Conservative or surgical treatment for subacromial impingement syndrome? A systematic review

Oscar Dorrestijn, MD\textsuperscript{a,}\textdagger, Martin Stevens, PhD\textsuperscript{a}, Jan C. Winters, MD, PhD\textsuperscript{b}, Klaas van der Meer, MD, PhD\textsuperscript{b}, Ron L. Diercks, MD, PhD\textsuperscript{a}

\textsuperscript{a}Department of Orthopedic Surgery, University Medical Center Groningen, University of Groningen, The Netherlands
\textsuperscript{b}Department of General Practice, University Medical Center Groningen, University of Groningen, The Netherlands

**Background:** Patients with subacromial impingement syndrome are often operated on when conservative treatments fail. But does surgery really lead to better results than nonoperative measures? This systematic review compared effects of conservative and surgical treatment for subacromial impingement syndrome in terms of improvement of shoulder function and reduction of pain.

**Methods:** A literature search for randomized controlled trials (RCTs) in PubMed, EMBASE, PEDro, and the Cochrane Central Register of Controlled Trials was conducted. Two reviewers assessed the methodological quality of the selected studies. A best-evidence synthesis was used to summarize the results.

**Results:** Four RCTs were included in this review. Two RCTs had a medium methodological quality, and 2 RCTs had a low methodological quality. No differences in outcome between the treatment groups were reported for any of the studies, irrespective of quality.

**Conclusion:** No high-quality RCTs are available so far to provide possible evidence for differences in outcome; therefore, no confident conclusion can be made. According to the best-evidence synthesis, however, there is no evidence from the available RCTs for differences in outcome in pain and shoulder function between conservatively and surgically treated patients with SIS.

© 2009 Journal of Shoulder and Elbow Surgery Board of Trustees.

**Keywords:** Shoulder impingement syndrome; randomized controlled trials; conservative treatment; surgical treatment

Shoulder disorders are encountered frequently in general practice. A recently published review has summarized 18 studies on the prevalence of shoulder complaints in the general population in the United States, United Kingdom, Scandinavia, Cuba, South Africa, Spain, and Nigeria.\textsuperscript{26}

Prevalence figures ranged from 6.9\% to 26\% for point prevalence, 18.6\% to 31\% for 1-month prevalence, 4.7\% to 46.7\% for 1-year prevalence, and 6.7\% to 66.7\% for lifetime prevalence. In a Dutch study, the cumulative incidence of shoulder problems was estimated at 19/1000 patients per year in Dutch general practice.\textsuperscript{5} For the neck and upper extremity it was, after neck symptoms, the second most commonly presented musculoskeletal problem. A differentiation between several diagnoses of shoulder problems in general practice was presented in another Dutch study.\textsuperscript{20} The rate of subacromial impingement syndrome (SIS) was 44\% and was the most frequently recorded disorder.

This study was supported by a grant from University Medical Center Groningen, The Netherlands.

*Reprint requests: O. Dorrestijn, University Medical Center Groningen, Department of Orthopedic Surgery, PO Box 30.001, 9700 RB Groningen, The Netherlands.
E-mail address: o.dorrestijn@orth.umcg.nl (O. Dorrestijn).
Treatments of SIS always start conservatively. A broad spectrum of conservative treatments for SIS is available in primary health care: rest, nonsteroidal anti-inflammatory drugs (NSAIDs), corticosteroid injections, physical therapy, and manual therapy. Several randomized controlled trials (RCTs) have been conducted to gather evidence on effectiveness of different treatments for SIS and are summarized in systematic reviews, but until now, these reviews have only focused on shoulder problems in general and on nonoperative treatments for SIS.12,27 These reviews provide little evidence to support or refute the efficacy of common interventions for shoulder pain, such as physical therapy, manual therapy, acupuncture, NSAID medication, and corticosteroid injections. Subacromial corticosteroid injections for rotator cuff disease and intra-articular injections for adhesive capsulitis may be beneficial, although their effect may be small and not well maintained. Furthermore, there is limited evidence to support the efficacy of therapeutic exercise and manual therapy to treat SIS.

Although there is limited evidence for most conservative interventions, a retrospective study among 616 participants showed approximately 60% had satisfactory results after nonoperative treatment (NSAIDs or physical therapy, or both) at an average follow-up of 27 months.28 A therapeutic dilemma arises when these nonoperative treatments fail. The literature recommends referral to an orthopedic surgeon for an evaluation for (arthroscopic) subacromial decompression.31 Several publications report good-to-excellent results for open and arthroscopic subacromial decompression.22,23,25,33,35,37,41 However, are the results of surgery really better than those of conservative treatments? To answer whether surgery produces better results for SIS than conservative treatments, we performed a systematic review of RCTs to compare effects of conservative and surgical treatments for SIS in terms of improvement of shoulder function and reduction of pain.

Materials and methods

Literature search

A search of the literature in PubMed (from 1948 onward), EMBASE (from 1947 onward), PEDro (from 1992 onward), and the Cochrane Central Register of Controlled Trials was conducted to identify relevant publications until October 2007, without language restrictions. The literature search strategy for PubMed is presented in Table I. Reference lists of retrieved articles and articles on surgical interventions for SIS were screened for additional publications. Names of first authors of selected articles were used for citation tracking.

Study selection

The publications had to meet the following selection criteria:

- Study design: RCT for SIS. Studies focusing on surgical repair of rotator cuff tears, adhesive capsulitis, and shoulder instability were excluded.

- Participants: adult patients (> 18 years) with SIS manifest as pain upon abduction of the shoulder with a diagnosis confirmed with a positive result on an impingement test. For this test, the examiner injects lidocaine into the subacromial space and then repeats tests for the impingement sign (eg, Neer and Hawkins sign). Elimination or a significant reduction of pain constitutes a positive impingement test result. Furthermore, patients had to have been resistant to conservative treatments for at least 3 months.

- Interventions: all studies comparing (arthroscopic) subacromial decompression with conservative treatment.

- Outcome measures: all outcome measures for shoulder function or pain.

Two of the authors (M.S. and O.D.) used these criteria to independently select the relevant articles for this review by reading all titles and abstracts retrieved by the search strategy. In case of disagreements, a third reviewer (R.L.D.) was consulted.

Methodological quality assessment

All publications were assessed by 2 reviewers (M.S. and J.C.W.) according to a methodological quality list for the assessment of RCTs (Table II).14 Requirement of blinding patients or care providers to the intervention was excluded because this kind of blinding is not possible in this type of RCT. An item concerning blinding the outcome assessor was present. The questions on whether “outcome measures were suitable” and “the duration of follow-up was adequate to measure clinical differences between treatments” (items J and K) were added because they were considered relevant to measuring treatment effect.

Each criterion was graded as positive/yes (+), negative/no (−), or unclear (?). Disagreements were discussed in a consensus meeting.

<table>
<thead>
<tr>
<th>Table I</th>
<th>Literature search strategy for Medline</th>
</tr>
</thead>
<tbody>
<tr>
<td>StepSearch</td>
<td>Results</td>
</tr>
<tr>
<td>1 “Shoulder Impingement Syndrome” (MeSH) or shoulder impingement (TW)</td>
<td>786</td>
</tr>
<tr>
<td>2 “Shoulder Joint” (MeSH) and “Bursitis” (MeSH)</td>
<td>680</td>
</tr>
<tr>
<td>3 Subacromial impingement (TW)</td>
<td>783</td>
</tr>
<tr>
<td>4 Acromion (TW)</td>
<td>1071</td>
</tr>
<tr>
<td>5 Rotator Cuff (TW)</td>
<td>4079</td>
</tr>
<tr>
<td>6 1 or 2 or 3 or 4 or 5</td>
<td>5591</td>
</tr>
<tr>
<td>7 6 not capsulitis (TW) or frozen shoulder (TW)</td>
<td>5287</td>
</tr>
<tr>
<td>8 7 and “surgery” (SH) or “operative surgical procedures” (TW) or “surgical procedures, operative” (Mesh) or “Surgery” (Mesh) or surgery (TW) or “arthroscopy” (Mesh) or arthroscopy (TW) and “therapeutics” (Mesh) or therapeutics (TW) or “therapy” (SH) or therapy (TW)</td>
<td>2462</td>
</tr>
<tr>
<td>9 “Randomized Controlled Trials” (Mesh) or “Randomized Controlled Trial” (PT) or “Clinical Trial” (PT) or “Clinical Trials” (Mesh) or “Controlled Clinical Trial” (PT)</td>
<td>625,063</td>
</tr>
<tr>
<td>10 8 and 9</td>
<td>162</td>
</tr>
</tbody>
</table>

MeSH, Medical Subject Headings; PT, publication type; SH, subheading; TW, text word.
When no consensus could be reached, a third reviewer (R.L.D.) was asked for a binding verdict. An intraclass correlation coefficient was used to calculate the overall agreement between the 2 reviewers.

A quality score was calculated for the selected studies by summing the positive answers. Items E or G, or both, were only answered if, respectively, D or F, or both, were scored negatively. The maximum attainable score was 9.

**Results**

**Study selection**

The PubMed search resulted in 162 citations (Table I). One more citation was found in the Cochrane Register. No other studies were identified through the EMBASE or PEDro databases by hand search or citation tracking. The title or abstract, or both, was used to exclude 155 articles (Figure 1), and 8 were retrieved for a more detailed evaluation. Next, 2 RCTs were excluded for reasons of poster presentation and commentary. Six articles describing 4 RCTs met our inclusion criteria.42 The best-evidence synthesis was modified for purposes of this review, based on the method presented in another systematic review.36

Two articles7,8 were related to the same trial, of which 1 reported long-term outcomes (2.5-year follow-up). Only the short-term results were used for the best-evidence synthesis, given that the long-term outcomes were analyzed as a prognostic cohort study rather than an RCT and contained changes in methodology and analysis that hampered use of these data for the present review.

**Methodological quality**

The results of the quality assessment are presented in Table IV. The quality scores of the 4 trials ranged from 2 to 6. According to our cutoff points for quality, 2 trials were classified as medium quality and 2 as low quality (Table III).36

Studies were considered to be methodologically high quality when at least 7 items scored positively; a score of 4 to 6 was medium quality and 0 to 3 was low quality.

**Data extraction and analysis**

Using standardized forms, 2 reviewers (M.S. and J.C.W.) independently extracted data from the selected studies on characteristics of the study population, description and standardization of interventions, outcome measures, and results.

Extraction of results focused on obtaining risk ratios and their respective confidence intervals for dichotomous data or means (or median scores) with standard deviations, and differences in means (or median scores) and their confidence intervals for continuous outcomes. When not given, these descriptive data were calculated if sufficient data were available. The intention was to perform a quantitative analysis (meta-analysis). However, meta-analysis was not possible because of the diversity in outcome measures among the included studies and the different and sometimes incomplete presentation form (median scores, mean scores, relative risk ratios). Efforts to retrieve raw data or means and their standard deviations to compute effect sizes by contacting the authors of the different articles were unsuccessful. We therefore chose to summarize the results by means of a qualitative analysis (best-evidence synthesis). Guidelines for systematic reviews from the Cochrane Collaboration Back Review Group were used.42 The best-evidence synthesis was modified for purposes of this review, based on the method presented in another systematic review (Table III).36

The overall agreement between the 2 reviewers for the 11 items applied to the 4 trials was quite good (Cohen κ coefficient, 0.66 ± 0.09 [SE]). Disagreements between the 2 observers arose in 1 trial for item D,7,8 in 1 trial for item F,7,8 in 3 trials for item I7,8,19,20,31,32 and in 3 trials for item J7,8,31,32 After the consensus meeting, consensus between the 2 reviewers was unclear in 56% of these items.

<table>
<thead>
<tr>
<th>Item</th>
<th>Quality variable</th>
<th>Ratinga</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Was the treatment allocation randomized?</td>
<td>++/−/?</td>
</tr>
<tr>
<td>B</td>
<td>Was the treatment allocation concealed?</td>
<td>++/−/?</td>
</tr>
<tr>
<td>C</td>
<td>Was the outcome assessor blinded to the intervention?</td>
<td>++/−/?</td>
</tr>
<tr>
<td>D</td>
<td>Were the groups similar at baseline regarding the most important prognostic indicators?</td>
<td>++/−/?</td>
</tr>
<tr>
<td>E</td>
<td>If not, were adjustments made in the analysis for differences of prognostic indicators at baseline and/or for confounding variables?</td>
<td>++/−/?</td>
</tr>
<tr>
<td>F</td>
<td>Was a sufficient proportion (≥ 80%) of included patients available for the full length of follow-up?</td>
<td>++/−/?</td>
</tr>
<tr>
<td>G</td>
<td>If not, was selective loss to follow-up excluded?</td>
<td>++/−/?</td>
</tr>
<tr>
<td>H</td>
<td>Was an intention-to-treat analysis included?</td>
<td>++/−/?</td>
</tr>
<tr>
<td>I</td>
<td>Were cointerventions avoided or similar?</td>
<td>++/−/?</td>
</tr>
<tr>
<td>J</td>
<td>Were the outcome measures suitable to measure clinically relevant differences in treatment effects?</td>
<td>++/−/?</td>
</tr>
<tr>
<td>K</td>
<td>Was the duration of follow-up adequate to measure clinical differences between treatments (≥ 1 year)?</td>
<td>++/−/?</td>
</tr>
</tbody>
</table>

a Variables were rated as positive/yes (+), negative/no (−), unclear (?) or not applicable (NA).

**Table II** Methodological quality list

The results of the quality assessment are presented in Table IV. The quality scores of the 4 trials ranged from 2 to 6. According to our cutoff points for quality, 2 trials were classified as medium quality and 2 as low quality (Table V).
Data extraction and analysis

Table V presents the characteristics of the selected studies, including a description of interventions, population characteristics, treatment effect, follow-up period, and study quality. All studies randomized participants between physiotherapeutic regimens and subacromial decompression. A similarity of all the physiotherapeutic regimens applied was the focus on strengthening the rotator cuff and the scapular stabilizing muscles. In one trial,7,8 participants started with relaxed repetitive movements; in another,20 with application of heat, cold packs, or soft-tissue treatments. In one study,32 strength training followed education about the shoulder problem and unloaded movements. In the last trial,31 participants in the conservative group were hospitalized for 2 weeks, in contrast to the other studies. During this period the participants received intensive physiotherapy training supported with NSAIDs and corticosteroid injections. One trial7,8 had added a placebo group.

In the presented studies, most participants improved through either conservative treatment or surgery. In 1 study, the physiotherapy group improved 23.0 (95% confidence interval [CI], 16.9-29.1) in mean Constant score (range, 0-100) from 34.7 (2.2 SME) at baseline; the surgery group improved 18.8 (95% CI, 11.5-26.1) from 33.7 (2.3 SME).20 In another RCT, the median Neer score at entry was 67.5 for the physiotherapy group, 64.0 for the surgery group, and 65.5 for the placebo group.8 After 6 months of follow-up, the
<table>
<thead>
<tr>
<th>Study</th>
<th>Interventions (number of patients)</th>
<th>Population characteristics</th>
<th>Treatment effecta (95% CI) Follow-up</th>
<th>Study quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haahr²⁹ (2005/6)</td>
<td>1. Arthroscopic subacromial decompression (41) 2. Supervised exercises (43)</td>
<td>1. F/M: 29/12; mean age: 44.3 (SEM 1.3); DoC: &lt;6 mon: 4; 6-12 mon: 3; &gt;1 y: 34 2. F/M: 29/14; mean age: 44.5 (SEM 1.2); DoC: &lt;6 mon: 3; 6-12 mon: 10; &gt;1 y: 29.</td>
<td>Constant score (12 mon) (0-100), mean change (CI): 1. 18.8 (11.5 to 26.1); 2. 23.0 (16.9-29.1) SMD = −0.003 (−0.010 to 0.004)</td>
<td>12 mon/4-8 y Medium (67%)</td>
</tr>
<tr>
<td>Brox⁸ (1993)</td>
<td>1. Arthroscopic subacromial decompression (45) 2. Supervised exercises and education (50) 3. Detuned soft laser treatment (30)</td>
<td>1. F/M: 16/29; mean age: 48; DoC: &lt;6 mon: 8; 6-12 mon: 8; 1-3 y: 9; &gt;3 y: 20 2. F/M: 28/32; mean age: 47; DoC: DoC: &lt;6 mon: 6; 6-12 mon: 6; 1-3 y: 13; &gt;3 y: 25. 3. F/M: 15/15; mean age: 48; DoC: &lt;6 mon: 5; 6-12 mon: 5; 1-3 y: 5; &gt;3 y: 14.</td>
<td>No differences between groups. PRIM score (4-8 y) (0-36), mean change (CI): 1. 9.1 (5.5 to 12.6); 2. 11.4 (8.7 to 14.1) SMD = 2.4 (−2 to 6.8).</td>
<td>6 mon No differences between groups. 6 mon Medium (67%)</td>
</tr>
<tr>
<td>Rahme³² (1998)</td>
<td>1. Open subacromial decompression (21) ± rotator cuff repair (5) 2. Physiotherapy and education (18)</td>
<td>1. &amp; 2.: F/M 23/19; mean age: 42 (range 28-63); DoC: almost 4 y avg.</td>
<td>Group 1: 16/21 (76%; RR₁→₂A = 1.1; RR₁→₂B = 1.1). Group 2A: 4/6 (67%). Group 2B (operated on): 7/12 (58%). No differences between groups. When those who were operated on or were lost to follow-up in group 2 were considered as failed; success for group 2C: 4/21 (RR₁→₂C = 4; P &lt; .0005). These data were excluded from the synthesis.</td>
<td>12 mon Low (33%)</td>
</tr>
</tbody>
</table>

(continued on next page)
With just low-to-medium quality, a small group of 323 patients in a small number of trials (4) failed conservative treatment of SIS. The confidence interval (CI) is expanded in the trial. Treatment effect (TE) is calculated by using the proportion of "success" or "failure." The CI is presented as the 95% CI, whereas the standard error of the mean (SEM) is presented as the standard deviation (SD).

### Discussion

Failed conservative treatment of SIS is often followed by surgery. This systematic review was designed to determine if the results of surgery for SIS are better than those of conservative treatment. According to the best-evidence synthesis, we conclude that the RCTs provide no evidence for differences in outcome in pain and shoulder function between conservatively and surgically treated patients with SIS. This conclusion is based on a relatively small group of 323 patients from a small number of trials (4) with just low-to-medium quality.

#### Validity of the trials

Failed conservative treatment of SIS is often followed by surgery. This systematic review was designed to determine if the results of surgery for SIS are better than those of conservative treatment. According to the best-evidence synthesis, we conclude that the RCTs provide no evidence for differences in outcome in pain and shoulder function between conservatively and surgically treated patients with SIS. This conclusion is based on a relatively small group of 323 patients from a small number of trials (4) with just low-to-medium quality.}

### Table V (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Interventions (number of patients)</th>
<th>Population characteristics</th>
<th>Treatment effect (95% CI)</th>
<th>Follow-up</th>
<th>Study quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peters31 (1997)</td>
<td>Open (n = 17) or arthroscopic (n = 15) subacromial decompression</td>
<td>F/M: 14/18; mean age: 49 (range 37-61); DoC: not given.</td>
<td>Differences between median scores at baseline and at maximum follow-up for Subjective Shoulder Rating Scale: Group 1: 30; Group 2: 15.</td>
<td>Low (22%)</td>
<td>48 men</td>
</tr>
<tr>
<td></td>
<td>2. Two-week hospital stay: physiotherapy supported with NSAIDs and corticosteroid injections (10)</td>
<td>F/M: 12/28; mean age: 56 (range 37-82); DoC: not given.</td>
<td>No statistical analysis for significance performed.</td>
<td>Low (38%)</td>
<td>3.98</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study</th>
<th>Interventions (number of patients)</th>
<th>Population characteristics</th>
<th>Treatment effect (95% CI)</th>
<th>Follow-up</th>
<th>Study quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peters31 (1997)</td>
<td>Open (n = 17) or arthroscopic (n = 15) subacromial decompression</td>
<td>F/M: 14/18; mean age: 49 (range 37-61); DoC: not given.</td>
<td>Differences between median scores at baseline and at maximum follow-up for Subjective Shoulder Rating Scale: Group 1: 30; Group 2: 15.</td>
<td>Low (22%)</td>
<td>48 men</td>
</tr>
<tr>
<td></td>
<td>2. Two-week hospital stay: physiotherapy supported with NSAIDs and corticosteroid injections (10)</td>
<td>F/M: 12/28; mean age: 56 (range 37-82); DoC: not given.</td>
<td>No statistical analysis for significance performed.</td>
<td>Low (38%)</td>
<td>3.98</td>
</tr>
</tbody>
</table>
The studies failed to reach a high-quality classification because most did not score positively on items B (“treatment allocation concealed”), C (“outcome assessor blinded”), D (“groups similar at baseline?”), E (“co-interventions similar?”), and J (“outcome measures clinically relevant?”). In the quality assessment, most of the discussion concerned item J (“Were the outcome measures suitable to measure clinically relevant differences in treatment effects?”). As outlined in an extensive review on shoulder disorders, there is no gold standard that provides a valid and reliable estimate for clinically relevant changes in any subgroup of patients with shoulder disorders.\(^{33}\) Few published studies can be found describing validity, reproducibility, responsiveness, or interpretability of the outcome measures used in the presented trials.

Item I (“Were co-interventions avoided or similar?”) also led to discussion because it was not always described clearly in the articles. The same goes for item B (“Treatment allocation concealed”). Several studies show empiric evidence that inadequate concealment of treatment allocation is associated with bias.\(^{10,24,34}\) An inadequate description of randomization procedures does not automatically mean bias was present, but it cannot be excluded.

Another potential source of bias was caused by the way the data analysis was done. Two studies incorrectly transformed their data for an intention-to-treat analysis, which violates the principles of the method. The actual outcome scores should have been used for the patients in the conservative group who had been operated on instead of assigning them the lowest available score as if they had failed. The intention-to-treat approach is often inadequately applied, which has also been noted in a survey of RCTs published in 4 major medical journals.\(^{21}\) This inadequate use of the intention-to-treat approach is a potential source of bias.

Bias could also have been caused by the differences in treatments between the intervention and the control groups. Blinding the care provider and the participant to the intervention can prevent such bias, but this would not have been possible in the presented RCTs.

Another important aspect is the heterogeneity of treatments of the different studies, which makes it difficult to compare them. Two trials used an arthroscopic technique to perform subacromial decompressions,\(^{7,8,19,20}\) 1 study used open surgery,\(^{32}\) and in another both were used.\(^{31}\) Both methods seem to result in adequate subacromial decompression,\(^{22,23,25}\) but according to some studies, the arthroscopic method results in earlier restoration in active range of motion and a quicker return to work.\(^{25,33}\) In 1 study, coexistent rotator cuff ruptures were also sutured.\(^{32}\) This of course makes it impossible to do a comparison with results of other studies. Furthermore, in 1 trial, participants in the conservative group were hospitalized for 2 weeks.\(^{31}\) This is not and will not become a common treatment method due to its significant economic and health implications.

Most participants in these RCTs had symptoms longer than a year and were resistant to previous conservative treatment and, therefore, were probably in favor of being assigned to surgery because the previous conservative treatment was not effective. This might have led to a source of bias for the surgery groups. By contrast, high expectations of surgery can lead to disappointing results, even more when there are side effects (eg, postoperative stiffness of the shoulder). Despite these possible biases, this did not result in significant differences between the study groups.

All 4 trials together scored 17 times positively, 14 times unclear, 7 times negatively, and 6 times not applicable for the different items in the methodological assessment, for a total of 44. Although for practical reasons certain methodological concessions can be made (eg, not blinding an outcome assessor), a great gain in quality could be achieved by a clear and full presentation of the study design.

Effectiveness of treatment

No confident conclusion can be made based on the results available. The RCTs included in this review failed to provide evidence for differences in outcome between conservatively and surgically treated patients with SIS. Whether this failure is due to impairments in methodological quality or a lack of difference in treatment outcome remains unclear. For several decades, patients with SIS have been operated on when conservative treatments failed. In that respect, observational studies have reported satisfactory results in 67% to 90% of such patients.\(^{22,33,35,41}\) However, the results of this review show that no conclusion can be made about whether surgery is better than conservative treatment.

Limitations

A major limitation of this review is that only 4 RCTs have been conducted for such a common shoulder disorder like SIS. A possible explanation for this could be that patients with chronic SIS do not want to risk being randomized to a nonoperative treatment after extensive previous conservative treatments. In addition, no data on cost-effectiveness of treatments or sick leaves are available from the 4 RCTs. This information is indispensable for the decision-making process of care providers. For example, in the short-term, surgery is more expensive than conservative treatments, but it can be more cost-effective than conservative treatments with a shorter patient sick leave.

Recommendations

To answer the question of whether surgery for SIS is indeed more effective than conservative treatment, high-quality trials are needed. These trials should use outcome measures that quantify improvement of shoulder function and
reduction of pain that are valid, reliable, and responsive in these study populations. Correct tests, such as the impingement test, should be used to diagnose patients with SIS, and strict inclusion and exclusion criteria should be observed to create homogenous study groups. Participants should present with a certain minimum severity of SIS to be potentially responsive to the study treatments. Proper power analysis is needed to determine sample size. Follow-up should be at least 1 year, and it would be important for studies to provide data on cost-effectiveness.

Furthermore, future trials should also take duration of symptoms into account. There is a trend for an earlier indication for surgery. Several observational studies report a significantly better outcome in operated-on patients who had not responded to nonoperative measures and who had a short symptom duration compared with those who had prolonged symptoms before surgery. So far, however, no RCTs have focused on duration of symptoms. Future RCTs on patients with SIS should therefore also investigate the influence of a shorter-than-usual preoperative duration of symptoms compared with usual medical care.

Acknowledgments

We thank Truus van Ittersum for her help in developing search strategies for the electronic databases used. We are also grateful to Daniëlle van der Windt for her comments. This study is supported by a grant from University Medical Center Groningen, The Netherlands.

References