, cardiovascular or gastrointestinal effects, deep venous thrombosis, pulmonary embolism, systemic or local infection were categorised as SAEs and shoulder symptoms like pain, swelling and decreased range of motion were categorised as MAEs. The number and severity of complications and AEs will be assessed.

Follow-up

The full follow-up process is shown in figure 1. In brief, the participants filled in the above noted (mailed) outcome questionnaires at 3 months, 6 months, 12 months and 24 months postrandomisation, in addition to which they were also assessed clinically at 6 months and 24 months (and 5 years and 10 years) postrandomisation by a study physiotherapist unaware of treatment allocation, treatment given or possible unblinding. Outcome assessors were instructed not to inquire anything about prior treatment. Further, participants wore a T-shirt on all follow-up examinations.

Adherence and loss to follow-up

Several procedures were implemented to limit loss to follow-up, including excluding individuals likely to pose suboptimal adherence to study follow-up, obtaining verified contact information from each consented participant and having a local research nurse remind participants of upcoming follow-up/clinic visits. All attempts were made to make follow-up as convenient for the patients as possible. Participants were required to visit the outpatient clinic only at 6 months and 24 months (and 5 years and 10 years) postrandomisation, while the 3-month and 12-month follow-ups were carried out using mailed questionnaires to minimise inconvenience to the participants. The follow-up visits had no more discomfort for the participant than the routine clinical shoulder examinations. The follow-up schedule did not involve extra costs to the participants. Follow-up rate was monitored throughout the trial and patients who did not return follow-up questionnaires would receive reminder telephone calls. Using strategies highly similar to these in our previous placebo-surgery controlled trial,[33] a 99% follow-up rate was achieved.

The number and proportion of individuals eligible for and compliant with each follow-up was documented. Individuals who died during the study (from causes unrelated to the study or procedure) will be tabulated. An analysis of the demographic and prognostic characteristics will be carried out between the individuals who withdrew and those who remained in the study. For continuous variables, parametrical or non-parametrical analysis of variance will be used. For categorical variables, χ^2 or Fisher's exact test will be applied.

Missing items

We will use multiple imputation to handle missing data for those statistical analyses that cannot handle occasional missing values. All variables to be included in the final analyses will be included in the chained equations imputation model. The imputation algorithm, a fully conditional specification, uses a specific univariate model for each variable and, for each specific imputed data set,

iteratively imputes each variable with missing values and uses the imputed values in the imputation of other variables.

Sample size

The sample size calculation was based on the two primary outcome measures, VAS at rest and at arm activity, at 24 months postrandomisation. The FIMPACT trial was powered to detect a minimal clinically important improvement (MCII) in a VAS Pain Score (improvement of at least 15; assumed SD 25) between ASD and DA (or ET). To achieve a somewhat unconventional (stringent) 90% study power and using a two-sided type I error rate (5%), our trial requires 68 patients per study group to show clinically meaningful advantage of ASD over DA (or ET). Acknowledging the stringent power threshold, we reserved only 3% surplus for potential loss to follow-up/crossovers (3%), and accordingly, we set the recruitment target at 70 patients per treatment group.

Safety analysis

There are no anticipated safety issues with the FIMPACT Study. Identically to our previous placebo-surgery controlled trial,[33] an interim analysis, as requested by the ethics board, was carried out after the enrolment of 45 participants by an independent data and safety monitoring board (the National Institute for Health and Welfare) to ensure that the rates of complications or reoperations were within acceptable limits (within the normal rate of complications and/or reoperations related to shoulder arthroscopy). Since we found no marked discrepancy in our crude assessment of the incidence of complications/reoperations, no unsealing of group assignments (unblinding) was carried out. No other interim analysis was carried out.

Data management

Questionnaire forms on paper were the primary data collection tools for the study. On receipt of the questionnaire forms, a study nurse made a visual check of the responses and queried missing data when possible. Research assistants, blinded to the group allocation, stored the forms into an electronic database by double data entry to minimise typing errors. The researchers, blinded to the group allocation, perform a visual check of the data in the electronic database and then queried all missing, implausible and inconsistent data. Patient records in the participating hospitals were used when collecting missing data or interpreting inconsistent or implausible data. The final analysis was performed on data transferred to the file 'FIMPACT-full data_final', having been documented as meeting the cleaning and approval requirements of our independent statistician and after the finalisation and approval of the accompanying statistical analysis plan (SAP) document. Participant files will be maintained in storage (both in electronic and paper formats) at the coordinating centre for a period of 10 years after completion of the study (10-year follow-up visits).

STATISTICAL METHODS

Statistical analysis plan

Please refer to the online supplementary appendix for a more complete SAP, which we briefly summarise here. An independent statistician who is unaware of the group assignments will perform all the analyses.

We will summarise the baseline characteristics of the participants by group, reported as a mean (SD) or median (first quartile, third quartile) for continuous variables and count (per cent) for categorical variables.

We will analyse the data in a blinded manner. All p values will be reported to three decimal places with those less than 0.001 reported as p<0.001. The criterion for statistical significance will be set at α=0.05.

Primary analysis

We will carry out the primary analysis according to the intention-to-treat principle: participants are retained in the groups to which they were initially randomised.

The primary comparison will be on the efficacy of ASD (ASD vs DA). We will perform the primary comparison on the efficacy of ASD (ASD vs DA) as a between-group comparison using a repeated measures mixed-effects model (RMMM). Study group and time of assessment (baseline, 3 months, 6 months, 12 months and 24 months) will be included as fixed factors and patient as a random factor. The model will include interactions between study group and time of assessment. The baseline value will be included as a covariate. The RMMM model will be used to quantify the treatment effect as the difference between the groups in pain scores (VAS) with the associated 95% CI and p -value at 24 months postprimary randomisation. To safeguard against potential multiplicity bias,[34] we will require a statistically significant treatment effect on both of our primary outcome variables, that is, pain at rest and pain at activity.

The same statistical model will also apply to the pragmatic comparison of the relative benefits of surgical versus non-operative treatment strategies on SIS (ASD vs ET).

Secondary analyses

We will also use the RMMM model to analyse secondary outcomes where applicable. The results will be reported as the differences between the groups with the associated 95% CI and p value at 24 months postprimary randomisation.

Categorical variables, reoperations or treatment conversions, and complications as well as AEs will be analysed using logistic regression analysis or Poisson regression dependent on whether subjects with complications or (multiple) complications (per subject) are analysed.

These secondary analyses will be supportive, explanatory and/or hypothesis-generating, which is why multiplicity is not a problem.[2]

Sensitivity analyses

We will carry out the following sensitivity analyses: (1) per-protocol analyses, in which the above noted primary and secondary analyses will be carried out again with patients who received the interventions as allocated; (2) and potential effects due to the treatment providing centres.

Subgroup analyses and hypothesised effects

We have identified three important subgroups. We will perform these three subgroup analyses with the primary end point as the outcome and the direction of hypothesised effect described[35]:

1. Duration of symptoms—Neer originally suggested that ASD should be considered for patients with persistent symptoms despite over 1 year of conservative treatment.[36] Recent randomised controlled trials failing to find efficacy on ASD (vs conservative treatment) have prompted arguments that ASD should be reserved to situations when long-term conservative treatment has failed.[37] Although a recent study specifically addressed this question and failed to support this hypothesis,[38] we still intend to compare the treatment effects of participants stratified based on the duration of symptoms. Accordingly, we will compare those with symptoms less than 12 months to those with symptoms longer than 12 months. We hypothesise that subacromial decompression will work better in patients with duration of symptoms longer than 12 months than for patients with symptoms less than12 months.
2. Severity of symptoms—A subgroup analysis will also be conducted comparing the treatment effects in patients with severe (VAS 70 or more), moderate (VAS 55 to 69) and mild (VAS less than 55) symptoms at baseline. We hypothesise that subacromial decompression will work better in patients with more severe (VAS 70 or more) than moderate (VAS 55 to 69) or mild (VAS less than 55) symptoms at baseline.
3. Acromial anatomy—A hook-type acromion has been suggested as an independent risk factor for subacromial impingement.[39] To assess the validity of this suggestion, a subgroup analysis will be conducted comparing the treatment effects in patients with flat (type I), curved (type II) or hooked (type III) acromion according to classification by Bigliani *et al*[40] We hypothesise that subacromial decompression will work better in patients with hooked (type III) than curved (type II) or flat (type I) acromion at baseline.

Effect modifying and mediating factors

Multiple regression models will be used to assess the potential effect modifying factors (eg, age, gender, psychological well-being, mental health, occupational shoulder load, education level and hand dominance) and effect mediating factors (eg, absence of complications and adherence to rehabilitation) on pain, functional disability and quality of life. These analyses are supportive, explanatory and/or hypothesis generating.

Blinded data interpretation

To safeguard against potential risk of bias during interpretation, we will use our recently introduced method of 'blinded data interpretation'.[41] So far, this method has been successfully applied to three previous trials.[33 42 43] Please refer to the online supplementary appendix for a more complete description of the process (blinded data interpretation plan), which we briefly summarise here. An independent statistician will provide the writing committee of the FIMPACT trial (authors of this protocol) with blinded results from the analyses with study groups labelled as group A, group B and group C. The writing committee will then contemplate on the interpretation of the results until a consensus is reached and agree in writing on all alternative interpretations of the findings. Once reaching a consensus, we will record the minutes of this meeting as a statement of interpretation document signed by all members of the writing committee. Only after reaching this common agreement will the data manager and independent statistician break the randomisation code.

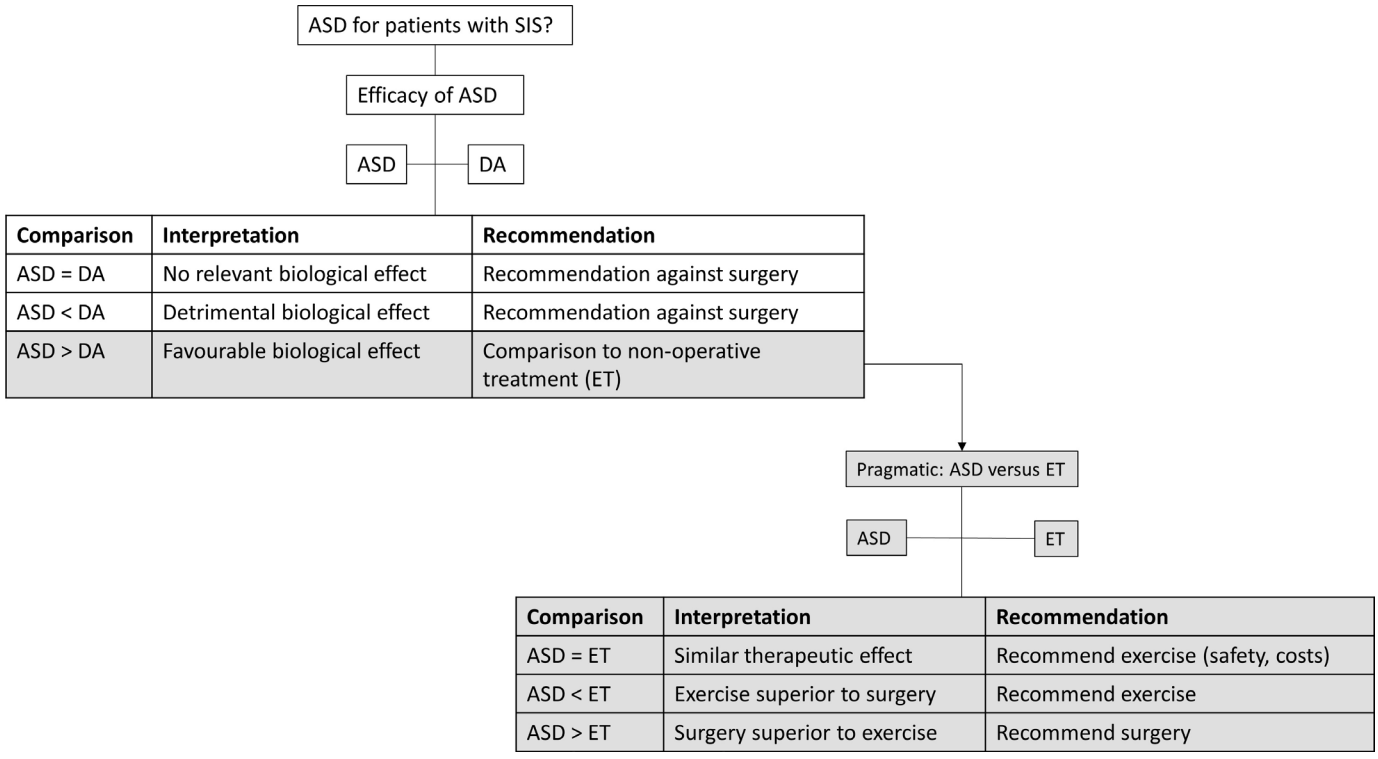
DISCUSSION

In this protocol paper, we describe the execution of a randomised, placebo-surgery controlled trial for the assessment of the efficacy of ASD in patients with SIS. Acknowledging the potential of surgery to produce powerful placebo effects,[44] our primary comparator is DA, differing from the ASD only by lacking the critical therapeutic element of the ASD (subacromial decompression). We will also conduct the pragmatic comparison of surgical and non-surgical treatment options of SIS by including a third group of progressive ET (figure 1, ASD vs ET).

Interpretations and generalisability

Our interpretation scheme primarily rests on the tenet that the minimum requirement for the clinical viability of ASD is that it needs to show superiority to DA—a therapeutically inert and thus a clinically non-viable option. To test this, we have chosen a classic *efficacy* or *'can it work'* design:[45–47] The recruited participants are those who—according to current evidence—should have an 'optimal response' to ASD and the participants and outcome assessors are blinded to the interventions given. This design should thus yield findings that are widely applicable to patients with characteristic clinical signs and symptoms of SIS. We will also compare ASD with a non-operative treatment option for SIS, the progressive ET, in a more pragmatic comparison, which is confounded by the lack of blinding of the participants (figure 2).

The generalisability of our primary (efficacy) comparison may be questioned as the patients are carefully



Comparison	Interpretation	Recommendation
ASD = DA	No relevant biological effect	Recommendation against surgery
ASD < DA	Detrimental biological effect	Recommendation against surgery
ASD > DA	Favourable biological effect	Comparison to non-operative treatment (ET)

Comparison	Interpretation	Recommendation
ASD = ET	Similar therapeutic effect	Recommend exercise (safety, costs)
ASD < ET	Exercise superior to surgery	Recommend exercise
ASD > ET	Surgery superior to exercise	Recommend surgery

Figure 2 Study design and interpretation of results. ASD, arthroscopic subacromial decompression; DA, diagnostic arthroscopy; ET, exercise therapy; SIS, subacromial impingement syndrome.

selected (strict eligibility criteria) and treated by experienced shoulder surgeons. Nevertheless, the eligibility criteria are in agreement with the existing treatment guidelines on SIS.[4] The results should thus be applicable to the specific populations currently receiving treatment for their SIS. As for the skill level of the surgeons, the index surgical procedure (ASD) is a relatively simple procedure and thus likely not very sensitive to individual surgeons' experience. For example, the amount of bone removed from the undersurface of the acromion seems to have at best a marginal effect on the outcome. Even bursectomy alone has been shown to produce the same therapeutic effect as standard acromioplasty.[48]

Rationale for outcome assessment and statistical analysis

Traditionally, the assessment of the treatment effects of two or more interventions has relied primarily on the statistical significance of the mean differences of the intervention groups. However, as described in a recent paper,[49] to truly assess the clinical relevance of a treatment, one also needs information about the distribution of individual responses. In essence, one needs to look at how many people on treatment and on comparator group(s) had a response at least as great as the MCID. Such individuals have been described as 'responders,' and this approach of comparing treatment groups as a 'responder analysis'.[50 51] The authors[49] suggested that '*Clinical trials should specify in their protocol that they will report the distribution of results in individual participants as well as the mean difference. Researchers should publish plots of individual results and responder analyses in clinical trial reports.*' The FIMPACT trial adheres to this suggested approach. Accordingly, we will elaborate several relevant and often interrelated issues, such as the study power, the primary outcomes and their interpretation, the MCID, as well as the approach we have chosen for carrying out a responder analysis.

Study power

Traditionally the sample size is calculated based on the MCID or MCII, that is, the smallest change in measurement that signifies an important/detectable improvement in a patient's symptom(s). MCII/MCID is not a static value even for one outcome instrument, but rather can have different values when assessed with different methods or *in different patient populations*. We chose VAS at rest and during arm activity as our primary outcomes, because shoulder pain is the primary complaint of patients with SIS. The FIMPACT trial was powered to detect an improvement of at least 15 on a 0–100 VAS scale[52] between ASD and ET. This yielded a sample size estimate of 70 participants per group. To safeguard against lack of study power, we chose a statistical threshold of 90% over the more conventional 80%. In this context, Norman *et al*[53] recently introduced a thought-provoking proposal arguing that a standard ('off-the-peg') sample size of 64 per group would be just as valid an estimate as one obtains by more traditional ('made-to-measure') sample size calculations.[53] Finally, although the statistical power is a vital step in the *planning phase* of any clinical trial, the actual quality of evidence (certainty in the obtained estimates) can only be appropriately assessed from the CI of the data obtained.[54]

Responder analysis

As noted above, instead of focusing only on the statistical significance of the mean differences between treatment groups in the VAS (ie, the mean improvement from baseline to 24 months), we will also carry out 'a responder analysis'. In principle, this analysis allows physicians to inform a patient of his or her chance of experiencing a clinically meaningful improvement from the treatment, both in absolute terms and in comparison, to a control group. The difference between responders and non-responders can be considered the net benefit of the treatment. One proposed means to carry out a responder analysis relies on the assessment of the proportion of patients reaching the patient-acceptable symptom state (PASS) and the patient-disappointing symptoms state (PDSS). As no universal consensus exists on either the PASS or the PDSS in the context of SIS, we chose to anchor our responder analysis to the patient's assessment of satisfaction with the shoulder treatment outcome: Patients reporting very satisfied or satisfied will be categorised as 'Responders' and those reporting very dissatisfied or dissatisfied as 'Non-responders'. Given the obvious coarseness of this approach, we plan to evaluate the appropriate criteria for PASS and PDSS in more detail in the future, exploring the potential contribution of, for example, arm pain at rest and at activity, shoulder function, and night pain.

Ethics of placebo surgery

A recent systematic review of the use of surgical placebo shows that in more than half of these studies the treatment group that included critical surgical/therapeutic element had no greater effect than a placebo group.[18] The review also showed that risks of AEs were small and the placebo group was safer than the surgery under investigation. These findings make a compelling case for the use of surgical placebo controls when a placebo effect may be present. Regarding the ethics of surgical placebo controls, the authors of the review state '*Placebo controlled surgical trials raise important ethical concerns but are justified when there is a genuine equipoise; that is, a disagreement in the medical community about whether one treatment is superior to another, because standard treatment does not exist or its efficacy is questioned.*' They continue by concluding: '*Placebo controlled trials in surgery are as important as they are in medicine, and they are justified in the same way. They are powerful, feasible way of showing the efficacy of surgical procedures. They are necessary to protect the welfare of present and future patients as well as to conduct proper cost effectiveness analyses. Only then may publicly funded surgical interventions be distributed fairly and justly. Without such studies ineffective treatment may continue unchallenged.*' Our views regarding the ethics of using a surgical placebo group are perfectly aligned with these notions.

Limitations of the study

One possible confounder in our trial is that subacromial pain is also the hallmark symptom of a RC tear, although the latter patients usually also represent with muscle weakness. To exclude patients with a (clinically relevant) RC tear, our eligibility screening included two preoperative assessments: (A) clinical exams targeted at finding obvious weakness of the RC muscles and (B) MRA, an imaging modality with a shown 92 specificity and 94 sensitivity for 'full-thickness' RC tears.[55] In addition to these, we also carried out (C) a DA in the ASD and DA groups prior to randomisation. Despite the thorough *preoperative* screening, 10% (14/136) of the participants allocated to the two surgical groups had to be excluded because of acromioclavicular arthrosis (n=1) or intra-articular pathology found at DA (n=13). Although this does not have any effect on our primary comparison (ASD vs DA), one could argue that the ET and operatively treated groups (ASD and DA) are not fully comparable. At the same time we don't know the clinical relevance of small RC tears or superior labrum anterior to posterior (SLAP) lesions, which don't result in obvious muscle weakness and/or are not apparent in MRA. In the end, if this bias proves clinically relevant in our analysis, it will skew our results by favouring the ASD group in the pragmatic comparison (ASD vs ET). Another concern related to the pragmatic comparison (ASD vs ET) is that the progressive ET regimen carried out in the ET group is different from the postoperative rehabilitation carried out by patients in the ET group, for obvious reasons; surgically treated patients need time to recover from the initial surgical trauma. Furthermore, patients with ASD are also subject to some degree of postoperative immobilisation, extended sick leave, and modifications in pain medication and activities, all of which potentially have an effect on the outcome of treatment.

Another obvious concern related to our study design is the discrepant timing of the start of the actual treatment between the ET and the two surgical groups due to the time required to arrange the surgery. Acknowledging this, the 2-year follow-up was chosen as our primary time point for assessing the benefits of treatment, as we assume that by this time the potential confounding effect of slightly different follow-up times should be diluted to a minimum. This is also the reason why we use data from the shorter-term follow-up visits (ie, visits performed at 3 months, 6 months and 12 months after randomisation) primarily to illustrate the trajectory of the treatment response in the three groups. Concerns over the varying time span from the randomisation of the patients to the trial to the actual induction of treatment (due to delay in surgery) also applies to the CSAW trial.[20] To compensate for the waiting list effects, the CSAW investigators have chosen a slightly different strategy: Although the primary outcome assessment is performed at 6 months after randomisation in the CSAW trial, they have introduced additional follow-up assessments, referenced from surgery, for patients waiting for longer than 4 months for their surgery after randomisation. They have also set a secondary outcome measurement point at 1 year postrandomisation

ETHICS AND DISSEMINATION

Ethics

FIMPACT trial is conducted in accordance with the principles of the Declaration of Helsinki. This trial has been approved by the Institutional Review Board of the Pirkanmaa Hospital District and each participating centre granted clinical trial authorisation prior to recruitment. The trial has been registered to ClinicalTrials.gov registry and any revisions about the protocol are documented in this registry. For each participant, informed consent is obtained prior to any study-related procedures.

Dissemination policy

We aim to produce high-impact publications of the results of the trial and present the findings to the clinicians who manage shoulder pain in the front line. The investigators will be involved in preparing drafts of the manuscripts, abstracts, press releases and any other publications arising from the trial. The final reporting will follow the Consolidated Standards of Reporting Trials (CONSORT) Statement guidelines. Authorship will be determined in accordance with the International Committee of Medical Journal Editors (ICMJE) guidelines and other contributors will be acknowledged. The funders will be acknowledged in all resulting publications. There is no intended use of professional writers.

Date and version identifier

30 August 2016: Original protocol, (V.1.0)

5 December 2016: Revised protocol, major revision (V.2.0)

13 February 2017: Re-revised protocol, minor revision (V.2.1)

All versions available at http://bmjopen.bmj.com

Author affiliations
[1]Department of Orthopedics and Traumatology, Helsinki University Hospital and University of Helsinki, Helsinki, Finland
[2]National Institute for Health and Welfare, Center for Health and Social Economics, Helsinki, Finland
[3]Department of Orthopedics and Traumatology, Hatanpää City Hospital, Tampere, Finland

Acknowledgements The FIMPACT investigators thank Professor Jonas Ranstam (Lund University) for the statistical analysis plan and Dr Pekka Paavolainen and Professor Markku Järvinen for their critical contributions and support on the early phases of the planning and execution of the trial.

Collaborators Participating Clinical Sites: Helsinki University Hospital, Jorvi Hospital: Kalevi Hietaniemi, Juha Kalske, Vesa Lepola, Jyrki Salmenkivi, Sikri Tukiainen. Helsinki University Hospital, Herttoniemi Hospital: Jarkko Pajarinen, Mikko Salmela, Vesa Savolainen, Ilkka Sinisaari. Hatanpää City Hospital, Tampere: Timo Järvelä, Kari Kanto, Janne Lehtinen, Mikael Salmela. FIMPACT Methods Center: Leena Caravitis, Sari Karesvuori, Pirjo Toivonen (Project management), Mathias Bäck (Data management), Ville Haapamäki (Imaging) (Helsinki University Hospital and University of Helsinki), Jari Inkinen (Physiotherapy) (Fysios Finlayson, Physiotherapy center Kunnon Klinikka Oy, Tampere), Esa Läärä (Randomization) (University of Oulu), Harri Sintonen (QoL outcomes) (National Institute for Health and Welfare).

Contributors Concept and design: MP, AM, ST and TLNJ. Drafting and critical revision of the article for important intellectual content: MP, AM, ST, TLNJ, KK. Final approval of the article: MP, AM, ST, TLNJ, KK. Ensuring the accuracy of the work: MP, AM, ST, TLNJ, KK. Statistical expertise: Jonas Ranstam. Obtaining of funding: TLNJ and Markku Järvinen. Primary sponsor: MP. Steering committee: MP, TLNJ, AM, ST. Writing committee: TLNJ, ST, AM, MP, KK.

Funding The FIMPACT trial was supported by the Sigrid Juselius Foundation, the State funding for university/level health research, Pirkanmaa and Helsinki University Hospital Districts, the Academy of Finland, and the Jane and Aatos Erkko Foundation. The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of this protocol.

Competing interests ST reports personal fees from Evalua group of companies, personal fees from DBC group of companies, and personal fees from insurance companies, outside the submitted work.KK reports an honorarium for a lecture from Linvatec, outside the submitted work. TLNJ reports an honorarium for a lecture on osteoporosis from AMGEN (donated to AllTrials campaign). Authors not named here have disclosed no conflicts of interest.

Patient consent Obtained.

Ethics approval The institutional review board of the Pirkanmaa Hospital District (R04200).

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. Paloneva J, Lepola V, Karppinen J, *et al*. Declining incidence of acromioplasty in Finland. *Acta Orthop* 2015;86:220–4.
2. Vitale MA, Arons RR, Hurwitz S, *et al*. The rising incidence of acromioplasty. *J Bone Joint Surg Am* 2010;92:1842–50.
3. Engebretsen K, Grotle M, Bautz-Holter E, *et al*. Radial extracorporeal shockwave treatment compared with supervised exercises in patients with subacromial pain syndrome: single blind randomised study. *BMJ* 2009;339:b3360.
4. Diercks R, Bron C, Dorrestijn O, *et al*. Dutch Orthopaedic Association. Guideline for diagnosis and treatment of subacromial pain syndrome: a multidisciplinary review by the Dutch Orthopaedic Association. *Acta Orthop* 2014;85:314–22.
5. Oh LS, Wolf BR, Hall MP, *et al*. Indications for rotator cuff repair: a systematic review. *Clin Orthop Relat Res* 2007;455:52–63.
6. Aurora A, McCarron J, Iannotti JP, *et al*. Commercially available extracellular matrix materials for rotator cuff repairs: state of the art and future trends. *J Shoulder Elbow Surg* 2007;16(5 Suppl):S171–S178.
7. Michener LA, McClure PW, Karduna AR. Anatomical and biomechanical mechanisms of subacromial impingement syndrome. *Clin Biomech* 2003;18:369–79.
8. Koester MC, George MS, Kuhn JE. Shoulder impingement syndrome. *Am J Med* 2005;118:452–5.
9. Neer CS. Anterior acromioplasty for the chronic impingement syndrome in the shoulder: a preliminary report. *J Bone Joint Surg Am* 1972;54:41–50.
10. Yu E, Cil A, Harmsen WS, *et al*. Arthroscopy and the dramatic increase in frequency of anterior acromioplasty from 1980 to 2005: an epidemiologic study. *Arthroscopy* 2010;26(9 Suppl):S142–S147.
11. Judge A, Murphy RJ, Maxwell R, *et al*. Temporal trends and geographical variation in the use of subacromial decompression and rotator cuff repair of the shoulder in England. *Bone Joint J* 2014;96-B:70–4.
12. Dong W, Goost H, Lin XB, *et al*. Treatments for shoulder impingement syndrome: a PRISMA systematic review and network meta-analysis. *Medicine* 2015;94:e510.
13. Shi LL, Edwards TB. The role of acromioplasty for management of rotator cuff problems: where is the evidence? *Adv Orthop* 2012;2012:1–5.
14. Brox JI, Staff PH, Ljunggren AE, *et al*. Arthroscopic surgery compared with supervised exercises in patients with rotator cuff disease (stage II impingement syndrome). *BMJ* 1993;307:899–903.

15. Haahr JP, Østergaard S, Dalsgaard J, *et al*. Exercises versus arthroscopic decompression in patients with subacromial impingement: a randomised, controlled study in 90 cases with a one year follow up. *Ann Rheum Dis* 2005;64:760–4.
16. Ketola S, Lehtinen J, Rousi T, *et al*. No evidence of long-term benefits of arthroscopicacromioplasty in the treatment of shoulder impingement syndrome: Five-year results of a randomised controlled trial. *Bone Joint Res* 2013;2:132–9.
17. Ernst E, Resch KL. Concept of true and perceived placebo effects. *BMJ* 1995;311:551–3.
18. Wartolowska K, Judge A, Hopewell S, *et al*. Use of placebo controls in the evaluation of surgery: systematic review. *BMJ* 2014;348:g3253.
19. Buchbinder R. Meniscectomy in patients with knee osteoarthritis and a meniscal tear? *N Engl J Med* 2013;368:1740–1.
20. Beard D, Rees J, Rombach I, *et al*. CSAW Study Group. The CSAW Study (Can Shoulder Arthroscopy Work?) - a placebo-controlled surgical intervention trial assessing the clinical and cost effectiveness of arthroscopic subacromial decompression for shoulder pain: study protocol for a randomised controlled trial. *Trials* 2015;16:210.
21. Ellman H. Arthroscopic subacromial decompression: analysis of one- to three-year results. *Arthroscopy* 1987;3:173–81.
22. Tashjian RZ, Deloach J, Porucznik CA, *et al*. Minimal clinically important differences (MCID) and patient acceptable symptomatic state (PASS) for visual analog scales (VAS) measuring pain in patients treated for rotator cuff disease. *J Shoulder Elbow Surg* 2009;18:927–32.
23. Constant CR, G. Murley AH. A clinical method of functional Assessment of the shoulder. *Clin Orthop Relat Res* 1987;160:164–4.
24. Henseler JF, Kolk A, van der Zwaal P, *et al*. The minimal detectable change of the Constant score in impingement, full-thickness tears, and massive rotator cuff tears. *J Shoulder Elbow Surg* 2015;24(3):376–81.
25. Matsen FA, Ziegler DW, DeBartolo SE. Patient self-assessment of health status and function in glenohumeral degenerative joint disease. *J Shoulder Elbow Surg* 1995;4:345–51.
26. Godfrey J, Hamman R, Lowenstein S, *et al*. Reliability, validity, and responsiveness of the simple shoulder test: psychometric properties by age and injury type. *J Shoulder Elbow Surg* 2007;16:260–7.
27. Tashjian RZ, Deloach J, Green A, *et al*. Minimal clinically important differences in ASES and simple shoulder test scores after nonoperative treatment of rotator cuff disease. *J Bone Joint Surg Am* 2010;92:296–303.
28. Sintonen H. The 15D instrument of health-related quality of life: properties and applications. *Ann Med* 2001;33:328–36.
29. Moock J, Kohlmann T. Comparing preference-based quality-of-life measures: results from rehabilitation patients with musculoskeletal, cardiovascular, or psychosomatic disorders. *Qual Life Res* 2008;17:485–95.
30. Bowling A. Measuring health. *A review of quality of life measurement scales*. 3rd ed. Berkshire: Open University Press, 2004.
31. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473–83.
32. Garratt A, Schmidt L, Mackintosh A, *et al*. Quality of life measurement: bibliographic study of patient assessed health outcome measures. *BMJ* 2002;324:1417.
33. Sihvonen R, Paavola M, Malmivaara A, *et al*. Arthroscopic partial meniscectomy versus sham surgery for a degenerative meniscal tear. *N Engl J Med* 2013;369:2515–24.
34. Ranstam J. Multiple P-values and Bonferroni correction. *Osteoarthritis Cartilage* 2016;24:763–4.
35. Sun X, Ioannidis JP, Agoritsas T, *et al*. How to use a subgroup analysis: users' guide to the medical literature. *JAMA* 2014;311:405–11.
36. Neer CS. 2nd. impingement lesions. *Clin Orthop Relat Res* 1983;173:70–7.
37. Dong W, Goost H, Lin XB, *et al*. Treatments for shoulder impingement syndrome: a PRISMA systematic review and network meta-analysis. *Medicine* 2015;94:e510.
38. Ketola S, Lehtinen J, Rousi T, *et al*. Which patients do not recover from shoulder impingement syndrome, either with operative treatment or with nonoperative treatment? *Acta Orthop* 2015;86:641–6.
39. Tangtrakulwanich B, Kapkird A. Analyses of possible risk factors for subacromial impingement syndrome. *World J Orthop* 2012;3:5–9.
40. Bigliani LU MD. April EW The morphology of the acromion and its relationship to rotator cuff tears. *Orthop Trans* 1986;10:228.
41. Järvinen TL, Sihvonen R, Bhandari M, *et al*. Blinded interpretation of study results can feasibly and effectively diminish interpretation Bias. *J Clin Epidemiol* 2014;67:769–72.
42. Skou ST, Roos EM, Laursen MB, *et al*. A randomized, controlled trial of total knee replacement. *N Engl J Med* 2015;373:1597–606.
43. Bhandari M, Jeray KJ, Petrisor BA, *et al*. A trial of Wound Irrigation in the initial management of Open Fracture Wounds. *N Engl J Med* 2015;373:2629–41.
44. Jonas WB, Crawford C, Colloca L, *et al*. To what extent are surgery and invasive procedures effective beyond a placebo response? a systematic review with meta-analysis of randomised, sham controlled trials. *BMJ Open* 2015;5:e009655.
45. Haynes B. Can it work? does it work? is it worth it? the testing of healthcareinterventions is evolving. *BMJ* 1999;319:652–3.
46. Djulbegovic B, Paul A. From efficacy to effectiveness in the face of uncertainty: indication creep and prevention creep. *JAMA* 2011;305:2005–6.
47. Järvinen TL, Sievänen H, Kannus P, *et al*. The true cost of pharmacological disease prevention. *BMJ* 2011;342:d2175.
48. Donigan JA, Wolf BR. Arthroscopic subacromial decompression: acromioplasty versus bursectomy alone--does it really matter? a systematic review. *Iowa Orthop J* 2011;31:121–6.
49. Cates C, Karner C. Clinical importance cannot be ruled out using mean difference alone. *BMJ* 2015;351:h5496.
50. Guyatt GH, Juniper EF, Walter SD, *et al*. Interpreting treatment effects in randomised trials. *BMJ* 1998;316:690–3.
51. Snapinn SM, Jiang Q. Responder analyses and the assessment of a clinically relevant treatment effect. *Trials* 2007;8:31.
52. Tashjian RZ, Deloach J, Porucznik CA, *et al*. Minimal clinically important differences (MCID) and patient acceptable symptomatic state (PASS) for visual analog scales (VAS) measuring pain in patients treated for rotator cuff disease. *J Shoulder Elbow Surg* 2009;18:927–32.
53. Norman G, Monteiro S, Salama S. Sample size calculations: should the emperor's clothes be off the peg or made to measure? *BMJ* 2012;345:e5278.
54. Goodman SN, Berlin JA. The use of predicted confidence intervals when planning experiments and the misuse of power when interpreting results. *Ann Intern Med* 1994;121:200–6.
55. Lenza M, Buchbinder R, Takwoingi Y, *et al*. Magnetic resonance imaging, magnetic resonance arthrography and ultrasonography for assessing rotator cuff tears in people with shoulder pain for whom surgery is being considered. *Cochrane Database Syst Rev* 2013;9:CD009020.