Exercises versus arthroscopic decompression in patients with subacromial impingement: a randomised, controlled study in 90 cases with a one year follow up

J P Haahr, S Østergaard, J Dalsgaard, K Norup, P Frost, S Lausen, E A Holm, J H Andersen


S

houlder pain is common. In a Dutch study the incidence of new cases of rotator cuff tendinitis in general practice was found to be around 3.2 to 4.2 per 1000 person–years, and the corresponding incidence of shoulder pain (all causes) was 11.2 per 1000 person–years.1

Rotator cuff disease with subacromial impingement has been graded in three stages: stage 1, acute inflammation, and either tendinitis or bursitis; stage 2, chronic inflammation with or without degeneration; stage 3, full rupture of the cuff.2 The anatomical basis for impingement is a mismatch between the structures in the subacromial space. This aggravates or provokes pain. The main idea of the treatments given is to control pain and remedy the mechanical problem in order to preserve or improve function. Improved function can be obtained through reduction of inflammatory oedema, strengthening of the muscles, which act as depressors and stabilisers of the humeral head, or by removing fibrotic tissue in the subacromial bursa and a part of the acromion itself.

The condition is often treated conservatively in the primary health care sector by general practitioners or physiotherapists.3 Studies of the effectiveness of physiotherapy versus corticosteroid injections have found inconsistent short term results. Hay et al, in a trial from general practice, found no difference in effectiveness at six months’ follow up.4

Patients with resistant or longstanding shoulder pain are often referred for specialist treatment, included surgery. However, according to a Cochrane review, there is little evidence to support or refute the efficacy of common interventions for shoulder pain.5 Also, the evidence supporting the superiority of subacromial decompression relative to physiotherapy with training has been un convincing.6–9

Our objective in this study was to test the effect of graded physiotherapeutic training versus arthroscopic subacromial decompression after 12 months’ follow up in patients referred to specialist treatment for shoulder pain with subacromial impingement.

METHODS

Patients and recruitment
The study was conducted between 1996 and 2001 at Herning Hospital, Ringkøbing County, Denmark, as a collaborative project between the departments of occupational medicine, rheumatology and physical rehabilitation, orthopaedic surgery, and radiology. The departments of occupational medicine and of rheumatology and physical rehabilitation receive shoulder patients referred from the primary health care sector. These referrals constituted the recruitment base. Diagnostic criteria required were: the presence of shoulder pain, pain on abduction of the shoulder with painful arch, a positive impingement sign (Hawkins sign) and a positive impingement test (relief of pain within 15 minutes after injection of local anaesthetic (bupivacaine 5 ml) into the subacromial space).

A rheumatologist (SO) assessed all patients. The eligibility criteria for participation were: fulfilment of all diagnostic criteria, report of shoulder symptoms between six months and three years (because surgery in general was not offered to cases with symptoms of shorter duration), and age between 18 and 55 years. Previous treatment with rest, non-steroidal anti-inflammatory drugs, subacromial injection, and physiotherapy were allowed. Normal passive glenohumeral movement was a requirement. Patients were excluded for the following reasons: impaired rotation in the glenohumeral joint, a history of acute trauma, previous surgery or previous fracture in the proximity of the affected shoulder.

Statements of evidence (SOE) were calculated for all outcomes. Patients were randomised either to arthroscopic subacromial decompression, or to physiotherapy with training. Further studies are needed to qualify treatment choice decisions, and it is recommended that samples are stratified according to disability level.

Abbreviations: GEE, generalised estimation equations; MIREDEF, minimum relevant clinical difference; VAS, visual analogue scale.
shoulder, known osteoarthritis in the acromioclavicular or glenohumeral joints, calcifications exceeding than 2 cm in the rotator cuff tendons, or signs of a rupture of the cuff or cervical root syndromes.

**Study protocol**
Consecutively referred patients who fulfilled the inclusion criteria were informed about the project. Those interested in participation underwent a clinical reappraisal by a specialist at the rheumatology department. The same specialist (SO) carried out all the assessments, obtained informed consent for participation, and randomised the patients into one of two intervention groups by opening a sealed envelope containing the result of randomisation, which was unknown to SO. A computer program was used to generate a random sequence of allocation. In patients with bilateral symptoms the most affected shoulder was chosen as the primary intervention shoulder. After assessment and randomisation the patient was referred to x ray and ultrasonography of the shoulders. The rheumatologist filled in a baseline registration card, and gave the patient a baseline questionnaire to be completed and submitted to the department of occupational medicine before the start of the intervention. Radiographic and ultrasonography findings are not presented here.

The study was approved by the hospital ethics committee.

**Intervention**
Intervention in both groups began four weeks after enrolment.

The physiotherapeutic treatment consisted of 19 sessions, each lasting up to 60 minutes, given by two experienced therapists (SL and EA). The treatments started with application of heat, cold packs, or soft tissue treatments. This was followed by active training of the periscapular muscles (rhomboid, serratus, trapezoid, levator scapulae, and pectoralis minor muscles) and strengthening of the stabilising muscles of the shoulder joint (the rotator cuff). The team instructing the surgery group was different from the group treating the same shoulder. Three experienced surgeons undertook all procedures and recorded their findings on a predetermined proforma. Before discharge, the patient was instructed in performing light movements of the arm within the limits of pain. Stitches were removed by general practitioners after 10 days. At the same time, the patient was instructed by a physiotherapist to carry out increasingly active exercises, including exercises for strengthening the rotator cuff muscles. The team instructing the surgery group was different from the group treating the shoulder, known osteoarthritis in the acromioclavicular or glenohumeral joints, calcifications exceeding than 2 cm in the rotator cuff tendons, or signs of a rupture of the cuff or cervical root syndromes.

**Outcome measures**
All the patients were evaluated at baseline immediately before the intervention, and after three, six, and 12 months.
Evaluation was done by two physiotherapists (SV and EH) using the Constant score,\(^\text{10}\) which is a joint measure of four subscores: pain (VAS: 0 = max, 15 = no pain), function (ADL and movement: 0–20), range of movement (0–40), and force (0–25). Pain was measured on a visual analogue scale (VAS); function limitations in activities of daily living; active range of motion in four directions in the shoulder joint; and isometric shoulder strength measured in kg with a portable muscle strength analyser (Isobex 2.1, Cursor AG, Bern, Switzerland). Each kg was allocated 2 points up to 25 points for strength of at least 12 kg. Based on measurements of shoulder force in healthy male and female workers the force measurements were adjusted by multiplying the measure-ments by a factor of 1.94 in order to compare the values for women. The total Constant score sums up to 100 points, which indicates normal function. Physiotherapists were not blinded to the treatment given when assessing the Constant score.

After one year, patients filled in a follow up questionnaire, which repeated various questions given at baseline. In a set of four questions the patients were asked to indicate pain and dysfunction for each shoulder by using a numerical box complaint scale (Likert scale) ranging from 0 (no complaints at all) to 9 (pain as bad as could be)\(^\text{12}\) for:

- severity of worst pain and discomfort within the past three months;
- average pain and discomfort within the past three months;
- level of average pain and discomfort within the past seven days.

The scale has been used previously in the Danish study project on research and intervention in monotonous work (PRIM).\(^\text{13}\)

Information was collected at baseline on workplace and job title for the actual or latest jobs held (up to five appointments), employment within the past three months, sick leave, and having a labour compensation claim. Jobs were classified as either strenuous or not strenuous.\(^\text{14}\)

### Statistical analysis

The study’s central hypothesis was tested by comparing change in the Constant score between the two groups for the intervention shoulder. The difference in the Constant score between treatment groups from baseline to three, six, and 12 months’ follow up was tested using one way analysis of variance (ANOVA). The difference in Constant score between the two treatment groups at each measurement time was tested by GEE (generalised estimation equation) analysis. GEE corrects for the correlation and lack of independence of an individual’s responses by using quasi-likelihood methods and robust variance estimators. We introduced all baseline characteristics (table 1) in the model. None of the variables produced changes in regression coefficients greater than 5%. In the final model we retained sex (p = 0.54), age (p = 0.99), workers’ compensation claim (p = 0.60), and the function subscale of the Constant score at baseline (p = 0.28) as potential confounding variables.

The sample size was set at a minimum of 40 patients in each group based on an expected improvement of 30% in the physiotherapy group (mean (SD) expected baseline Constant score, 25 (14)), an x value set at 0.05 (type I error), and β at 0.20 (type II errors), and a minimum relevant clinical difference (MIREDIF) of 50% between the two groups in favour of surgery (corresponding to 9 to 10 points). Thus, a

---

### Table 2

<table>
<thead>
<tr>
<th>Change in score</th>
<th>Physiotherapy group</th>
<th>Surgery</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain (VAS: 0 = max, 15 = no pain)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline to 3 months</td>
<td>3.1 (2.1 to 4.3)</td>
<td>2.8 (1.7 to 4.0)</td>
<td>0.69</td>
</tr>
<tr>
<td>Baseline to 6 months</td>
<td>3.7 (2.6 to 4.8)</td>
<td>3.8 (2.6 to 5.0)</td>
<td>0.92</td>
</tr>
<tr>
<td>Baseline to 12 months</td>
<td>3.7 (2.7 to 4.6)</td>
<td>3.6 (2.3 to 4.9)</td>
<td>0.93</td>
</tr>
<tr>
<td>Function (ADL and movement: 0–20)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline to 3 months</td>
<td>3.7 (2.6 to 4.8)</td>
<td>3.7 (2.1 to 5.3)</td>
<td>0.96</td>
</tr>
<tr>
<td>Baseline to 6 months</td>
<td>4.6 (3.2 to 6.1)</td>
<td>3.7 (2.0 to 5.4)</td>
<td>0.38</td>
</tr>
<tr>
<td>Baseline to 12 months</td>
<td>4.5 (3.1 to 6.0)</td>
<td>3.8 (2.1 to 5.4)</td>
<td>0.46</td>
</tr>
<tr>
<td>Range of movement (0–40)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline to 3 months</td>
<td>10.7 (7.7 to 13.5)</td>
<td>6.8 (3.4 to 10.3)</td>
<td>0.09</td>
</tr>
<tr>
<td>Baseline to 6 months</td>
<td>10.3 (7.1 to 13.5)</td>
<td>9.6 (6.2 to 12.9)</td>
<td>0.76</td>
</tr>
<tr>
<td>Baseline to 12 months</td>
<td>11.6 (8.3 to 14.8)</td>
<td>8.2 (4.6 to 11.8)</td>
<td>0.17</td>
</tr>
<tr>
<td>Force (0–25)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline to 3 months</td>
<td>2.4 (1.1 to 3.7)</td>
<td>2.1 (0.4 to 3.8)</td>
<td>0.71</td>
</tr>
<tr>
<td>Baseline to 6 months</td>
<td>2.7 (1.6 to 3.9)</td>
<td>2.9 (0.8 to 5.0)</td>
<td>0.88</td>
</tr>
<tr>
<td>Baseline to 12 months</td>
<td>3.2 (1.7 to 4.7)</td>
<td>3.3 (1.1 to 5.4)</td>
<td>0.96</td>
</tr>
<tr>
<td>Constant score (0–100)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline to 3 months</td>
<td>20.1 (15.0 to 25.0)</td>
<td>15.5 (9.1 to 21.9)</td>
<td>0.27</td>
</tr>
<tr>
<td>Baseline to 6 months</td>
<td>21.3 (15.4 to 27.2)</td>
<td>19.9 (12.7 to 27.1)</td>
<td>0.76</td>
</tr>
<tr>
<td>Baseline to 12 months</td>
<td>23.0 (16.9 to 29.1)</td>
<td>18.8 (11.5 to 26.1)</td>
<td>0.38</td>
</tr>
</tbody>
</table>

Values are mean (95% confidence interval) by one way analysis of variance. Constant’s shoulder score = sum of pain, function, range of movements, and force. ADL, activities of daily living; VAS, visual analogue scale.

---

### Table 3

<table>
<thead>
<tr>
<th>Constant score</th>
<th>β Coefficient</th>
<th>95% CI</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>0.007</td>
<td>-0.005 to 0.019</td>
<td>0.27</td>
</tr>
<tr>
<td>3 months</td>
<td>-0.006</td>
<td>-0.015 to 0.003</td>
<td>0.20</td>
</tr>
<tr>
<td>6 months</td>
<td>0.005</td>
<td>-0.004 to 0.013</td>
<td>0.31</td>
</tr>
<tr>
<td>12 months</td>
<td>-0.003</td>
<td>-0.010 to 0.004</td>
<td>0.41</td>
</tr>
</tbody>
</table>

GEE analysis with adjustment for sex, age, workers’ compensation claim, and the function subscale of the Constant score at baseline. CI, confidence interval; GEE, generalised estimation equations.
prior, we intended to include 100 patients in expectation of a number of dropouts.

For the secondary outcome measure of pain and dysfunction the subscores of the four pain and function questions were added to a single total score ranging from 0 to 36, and this score was compared for the two intervention groups by ANOVA. Analyses were done as per intention to treat.

RESULTS

Ninety consecutive patients with subacromial impingement agreed to participate. Forty five cases were randomised to conservative treatment and 45 to surgical treatment. Among those assigned to conservative treatment, one withdrew from participation because of work problems and one failed to fill in the baseline questionnaire, leaving 43 cases in this group. In the surgery group, four cases dropped out before the start of the study (one because of work problems and one failed to fill in the baseline questionnaire, leaving 43 cases in this group. Within the conservative treatment group, a further six patients were operated on within the 12 months of the study (five because of unsatisfactory improvement during exercises and in one case because a labral lesion was suspected).

In the physiotherapy group 42 persons (93%) were followed for 12 months with the main outcome measure (Constant score). In the surgery group 40 persons (89%) had complete follow up data.

The distribution of the baseline characteristics among the 84 patients is shown in table 1 by treatment group. The two groups were very similar, though a slightly greater proportion within the surgery group had been on sick leave owing to shoulder pain within the past three years. Within the surgery group no cases with stage III impingement (complete tear of the cuff) were found.

The baseline Constant score was 34.8 in the physiotherapy group and 33.7 in the surgery group. Within the physiotherapy group the Constant score improved to 54.8, 55.3, and 57.0 after three, six, and 12 months. In the surgery group the corresponding values were 49.2, 53.8, and 52.7. Only 20 cases obtained a Constant score of 80 or more after one year (10 in each group). The mean improvement in Constant score in the physiotherapy group was 23.0 (95% confidence interval (CI), 16.9 to 29.1), and in the surgery group the improvement was 18.8 (11.5 to 26.1). Two patients in the physiotherapy group and eight in the surgery group had a reduction in the Constant score.

Table 2 shows the mean change in score with 95% confidence intervals from baseline by treatment group. Table 3 shows the GEE analysis of the difference between the two groups in Constant score at the different times of measurement. There was no difference at any point of follow up, neither did the results suggest any trends during the study period.

The secondary outcome measure of pain and discomfort is shown in table 4. No differences were found between the two treatment groups, and both groups improved during follow up.

DISCUSSION

We found similar improvements in the two treatment groups, as measured by the Constant score and the pain and dysfunction score. The greatest improvement occurred within the first three months of treatment. The patients had lower scores, both at the beginning and at the end of the study, compared with previously reported studies of treatment for rotator cuff disease with impingement syndrome.\(^*\)\(^\text{15}\)

Internal validity

The unblinded assessment of Constant scores may have introduced a bias in favour of the conservative approach, because the same physiotherapists who instructed the physiotherapy group also carried out assessment of the Constant scores. It is a weakness that the baseline Constant scoring was not done before randomisation and was postponed until just before the start of the treatment. The self reported pain and dysfunction score may also be biased by the patients’ own preferences for a particular treatment, which have not been recorded. If this bias is present, it can be assumed to be small, because the randomisation was otherwise successful and there was a low drop out rate. It is also reassuring that the results for the two different outcomes are in good agreement. Six patients in the conservative treatment group were operated on during follow up. Among these one might expect a better outcome and more improvement, but this was not so (mean constant score at 12 months was 41 (range 17 to 78). The a priori power of the study was intended to be 0.80. The actual standard deviation of the difference in the Constant score after 12 months was higher than estimated in the power calculation (21). Consequently with a power of 0.80, the MIREDIF is 13 points.

External validity

The patients differed in some ways from those in previous studies. Our cases were younger than those studied by Brox et al and Andersen et al.\(^*\)\(^\text{15}\) A greater proportion had been sick listed (73% v 54–75% in the study by Brox et al), and more cases had filed a work compensation claim (75% v 25% in the
study by Andersen et al). On the other hand, the cases in the study by Rahme et al had the same mean age (42 years) and the same proportion were sick listed (76%). Another difference from previous studies is the very low baseline Constant score. Our patients may therefore appear to differ from those normally referred for subacromial decompression. The reason for this could be related to the setting. Medical services in a provincial hospital setting, where the study took place—with fewer resources and longer waiting times to specialised treatment compared with counties with well-staffed and equipped university hospitals—may be related to lower Constant scores. This could also explain the large numbers reporting sick leave. The high number of work compensation claims may in part be explained by many cases seen at the department of occupational health. Furthermore Danish legislation requires a claim to be filed whenever an occupational disorder is suspected.

Even though the effects of surgical treatment have been unconvincing compared to physiotherapy, surgical treatment of subacromial impingement has been widely adopted in the secondary health care sector, and the predominant treatment is now arthroscopic subacromial decompression. Brox et al, in their study of 125 patients with a 2.5 year follow up, defined a successful outcome as the acquisition of a Neer score greater than 80, and found an odds ratio for success after surgery compared to conservative treatment of 1.5 (95% CI, 0.6 to 3.7). In their study of 72 patients Peters and Kohn used a modified questionnaire based Constant score and found that the surgical group had slightly higher scores after four years of follow up (mean value 84 v 74 points and total improvement of 30 v 15 points). They concluded that both treatment approaches were justified. In comparison, as mentioned above, our study found a lower score at baseline, and a lower score was also attained after treatment. This may reflect the fact that we did not exclude patients with bilateral pain and muscular tenderness, and that other studies have included patients with pain for less than three months. It has previously been mentioned that filling a work compensation claim predicts a poor prognosis. However, even though the Constant score in these patients was lower at baseline (32.6 v 39.5), they improved as much from baseline to the one year follow up as the other patients (mean increase 21.1 v 20.6). In a case–only study of 60 patients with shoulder impingement, the patients with workers’ compensation claims had a lower baseline Constant score, but improvement was 24 points v 29 points in the group without compensation claims. Thus claimants improve equally well, but to a more modest level.

Then who should be operated on?

From the results of their study, Brox et al recommended that patients who do not improve within six months using a supervised exercise regimen should be evaluated for surgery. The reason for this is not well documented in their own data or in other studies. Some patients may be harmed by treatment, as illustrated by our finding that some got worse in both treatment groups. This has been afforded little or no attention in the past. This risk should lead to greater caution in treatment choice decisions. In view of our findings, we are now more reluctant to recommend surgery in cases with stage II impingement. There is a need for larger scale studies with sufficient numbers of participants to allow for stratification into subgroups with different baseline levels of disability, whatever functional score one uses, before rigorous recommendations are made about who should have arthroscopic decompression and who could benefit from physiotherapy with training, maybe in combination with other medical treatments. This ought to be a prerequisite for the continually expanding industry of arthroscopic decompression operations in the shoulder.

ACKNOWLEDGEMENTS

Funded by the Medical Research Unit of Ringkøbing County, Denmark

Authors’ affiliations

J P Haahr, P Frost, J H Andersen, Department of Occupational Medicine, Herning Hospital, Herning, Denmark
S Østergaard, S Lausen, E A Holm, Department of Rheumatology and Physical Rehabilitation, Herning Hospital
J Dalsgaard, Department of Orthopaedic Surgery, Herning Hospital
K Norup, Department of Radiology, Herning Hospital

REFERENCES


www.annrheumdis.com

Haahr, Østergaard, Dalsgaard, et al.
Exercises versus arthroscopic decompression in patients with subacromial impingement: a randomised, controlled study in 90 cases with a one year follow up
J P Haahr, S Østergaard, J Dalsgaard, K Norup, P Frost, S Lausen, E A Holm and J H Andersen

Ann Rheum Dis 2005 64: 760-764
doi: 10.1136/ard.2004.021188