Effectiveness of early versus delayed rehabilitation following rotator cuff repair: systematic review and meta-analyses

Bruno Mazuquin1*, Maria Moffatt1, Peter Gill1,2, James Selfe1, Jonathan Rees3, Steve Drew4, Chris Littlewood1

1 Department of Health professions, Faculty of Health, Psychology and Social care, Manchester Metropolitan University, Manchester, United Kingdom.

2 Northern Care Alliance NHS Group, Manchester, United Kingdom.

3 Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Science, University of Oxford and NIHR Oxford Biomedical Research Centre, Oxford, United Kingdom.

4 University Hospitals Coventry and Warwickshire, Coventry, United Kingdom.

* Corresponding author

Email: b.mazuquin@mmu.ac.uk

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
ABSTRACT

Objective: To investigate the effectiveness of early rehabilitation compared with delayed/standard rehabilitation after rotator cuff repair for pain, function, range of movement, strength, and repair integrity.

Design: Systematic review and meta-analyses.

Methods: We searched databases and included randomised controlled trials (RCTs) comparing early with delayed/standard rehabilitation for patients undergoing rotator cuff repair surgery. We assessed risk of bias of the RCTs using the Cochrane RoB 2 tool.

Results: Twenty RCTs, with 1841 patients, were included. The majority of the RCTs were of high or unclear risk of overall bias. We found substantial variations in the rehabilitation programmes, time in the sling and timing of exercise progression. We found no statistically significant differences for pain and function at any follow-up except for the outcome measure Single Assessment Numeric Evaluation at six months (MD: 6.54; 95%CI: 2.24-10.84) in favour of early rehabilitation. We found statistically significant differences in favour of early rehabilitation for shoulder flexion at six weeks (MD: 7.36; 95%CI: 2.66-12.06), three (MD: 8.45; 95%CI: 3.43-13.47) and six months (MD: 3.57; 95%CI: 0.81-6.32) and one year (MD: 1.42; 95%CI: 0.21-2.64) and similar differences for other planes of movement. In terms of repair integrity at one year, regarding early mobilisation, the number needed to harm was 651.

Discussion: Current approaches to early mobilisation, based largely on early introduction of passive movement, did not demonstrate significant differences in most clinical outcomes, although we found statistically significant differences in
favour of early rehabilitation for range of movement. Importantly, there were no
differences in repair integrity between early and delayed/standard rehabilitation.
Most rehabilitation programmes did not consider early active movement as soon
as the patient feels able. With ongoing variation in rehabilitation protocols there
remains a need for large high quality RCT to inform the optimal approach to
rehabilitation after rotator cuff repair surgery.
INTRODUCTION

Shoulder pain is experienced by one in four people at any one time (1) and is one of the most common musculoskeletal pain presentations (2). Tears of the rotator cuff, the muscles and tendons that closely surround the shoulder joint, are a common cause of shoulder pain (1). Many people with symptomatic rotator cuff tears can be successfully managed non-surgically; where this treatment proves insufficient, surgery to repair the torn rotator cuff might be offered (3). However, as the number of operations to repair the rotator cuff increases and surgical techniques advance, there are ongoing uncertainties about the optimal approach to postoperative rehabilitation, a key component of the recovery process (4).

A recent international survey of practice reported that postoperative rehabilitation has not evolved over the last two decades (5). A generally cautious approach to standard postoperative rehabilitation still remains, and includes sling immobilisation for several weeks (6). In 2018 a systematic review (7) was published that reported conflicting evidence in relation to early versus delayed/standard rehabilitation following rotator cuff repair. The meta-analyses suggested that early mobilisation did not provide additional clinical benefit in terms of pain or function, although recovery of range of movement of movement was more rapid. It also reported no statistically significant difference in repair integrity, which is one of the historical concerns for this surgery and the justification for cautious approaches to postoperative rehabilitation. However, most of the randomised controlled trials (RCTs) included in the review were rated as presenting a high risk of bias and data for large tears were lacking.
A number of new RCTs (eight) have been published recently providing the opportunity for an up to date evidence. Our aim was to investigate the effectiveness of early compared to standard or delayed rehabilitation following rotator cuff repair for pain, function, range of movement, strength and repair integrity.

METHODS

We reported the systematic review according to the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) (8). The protocol was registered in the International prospective register of systematic reviews database (PROSPERO) - PROSPERO 2020 CRD42020209330.

Eligibility criteria

We included studies that met the following criteria:

- Design: RCT

- Participants: patients aged 18 years or older who had undergone surgical repair of the rotator cuff

- Intervention and comparison: early rehabilitation compared with delayed rehabilitation (as per study definitions).

- Outcomes: pain, function, range of movement, strength and repair integrity.

Search

We searched MEDLINE, CINAHL, Scopus, SPORTDiscus, SciELO and the Cochrane Library for relevant papers from inception to December 2020. We
decided not to limit the searches by date of publication to identify other RCTs that could be relevant but were not included in the previous version. The electronic search strategies were supplemented by hand searching the reference lists and citations of the included RCTs. There was no restriction to date or language of publication. For the search strategies, we combined MeSH terms and keywords such as: rotator cuff, shoulder, shoulder joint, rehabilitation, physiotherapy, immobilisation and RCT. The detailed search strategy is available in the supplementary file 1.

**Screening**

Searches results were imported to Rayyan (https://rayyan.qcri.org) where removal of duplicates and screening for eligibility was undertaken by two independent authors (BM and MM). Studies were first screened by their titles and abstracts; the full text was retrieved if further information was needed for a decision.

**Data extraction**

Data was extracted by one author (BM) and reviewed by a second author (PG) using a pre-established Excel template. Data extraction included author names, year of publication, country, participants characteristics, characteristics of the intervention and comparator, tools used for outcomes assessment, results for the variables of interest and public and patient involvement and engagement activities. In case of missing data, we contacted the authors via email to request additional available data.

**Risk of bias and grading of evidence**
The risk of bias of the included RCTs was assessed by one author (BM) and reviewed by a second author (MM) using the Cochrane risk of bias tool (RoB2) (9). The Grades of Recommendation, Assessment, Development and Evaluation (GRADE) framework and GRADEpro GDT (10) software were used to rate the certainty of the effect. Outcomes were rated and downgraded according to the presence or absence of factors (risk of bias, inconsistency, indirectness, imprecision) affecting the quality of the body of RCTs included in each outcome.

**Measurement of treatments effect**

We generated meta-analyses and forest plots using RevMan 5 (The Nordic Cochrane Centre, Copenhagen, Denmark)(11). The meta-analyses were presented according to the different outcome measures used, for example American Shoulder and Elbow Score (ASES), and follow-up timing. Continuous data was expressed as mean difference and 95% confidence interval. For dichotomous outcomes, the odds ratio was used with 95% confidence interval. As re-tear or failure to heal, indicative of repair integrity, is regarded as an unfavourable outcome, we also calculated the Absolute Risk Increase and Number Needed to Harm (NNH) for the repair integrity data. Statistical heterogeneity was tested using the $I^2$ test. In addition, we observed variation in the rehabilitation protocols for exercise dosage and time in the sling; therefore, we used the random effects methods for the meta-analyses. Funnel plots were not created to check for heterogeneity and bias as this is not recommended where meta-analyses have fewer than 10 RCTs (12, 13).
RESULTS

Study selection

Initially 2142 records were found, after removing duplicates 1805 studies were screened. Of these, 22 were selected for full text review. Another two studies were excluded due to wrong interventions: Baumgarten, Osborn (14) investigated the use of pulley exercises at six weeks and Michael, König (15) explored the effect of continuous passive movement and physiotherapy versus physiotherapy alone. Twenty RCTs were eligible for inclusion; in total, 13 RCTs were included in the meta-analyses (Figure 1).

Figure 1. PRISMA diagram for studies selection.

Study characteristics

Participants mean age ranged from 54 to 65.4 years. Sample sizes varied from 18 to 206 participants; only nine out of the 20 RCTs (45%) reported a sample size calculation and the majority, 16/20 (80%), recruited patients from a single hospital. The total number of patients from the 20 RCTs was 1841. Only eight of 20 RCTs included patients with large tears (3, 16-22). There was variation in the use of single or double row repairs as well as additional surgical procedures such as subacromial decompression, acromioclavicular joint excision and biceps tenodesis. Further details are available in supplementary file 2.

Risk of bias
Figure 2 summarises the results of the risk of bias assessment. Supplementary file 3 reports the risk of bias assessment of each RCT. Most RCTs were rated as of unclear or high risk of bias for randomisation process and overall bias.

**Figure 2.** Risk of bias summary.

**Rehabilitation protocols**

The method of immobilisation after rotator cuff repair was variable. The majority of the RCTs (3, 19, 20, 23-30)(11/20, 55%) reported the use of a standard sling, six RCTs (18, 21, 22, 31-33) used a sling with an abduction component and one RCT (34) used a standard sling for the early group and a sling with an abduction component for the delayed group. The time in the sling was also variable; eight RCTs (40%) (18, 21, 23-26, 29, 32) used a six-week period for both groups. Three RCTs prescribed a sling for the early group to be used for comfort only and discarded as the patient felt appropriate (3, 19, 20). One RCT did not prescribe a sling for those in the early group (30).

Exercise progression was similar across RCTs, starting with passive exercises, moving to active-assisted, active and then resisted exercise. However, there was variation in the timing for the exercise progression and information about frequency and intensity of exercises. All studies had a time driven protocol at all stages except Littlewood, Bateman (3), Sheps, Bouliane (19), Sheps, Silveira (20). Littlewood, Bateman (3) used a patient-led approach using acceptable symptom response to prescribe and progress exercises, regardless of the postoperative time. Sheps, Bouliane (19) used a similar protocol to Littlewood,
Bateman (3); however, there is limited information about the protocol after six weeks postoperative. Sheps, Silveira (20) also used a patient-led approach but only pain-free active movements were permitted but exercise progression was based on the number of weeks postoperative. More details about the rehabilitation programmes are available in supplementary file 4.

Outcomes

Pain at rest

Meta-analyses for pain at rest, measured with the visual analogue scale, were possible at six weeks, three and six months, and one- and two-years follow-up. No statistically significant differences were found for any comparisons (Table 1). Figures 3 and 4 shows the forest plots at six weeks and three months, respectively. All forest plots for meta-analyses in Tables 1-4 that are not shown in the main text are available in supplementary file 5. The GRADE summary of findings with reasons for downgrading the certainty of effect is available in supplementary file 6.
Table 1. Meta-analyses of pain at rest by visual analogue scale (A negative effect estimate favours early rehabilitation).

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Follow-up</th>
<th>Number of studies</th>
<th>Total sample size</th>
<th>Effect estimate MD [95% CI]</th>
<th>P value</th>
<th>GRADE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual Analogue Scale</td>
<td>6 weeks</td>
<td>6</td>
<td>707</td>
<td>0.39 [-1.35, 2.13]</td>
<td>0.66</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>3 months</td>
<td>6</td>
<td>692</td>
<td>-0.04 [-0.36, 0.29]</td>
<td>0.83</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>6 months</td>
<td>7</td>
<td>722</td>
<td>-0.06 [-0.30, 0.18]</td>
<td>0.62</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>1 year</td>
<td>4</td>
<td>521</td>
<td>-0.10 [-0.35, 0.15]</td>
<td>0.45</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>2 years</td>
<td>4</td>
<td>551</td>
<td>0.11 [-0.12, 0.35]</td>
<td>0.34</td>
<td>High</td>
</tr>
</tbody>
</table>

MD: mean difference, CI: confidence interval

1Scale from 0-10, a lower value is a better outcome.

Figure 3. Meta-analysis of pain level at six weeks.

Figure 4. Meta-analysis of pain level at three months.

Function

There was variation in the use of composite measures of shoulder pain and function and outcome measures for shoulder function. Given the wide range of data available, we combined studies according to the outcome measure used. Meta-analyses were possible for the ASES score (three and six months, and one and two years), Constant-Murley score (CM) (three and six months, and one and two years), Single Assessment Numeric Evaluation (SANE) (three and six months), Simple Shoulder Test (three and six months, and one year) and Western Ontario Rotator Cuff Index (three and six months, and one year).

There were no statistically significant differences for any outcome measures at any follow-up except for the SANE at six months in favour of early rehabilitation (Table 2). Overall, the mean differences were small in the short- and long-term follow-ups as illustrated in Figures 5 and 6 for the ASES. Other outcome
measures used by RCTs included the Disabilities of the Arm, Shoulder and Hand, Oxford Shoulder Score, Shoulder Pain and Disability Index and University of California Los Angeles shoulder rating scale. The certainty of effects ranged from very low to moderate (Table 2).

Table 2. Meta-analyses of function by outcome measures (A positive effect estimates favours early rehabilitation).

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Follow-up</th>
<th>Number of studies</th>
<th>Total sample size</th>
<th>Effect estimate MD [95% CI]</th>
<th>P value</th>
<th>GRADE</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Shoulder and Elbow score</td>
<td>3 months</td>
<td>3</td>
<td>243</td>
<td>3.43 [-1.07, 7.92]</td>
<td>0.14</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>6 months</td>
<td>4</td>
<td>365</td>
<td>-0.26 [-4.76, 4.25]</td>
<td>0.91</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>1 year</td>
<td>4</td>
<td>441</td>
<td>-0.57 [-2.39, 1.25]</td>
<td>0.54</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>2 years</td>
<td>2</td>
<td>202</td>
<td>-2.67 [-6.35, 1.02]</td>
<td>0.16</td>
<td>Moderate</td>
</tr>
<tr>
<td>Constant-Murley score</td>
<td>3 months</td>
<td>4</td>
<td>313</td>
<td>3.18 [-1.53, 7.90]</td>
<td>0.19</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>6 months</td>
<td>6</td>
<td>513</td>
<td>1.19 [-1.33, 3.71]</td>
<td>0.35</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>1 year</td>
<td>5</td>
<td>559</td>
<td>-1.18 [-2.62, 0.25]</td>
<td>0.11</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>2 years</td>
<td>2</td>
<td>202</td>
<td>-1.80 [-5.10, 1.49]</td>
<td>0.28</td>
<td>Moderate</td>
</tr>
</tbody>
</table>
**Table 2 (continue).** Meta-analyses of function by outcome measures.

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Follow-up</th>
<th>Number of studies</th>
<th>Total sample size</th>
<th>Effect estimate (MD [95% CI])</th>
<th>P value</th>
<th>GRADE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Assessment Numeric Evaluation¹</td>
<td>3 months</td>
<td>2</td>
<td>138</td>
<td>2.23 [-4.62, 9.07]</td>
<td>0.52</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>6 months</td>
<td>2</td>
<td>138</td>
<td>6.54 [2.24, 10.84]</td>
<td>0.003</td>
<td>Moderate</td>
</tr>
<tr>
<td>Simple Shoulder Test²</td>
<td>3 months</td>
<td>2</td>
<td>163</td>
<td>-0.77 [-2.82, 1.28]</td>
<td>0.46</td>
<td>Very low</td>
</tr>
<tr>
<td></td>
<td>6 months</td>
<td>3</td>
<td>277</td>
<td>0.63 [-0.36, 1.62]</td>
<td>0.21</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>1 year</td>
<td>3</td>
<td>277</td>
<td>0.39 [-0.40, 1.18]</td>
<td>0.33</td>
<td>Moderate</td>
</tr>
<tr>
<td>Western Ontario Rotator Cuff Index¹</td>
<td>3 months</td>
<td>2</td>
<td>309</td>
<td>-1.82 [-5.96, 2.32]</td>
<td>0.39</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>6 months</td>
<td>2</td>
<td>305</td>
<td>-1.29 [-5.17, 2.59]</td>
<td>0.52</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>1 year</td>
<td>2</td>
<td>300</td>
<td>-1.91 [-5.21, 1.40]</td>
<td>0.26</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

MD: mean difference, CI: confidence interval

¹Scale from 0-100, a higher value is a better outcome.

²Scale from 0-12, a higher value is a better outcome.

**Figure 5.** Forest plot of function at 3 months by American Shoulder and Elbow Surgery score.
Figure 6. Forest plot of function at 1 year by American Shoulder and Elbow Surgery score.

Range of movement

Meta-analyses were possible for shoulder flexion, abduction, external rotation at 90° of abduction and internal rotation at 90° of abduction at six weeks, three and six months, and one and two years. All measurements were made using a goniometer. Statistically significant differences were found for flexion at six weeks, three and six months and one year, abduction at six weeks, external rotation at three and six months and internal rotation at six weeks, three and six months (Table 3). The certainty of effects ranged from very low to high.
Table 3. Meta-analyses of range of movement by movement (A positive effect estimates favours early rehabilitation).

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Follow-up</th>
<th>Number of studies</th>
<th>Total sample size</th>
<th>Effect estimate</th>
<th>P value</th>
<th>GRADE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flexion</td>
<td>6 weeks</td>
<td>6</td>
<td>753</td>
<td>7.36 [2.66, 12.06]</td>
<td>0.002</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>3 months</td>
<td>10</td>
<td>1030</td>
<td>8.45 [3.43, 13.47]</td>
<td>0.001</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>6 months</td>
<td>12</td>
<td>1275</td>
<td>3.57 [0.81, 6.32]</td>
<td>0.01</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>1 year</td>
<td>9</td>
<td>1062</td>
<td>1.42 [0.21, 2.64]</td>
<td>0.02</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>2 years</td>
<td>4</td>
<td>543</td>
<td>1.61 [-1.11, 4.33]</td>
<td>0.25</td>
<td>High</td>
</tr>
</tbody>
</table>

| Abduction       | 6 weeks   | 4                 | 615              | 6.82 [2.30, 11.33] | 0.003   | Moderate |
|                 | 3 months  | 5                 | 581              | 6.68 [-1.47, 14.83] | 0.11    | Low    |
|                 | 6 months  | 5                 | 574              | 1.14 [-2.78, 5.05]  | 0.57    | Low    |
|                 | 1 year    | 4                 | 529              | 0.05 [-4.04, 4.14]  | 0.98    | Moderate |
|                 | 2 years   | 2                 | 341              | -1.73 [-7.41, 3.94] | 0.55    | Moderate |
**Table 3 (continue).** Meta-analyses of range of movement by movement.

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Follow-up</th>
<th>Number of studies</th>
<th>Total sample size</th>
<th>Effect estimate MD [95% CI]</th>
<th>P value</th>
<th>GRADE</th>
</tr>
</thead>
<tbody>
<tr>
<td>External rotation at 90° abduction</td>
<td>6 weeks</td>
<td>5</td>
<td>633</td>
<td>2.06 [-2.65, 6.76]</td>
<td>0.39</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>3 months</td>
<td>8</td>
<td>805</td>
<td>8.11 [3.85, &lt;0.001]</td>
<td>&lt;0.001</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>6 months</td>
<td>9</td>
<td>964</td>
<td>1.77 [-0.05, 3.60]</td>
<td>0.06</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>1 year</td>
<td>7</td>
<td>839</td>
<td>0.76 [-2.01, 3.53]</td>
<td>0.59</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>2 years</td>
<td>3</td>
<td>461</td>
<td>-0.70 [-5.69, 4.28]</td>
<td>0.78</td>
<td>Moderate</td>
</tr>
<tr>
<td>Internal rotation at 90° abduction</td>
<td>6 weeks</td>
<td>3</td>
<td>495</td>
<td>3.34 [0.40, 6.28]</td>
<td>0.03</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>3 months</td>
<td>4</td>
<td>461</td>
<td>8.19 [0.99, 15.39]</td>
<td>0.03</td>
<td>Very low</td>
</tr>
<tr>
<td></td>
<td>6 months</td>
<td>5</td>
<td>620</td>
<td>3.60 [0.06, 7.13]</td>
<td>0.05</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>1 year</td>
<td>4</td>
<td>580</td>
<td>1.26 [-1.64, 4.15]</td>
<td>0.39</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>2 years</td>
<td>2</td>
<td>341</td>
<td>0.54 [-2.39, 3.46]</td>
<td>0.72</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

1 Scale from 0-180, a higher value is a better outcome.
2 Scale from 0-90, a higher value is a better outcome.
3 Scale from 0-70, a higher value is a better outcome.
**Strength**

Nine RCTs reported muscle strength (18-21, 23, 26-29). Due to heterogeneity of testing position and data reporting, we did not pool the data into meta-analyses. Six RCTs used a hand-held dynamometer (18, 20, 21, 26, 27, 29); other instruments included isokinetic dynamometer, tensiometer and the strength composite of the CM. The follow-up assessment ranged from six weeks to two years. Muscle strength was tested for flexion (five RCTs)(18, 20, 21, 26, 27), abduction (four RCTs)(19-21, 26), internal (four RCTs)(18, 21, 27, 28) and external rotation (six RCTS) (18, 21, 26-29). All RCTs reported that both groups improved in strength; another consistent finding across all RCTs was that no statistically significant differences between group were found for any strength test at any follow-up assessed.

**Repair integrity**

Meta-analyses were possible for three and six months and one-year follow-up. Only Arndt, Clavert (23) used a CT arthroscopy to assess the repair integrity, all other RCTs used ultrasound or MRI scan. There were no statistically significant differences between groups at any follow-up. At one year, the number needed to harm (NNH) was 42.5. We carried out a sensitivity analysis for the one-year meta-analysis by including only RCTs with overall low risk of bias. The sensitivity analysis showed a reduction in the odds ratio from 1.26 (95% CI: 0.82-1.93) (Figure 7) to 1.05 (95% CI: 0.64 – 1.75) (Figure 8); the NNH increased to 651. The certainty of effects ranged from low to moderate.
**Table 4.** Meta-analyses of repair integrity (re-tear events).

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Follow-up</th>
<th>Number of studies</th>
<th>Total sample size</th>
<th>Effect Estimate</th>
<th>P value</th>
<th>OR [95% CI]</th>
<th>GRADE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repair integrity</td>
<td>3 months</td>
<td>2</td>
<td>168</td>
<td>0.94 [0.39, 2.27]</td>
<td>0.92</td>
<td>Moderate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6 months</td>
<td>3</td>
<td>221</td>
<td>1.34 [0.59, 3.04]</td>
<td>0.48</td>
<td>Moderate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 year</td>
<td>8</td>
<td>960</td>
<td>1.26 [0.82, 1.93]</td>
<td>0.29</td>
<td>Low</td>
<td></td>
</tr>
</tbody>
</table>

OR: odds ratio, CI: confidence interval

**Figure 7.** Meta-analysis of repair integrity (re-tear events) at 1 year.

**Figure 8.** Sensitivity meta-analysis of repair integrity (re-tear events) at 1 year.

### Complications

Overall, the number of complications were low. Nine RCTs (3, 19, 20, 22, 23, 26, 27, 30, 34) reported post-operative complications. The most common complication, reported in five RCTs (19, 20, 22, 23, 34), was related to limited shoulder range of movement. Jenssen, Lundgreen (34) reported two cases of capsulitis in the early group and none in the delayed group. Sheps, Silveira (20) found one case of frozen shoulder in the early group and two cases in the delayed group. Koh, Lim (22) found no difference in the proportion of patients with stiffness (defined as having any one of the following three: forward
elevation of $<120^\circ$, internal rotation lower than L3, and/or external rotation with the arm at the side of $<20^\circ$) at three months postoperatively (early: 53% vs delayed: 50%). However, at 2 years follow-up the proportion of patients with shoulder stiffness in the delayed group (38%, n= 18/48) was greater than the early group (18%, n=7/40). Arndt, Clavert (23) and Sheps, Bouliane (19) did not report the number of patients with complications by group. Other complications reported included deep and superficial infection, loose anchors, suture pull-out, persistent shoulder pain, biceps subluxation, detached biceps tendon, deep vein thrombosis and pulmonary embolism; no differences between groups were observed for these other complications.

**Patient and public involvement**

The only RCT to describe the participation of patients in the study development was Littlewood, Bateman (3). Their patient and public involvement and engagement activities involved three meetings facilitated by the lead researcher. The group supported the co-production of patient-facing materials and development of study processes such as recruitment and informed consent. They also had a patient representative as part of their trial management group (35).

**DISCUSSION**

We aimed to summarise the effectiveness of early compared to delayed/standard rehabilitation following rotator cuff repair on clinical outcomes and repair integrity. We found no statistically significant differences for pain. For function, the only statistically significant difference was for the SANE at six
months in favour of early rehabilitation (MD:6.54, 95%CI:2.24, 10.84, p=0.003).

The mean difference found for outcome measures for function were small and
did not reach the minimal clinically important differences (MCID)(36-38). The
MCID for the SANE for patients undergoing rotator cuff repair is reported to be
16.9 (37). The meta-analyses for range of movement showed statistically
significant differences for flexion, abduction, internal and external rotation,
mainly in the short-term in favour of early rehabilitation. However, similar to
function, the mean differences were small and unlikely to be clinically
significant (39).

The majority of RCTs were rated at high risk or unclear for overall bias, this was
mainly due to issues with the randomisation process, but also partially related to
the blinding of personnel and patients, which is not often possible in RCTs
delivering exercises interventions. Risk of bias was one of the main reasons for
downgrading the certainty of the evidence.

Despite rotator cuff repair being such a common operation, we observed
substantial variations in the rehabilitation protocols in relation to when patients
were permitted to begin moving their shoulder and timelines for exercise
progression. Almost all RCTs only allowed passive shoulder exercises in the
first few weeks post-surgery for their early rehabilitation groups. Only
Littlewood, Bateman (3), Sheps, Bouliane (19) used an individualised patient-
directed approach, which facilitated a controlled and progressive introduction of
active movements according to an acceptable symptoms response (40). Sheps,
Silveira (20) also used a patient-directed approach; however only pain-free
active movements were allowed and exercise progression was still based on
the number of weeks post-surgery.
Restricting to only passive exercises in early postoperative stages may not provide sufficient load to stimulate and assist tendon healing (41) and, therefore, may not provide an optimal stimulus for tissue repair and remodelling to enable patients return to their usual activities, including leisure and work. Only four RCTs had clear distinctions between their groups regarding the sling usage; three of them prescribed the sling for comfort only and one RCT recommended patients to not use a sling at all. Although sling immobilisation has been traditionally viewed by many as important in protecting the tendon to facilitate healing, its use is open to question. Stephens, Littlewood (42) interviewed patients who were part of the RaCeR RCT (3), where patients were supported to remove their sling after rotator cuff repair surgery as soon as they felt able. They reported that patients who were in the delayed group, and had to use a sling for four weeks, found that the sling contributed to their pain instead of relieving it. The restrictions imposed by the sling also impacted patients self-efficacy and in some cases even their self-identity. In contrast, those patients in the early group felt more confident and had the perception that moving their shoulder was contributing to their recovery. Another qualitative study (43) exploring patients perceptions of rehabilitation after shoulder arthroplasty reported that the sling, especially when using an abduction component, was impairing patients sleep. Patients reported that they were unable to sleep because of the position, itching and temperature changes.

One of the main justifications for sling immobilisation and delayed rehabilitation following rotator cuff repair is the risk to the integrity of the repair, i.e. re-tear or failure to heal. However, as observed in our meta-analysis of repair integrity at one-year follow-up with RCTs at low risk of bias, the absolute risk reduction is

21
very small (0.2%) and the NNH is 651. Thus, the chances of a re-tear that is
caused by starting an early controlled and progressive mobilisation of the
shoulder is no higher than delayed/standard rehabilitation and needs to be
considered in the context of the negative effects of immobilisation. There is
debate that for small and medium tears early rehabilitation may be appropriate;
however, for large tears the risk is less acceptable. Subgroups analyses by tear
size was not possible and further recommendations for specific groups cannot
be made. Only eight RCTs included patients with large tears in their sample, but
data reported by tear size were not available. The RCT from Sheps, Silveira
(20) (n=206; low risk of bias), the largest RCT included in our systematic review,
included patients with large tears. They found that despite patients with a large
tear having a higher risk of a re-tear, this risk was not affected by the
postoperative rehabilitation protocol. In contrast, using a sling for weeks may
possibly cause problems such as stiffness and may contribute to further
deconditioning of an already weakened rotator cuff muscle. As observed by
Koh, Lim (22), delaying mobilisation and restricting movement may increase the
risk of patients having range of movement limitations in the long-term.

Strengths and limitations

We followed strict methods for this systematic review. However, the certainty of
effects were affected by the methodological quality of the body of evidence. The
majority of the RCTs were considered of high or unclear overall risk of bias, had
small sample sizes and their definition of early and delayed rehabilitation were
not consistent. Further subgroup analyses were not possible due to the lack of
data reported by tear size. Therefore, our results should be interpreted with
caution, especially for large tears where the tendon integrity may be at a higher risk of failure (44).

**Implications for clinical practice**

Current approaches to early rehabilitation following rotator cuff lead to more rapid restoration of range of movement and, importantly, based on current data, risk of re-tear does not seem to be increased.

We found substantial variation in the time that patients used a sling, how exercises were progressed, and a lack of information about exercise dosage. This limits the ability to make specific clinical recommendations in relation to an optimal rehabilitation programme.

**Implications for research**

Further large, high-quality RCTs are needed to investigate the effectiveness of early rehabilitation, particularly in relation to individual, patient-directed rehabilitation. Future RCTs must ensure that their sample size is of adequate power to allow for definitive national healthcare recommendations on whether early rehabilitation is beneficial for different tear sizes, especially for large tears. We also observed a lack of any cost-effectiveness analyses in all RCTs included. A health economics analysis should not be ignored in any future RCTs and must be included to inform decision making about healthcare provision to this common group of patients.
CONCLUSION

Although rotator cuff repair is a common surgery, postoperative rehabilitation has not evolved for over twenty years. Protocols are still variable and cautious with regards to sling use and exercise progression. Our systematic reviewed suggests that patients treated with early rehabilitation may have a faster recovery of their range of movement and are not at a higher risk of compromising the repair integrity, which has been a concern that for clinicians.

A large, definitive high-quality multicentre RCT, including all rotator cuff tear sizes and with a more progressive and controlled approach is needed to advance knowledge and for conclusive recommendations on the optimal rehabilitation programme following rotator cuff repair.
REFERENCES


7. Mazuquin BF, Wright AC, Russell S, Monga P, Selfe J, Richards J. Effectiveness of early compared with conservative rehabilitation for patients


27. Kjær BH. Physiotherapy for patients with traumatic rotator cuff tear and associations within patient characteristics.: University of Southern Denmark; 2019.


33. Oyarzún DR, Poblete LA, Nieto FF, Huerta CO, Lange CC, Cifuentes AT. Efectividad de ejercicios submaximales versus protocolo convencional en el...


39. Muir SW, Corea CL, Beaupre L. Evaluating change in clinical status: reliability and measures of agreement for the assessment of glenohumeral


Supporting information captions

Supplementary file 1. Search strategy.
Supplementary file 2. Characteristics of the included RCTs.
Supplementary file 3. Risk of bias by RCT.
Supplementary file 4. Characteristics of the rehabilitation programmes.
Supplementary file 5. Forest plots.
Supplementary file 6. GRADE summary of findings.
<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Early Mean (SD)</th>
<th>Total</th>
<th>Delayed Mean (SD)</th>
<th>Total</th>
<th>Weight IV, Random, 95% CI</th>
<th>Mean Difference IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fawzy, 2017</td>
<td>4.6 (1.4)</td>
<td>78</td>
<td>1.1 (0.9)</td>
<td>86</td>
<td>17.4%</td>
<td>3.50 [3.14, 3.86]</td>
</tr>
<tr>
<td>Mazzoca, 2017</td>
<td>2.5 (2.1)</td>
<td>35</td>
<td>3.1 (2.2)</td>
<td>33</td>
<td>16.5%</td>
<td>-0.60 [-1.62, 0.42]</td>
</tr>
<tr>
<td>Raschhofer 2017</td>
<td>2.3 (2.3)</td>
<td>14</td>
<td>3.3 (2.1)</td>
<td>16</td>
<td>15.3%</td>
<td>-1.00 [-2.58, 0.58]</td>
</tr>
<tr>
<td>Sheps, 2015</td>
<td>2.4 (2.5)</td>
<td>80</td>
<td>2.3 (2.1)</td>
<td>85</td>
<td>17.0%</td>
<td>-0.10 [-0.59, 0.79]</td>
</tr>
<tr>
<td>Sheps, 2019</td>
<td>2.2 (2.0)</td>
<td>101</td>
<td>2.1 (1.9)</td>
<td>99</td>
<td>17.2%</td>
<td>0.20 [-0.34, 0.74]</td>
</tr>
<tr>
<td>Tirefort, 2019</td>
<td>3.4 (2.2)</td>
<td>40</td>
<td>3.5 (2.2)</td>
<td>40</td>
<td>16.6%</td>
<td>-0.10 [-1.06, 0.86]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td></td>
<td>348</td>
<td></td>
<td>359</td>
<td></td>
<td>0.39 [-1.35, 2.13]</td>
</tr>
</tbody>
</table>

Heterogeneity: $\tau^2 = 4.48; \chi^2 = 187.76, df = 5 (P < 0.000001); I^2 = 97\%$

Test for overall effect: $Z = 0.44 (P = 0.66)$
<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Early Mean</th>
<th>Early SD</th>
<th>Total</th>
<th>Delayed Mean</th>
<th>Delayed SD</th>
<th>Total</th>
<th>Weight</th>
<th>IV, Random, 95% CI</th>
<th>Mean Difference IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kim, 2012</td>
<td>65.2</td>
<td>20</td>
<td>56</td>
<td>64.7</td>
<td>21.7</td>
<td>49</td>
<td>31.4%</td>
<td>0.50 [-7.52, 8.52]</td>
<td></td>
</tr>
<tr>
<td>Mazzocca, 2017</td>
<td>70</td>
<td>13</td>
<td>31</td>
<td>64</td>
<td>18</td>
<td>27</td>
<td>30.2%</td>
<td>6.00 [-2.19, 14.51]</td>
<td></td>
</tr>
<tr>
<td>Tirefort, 2019</td>
<td>72.9</td>
<td>16.7</td>
<td>40</td>
<td>69.1</td>
<td>16.4</td>
<td>40</td>
<td>38.4%</td>
<td>3.80 [-3.45, 11.15]</td>
<td></td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>127</strong></td>
<td></td>
<td></td>
<td><strong>116</strong></td>
<td></td>
<td></td>
<td><strong>100.0%</strong></td>
<td><strong>3.43 [-1.07, 7.92]</strong></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 0.90$, df = 2 ($P = 0.64$); $I^2 = 0\%$

Test for overall effect: $Z = 1.49$ ($P = 0.14$)
<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Early Mean</th>
<th>Early SD</th>
<th>Early Total</th>
<th>Delayed Mean</th>
<th>Delayed SD</th>
<th>Delayed Total</th>
<th>Weight</th>
<th>Mean Difference</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fawzy, 2017</td>
<td>90.9</td>
<td>2.8</td>
<td>78</td>
<td>91.6</td>
<td>3.5</td>
<td>86</td>
<td>84.8%</td>
<td>-0.70</td>
<td>[-1.67, 0.27]</td>
</tr>
<tr>
<td>Keener, 2014</td>
<td>88.1</td>
<td>15.8</td>
<td>61</td>
<td>89.1</td>
<td>14.1</td>
<td>53</td>
<td>10.00%</td>
<td>-1.00</td>
<td>[-6.49, 4.49]</td>
</tr>
<tr>
<td>Kim, 2012</td>
<td>73.3</td>
<td>56.2</td>
<td>56</td>
<td>82.9</td>
<td>27.5</td>
<td>49</td>
<td>1.2%</td>
<td>-9.60</td>
<td>[-26.21, 7.01]</td>
</tr>
<tr>
<td>Mazzoca, 2017</td>
<td>90</td>
<td>15</td>
<td>31</td>
<td>84</td>
<td>19</td>
<td>27</td>
<td>4.0%</td>
<td>6.00</td>
<td>[-2.90, 14.90]</td>
</tr>
<tr>
<td>Total</td>
<td>226</td>
<td></td>
<td></td>
<td>215</td>
<td></td>
<td>100.0%</td>
<td></td>
<td>-0.57</td>
<td>[-2.39, 1.25]</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau^2 = 0.77; Chi^2 = 3.28, df = 3 (P = 0.35); I^2 = 9%
Test for overall effect: Z = 0.61 (P = 0.54)
<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Events Early</th>
<th>Total Early</th>
<th>Events Delayed</th>
<th>Total Delayed</th>
<th>Weight</th>
<th>Odds Ratio (M-H, Random, 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cuff, 2012</td>
<td>5</td>
<td>33</td>
<td>3</td>
<td>35</td>
<td>7.6%</td>
<td>1.90 [0.42, 8.70]</td>
</tr>
<tr>
<td>Fawzy, 2017</td>
<td>8</td>
<td>78</td>
<td>2</td>
<td>86</td>
<td>7.0%</td>
<td>4.80 [0.99, 23.34]</td>
</tr>
<tr>
<td>Jenssen, 2018</td>
<td>6</td>
<td>56</td>
<td>6</td>
<td>57</td>
<td>11.9%</td>
<td>1.02 [0.31, 3.38]</td>
</tr>
<tr>
<td>Keener, 2014</td>
<td>6</td>
<td>63</td>
<td>3</td>
<td>53</td>
<td>8.4%</td>
<td>1.75 [0.42, 7.38]</td>
</tr>
<tr>
<td>Kim, 2012</td>
<td>7</td>
<td>56</td>
<td>9</td>
<td>49</td>
<td>14.7%</td>
<td>0.63 [0.22, 1.86]</td>
</tr>
<tr>
<td>Lee, 2012</td>
<td>7</td>
<td>30</td>
<td>3</td>
<td>34</td>
<td>8.2%</td>
<td>3.14 [0.73, 13.49]</td>
</tr>
<tr>
<td>Sheps, 2015</td>
<td>4</td>
<td>80</td>
<td>3</td>
<td>85</td>
<td>7.5%</td>
<td>1.44 [0.31, 6.64]</td>
</tr>
<tr>
<td>Sheps, 2019</td>
<td>24</td>
<td>79</td>
<td>28</td>
<td>86</td>
<td>34.6%</td>
<td>0.90 [0.47, 1.75]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>67</td>
<td>475</td>
<td>57</td>
<td>485</td>
<td>100.0%</td>
<td>1.26 [0.82, 1.93]</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.02; Chi² = 7.46, df = 7 (P = 0.38); I² = 6%

Test for overall effect: Z = 1.05 (P = 0.29)