Effectiveness of early compared with conservative rehabilitation for patients having rotator cuff repair surgery: an overview of systematic reviews

Bruno Fles Mazuquin,1 Andrew Charles Wright,2 Sarah Russell,3 Puneet Monga,3 James Selfe,4 Jim Richards1

ABSTRACT

Aim/objective The aim is to critically analyse and discuss the current literature and determine the effectiveness of rehabilitation for patients after surgical repair of rotator cuff tears for range of motion (ROM), pain, functional status and retear rates; in addition, an update of new literature is included.

Design Overview of systematic reviews.

Data sources A search was performed with no restrictions to date of publication and language in the following databases: EBSCO, AMED, CINAHL, SPORTDiscus, EMBASE, Cochrane, LILACS, MEDLINE, PEDro, Scielo, SCOPUS and Web of Knowledge. The PRISMA guideline was followed to develop this review and the R-AMSTAR tool was used for critical appraisal of included reviews.

Eligibility criteria Only systematic reviews and randomised controlled trials (RCTs) comparing the effectiveness of early with conservative rehabilitation, after surgical repair of the rotator cuff, were included. Moreover, the studies should report ROM, pain, functional status and/or retear rates before and after 3–24 months of the surgery.

Results 10 systematic reviews and 11 RCTs were included for the final analysis. Conflicting results and conclusions were presented by the systematic reviews, the use of primary studies varied; also the methodological quality of the reviews was diverse. This updated review, with new meta-analysis, showed no difference for function, pain, ROM or retear ratio between early and conservative rehabilitation.

Summary/Conclusions Early mobilisation may be beneficial, particularly for small and medium tears; however, more studies with higher quality are required, especially for patients with large tears who have been given less attention.

INTRODUCTION

Following surgical rotator cuff repair, a period of movement restriction is advised;1 however, the optimal time of immobilisation is unknown. It is common practice to ask patients to use a sling for 6 weeks and avoid activities with the affected shoulder.2 3 This period is important to protect the tendon, allow healing and to prevent retear episodes.4 However, delayed motion may increase the risk of postoperative shoulder stiffness, muscle atrophy and delay functional recovery.5 Based on the available evidence, it is difficult to make a clinical decision about the best rehabilitation regime and establish the most favourable time to start postoperative rehabilitation. One of the problems is the variation in the rehabilitation protocols and evidence provided from multiple systematic reviews. This lack of consensus may lead therapists to a variety of contradictory clinical decisions.6 Previous systematic reviews have included different primary studies and focused on different periods within which early or conservative intervention were applied, and these discrepancies may account for some of the inconsistent findings. In addition, the majority of these systematic reviews were published between 2014 and 2015, highlighting that this is currently an area of much debate.

Therefore, the aims of our systematic review of systematic reviews were to (1) comprehensively review the available evidence in the topic and (2) assess the effectiveness of early mobilisation for pain, functional status, range of motion (ROM) and retear rate for this patient population.

METHODS

Design

This study is a review of systematic reviews which followed the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) statement7 requirements. Moreover, each systematic review included was assessed and scored using the Revised Assessment of Multiple Systematic Reviews (R-AMSTAR) tool.7

Inclusion/exclusion criteria

Systematic reviews and randomised controlled trials (RCTs) that compared the effectiveness of early rehabilitation with conservative rehabilitation, after surgical repair of the rotator cuff, and were allocated to groups that had different starting times of their rehabilitation (physiotherapy and exercises). Moreover, studies should report at least one of: shoulder ROM, pain, functional scores and retear rates.8

For inclusion, studies must have:

▸ Reported at least one of: shoulder ROM, pain, functional scores and retear rates.

▸ Included patients who had a surgical repair of the rotator cuff and were allocated to groups that had different starting times of their rehabilitation (physiotherapy and exercises).

▸ Reported a clinically relevant follow-up period of between 3 and 24 months.

Studies that included patients with acute tears and studies where the aim was not to compare the impact of the rehabilitation start time application were excluded. Only chronic tears were considered, which were defined as not being caused by a traumatic event (i.e., accidents).
Search strategy

The search strategy planning was supported by a librarian and applied independently by two reviewers in the databases. The main MeSH terms and key words: rotator cuff, shoulder, shoulder joint, rehabilitation*, physiotherapy, physical therapy, immobilization, stiffness, accelerat* and sling were used in the following databases: EBSCO, AMED, CINAHL, SPORTDiscus, EMBASE, Cochrane, LILACS, MEDLINE, PEDro, Scielo, SCOPUS and Web of Knowledge. There were no restrictions on language or date of publication. Secondary searching of reference lists of key articles and grey literature was undertaken to identify any additional studies missed on the electronic database search. In order to permit the search to return other primary studies, which were not included to the published reviews, MeSH terms and key words such as review, systematic review and meta-analysis were not used in the search strategy. The electronic databases were searched in October 2015. Further information about how the searches were structured in each database is available in the online supplementary material 1. The selection process was based first on the title; further, the abstract and the full text were reviewed for inclusion.

Data extraction

The data extracted and synthesised by two independent reviewers were: author names and publication years, design of the included primary studies, inclusion criteria for primary studies, group intervention and comparison of the primary studies, tools used for outcomes assessment, the results for the variables of interest (ie, ROM, functional scores and retears rate) and references of the primary studies.

Risk of bias assessment

Although every systematic review did their own risk of bias assessment for primary studies, the inconsistency on final rates from these reviews leads to the decision of independently scoring the primary studies already scored in other reviews, in addition to new studies that were included in the update. Therefore, we used the Cochrane Handbook for Systematic Reviews of Interventions to assess the risk of bias of each primary study. The areas assessed were: method of randomisation, allocation concealment, patient blinding, care provider blinding, outcome assessor blinding, dropout rate, intention-to-treat analysis, reports on the study free of suggestion of selective outcome reporting, similarity of participants at the baseline, cointerventions avoided, compliance, timing of the outcome assessment and follow-up. Each item was scored as low, high or unclear risk of bias.

Two reviewers independently completed the R-AMSTAR and Cochrane risk of bias assessments, and any disagreements were discussed until consensus was reached. We assessed inter-reviewer agreement using the weighted κ statistic.

The Grades of Recommendation, Assessment, Development and Evaluation (GRADE) approach was used to rate the quality and strength of the evidence synthesised from the primary studies. Following the GRADE system, whether the outcome was based on a body of evidence of RCTs, the recommendation is rated as high; however, if factors affecting the quality of the study were observed (limitations in the design and implementation, indirectness of evidence, unexplained heterogeneity or inconsistency of results, imprecision of results and publication bias), the score was downgraded accordingly.

Meta-analyses

For the systematic reviews update, meta-analyses for the outcomes were performed. They were separate according to different questionnaires and tools used to score the outcomes: the American Shoulder and Elbow Surgeons (ASES), the Constant-Murley score (CM), the Simple Shoulder Test (SST), visual analogue scale (VAS) and ROM. Continuous data were expressed as mean difference and 95% CI, while for dichotomous outcomes, the OR was used with 95% CI. We assessed statistical heterogeneity using Higgins’ I². Where the studies were homogenous (p>0.10), we used a fixed effects meta-analysis model and in all other cases, we used a random effects model.

The software for the inter-reviewers’ agreement on R-AMSTAR and risk of bias was the MedCalc, V.15.4 (MedCalc Software, Ostend, Belgium) and for all meta-analyses was the RevMan V.5.3.5 (The Nordic Cochrane Centre, Copenhagen, Denmark).

RESULTS

Review of systematic reviews

Initially, 1722 records were screened regarding the inclusion criteria; from the total, 13 were selected for final decision (figure 1). Thirteen systematic reviews were analysed and three others were excluded. These were: van der Meijden et al, as the primary objective was not to compare the influence of the rehabilitation time during the recovery process; Ross et al, which used a non-systematic review method and Shen et al, which was published in Chinese. It is noteworthy that another review from Shen et al, in the same topic, was published in English in the

Meta-analyses

For the systematic reviews update, meta-analyses for the outcomes were performed. They were separate according to different questionnaires and tools used to score the outcomes: the American Shoulder and Elbow Surgeons (ASES), the Constant-Murley score (CM), the Simple Shoulder Test (SST), visual analogue scale (VAS) and ROM. Continuous data were expressed as mean difference and 95% CI, while for dichotomous outcomes, the OR was used with 95% CI. We assessed statistical heterogeneity using Higgins’ I². Where the studies were homogenous (p>0.10), we used a fixed effects meta-analysis model and in all other cases, we used a random effect model.

The software for the inter-reviewers’ agreement on R-AMSTAR and risk of bias was the MedCalc, V.15.4 (MedCalc Software, Ostend, Belgium) and for all meta-analyses was the RevMan V.5.3.5 (The Nordic Cochrane Centre, Copenhagen, Denmark).

RESULTS

Review of systematic reviews

Initially, 1722 records were screened regarding the inclusion criteria; from the total, 13 were selected for final decision (figure 1). Thirteen systematic reviews were analysed and three others were excluded. These were: van der Meijden et al, as the primary objective was not to compare the influence of the rehabilitation time during the recovery process; Ross et al, which used a non-systematic review method and Shen et al, which was published in Chinese. It is noteworthy that another review from Shen et al, in the same topic, was published in English in the

Meta-analyses

For the systematic reviews update, meta-analyses for the outcomes were performed. They were separate according to different questionnaires and tools used to score the outcomes: the American Shoulder and Elbow Surgeons (ASES), the Constant-Murley score (CM), the Simple Shoulder Test (SST), visual analogue scale (VAS) and ROM. Continuous data were expressed as mean difference and 95% CI, while for dichotomous outcomes, the OR was used with 95% CI. We assessed statistical heterogeneity using Higgins’ I². Where the studies were homogenous (p>0.10), we used a fixed effects meta-analysis model and in all other cases, we used a random effect model.

The software for the inter-reviewers’ agreement on R-AMSTAR and risk of bias was the MedCalc, V.15.4 (MedCalc Software, Ostend, Belgium) and for all meta-analyses was the RevMan V.5.3.5 (The Nordic Cochrane Centre, Copenhagen, Denmark).

RESULTS

Review of systematic reviews

Initially, 1722 records were screened regarding the inclusion criteria; from the total, 13 were selected for final decision (figure 1). Thirteen systematic reviews were analysed and three others were excluded. These were: van der Meijden et al, as the primary objective was not to compare the influence of the rehabilitation time during the recovery process; Ross et al, which used a non-systematic review method and Shen et al, which was published in Chinese. It is noteworthy that another review from Shen et al, in the same topic, was published in English in the

Figure 1 Flow diagram of selected systematic reviews.
same year, which might indicate that the same review has been published in two different journals.

Population
The number of patients in each review varied between 265 and 1776. The majority did not stipulate an age range as inclusion criteria, only two reviews\textsuperscript{16,17} included this information and both chose the age of 18 as the lower limit. Only Chang et al\textsuperscript{18} used traumatic tears as exclusion criteria; other studies did not consider different tear types. More information about the selected systematic reviews is summarised in the online supplementary material 2.

Group categorisation
Classification of participants as early or conservative/delayed group had extensive variations. Four studies did not specify how the groups were defined and the other had different thresholds. The starting time for rehabilitation also had great variance among the primary studies, from the same day postsurgery to 4 weeks in the early management, and from 4 to 8 weeks in the conservative group.

Outcomes
The most reported tool was the American Shoulder and Elbow Surgeons questionnaire. The majority of the reviews performed meta-analyses for ROM and retears ratio. Only Chan et al\textsuperscript{16} reported separated meta-analysis for clinical scores (ASES, CM, SST). One systematic review evaluated retear rates only.\textsuperscript{19}

Clinical disclosures
The conclusions were divergent about ROM, functionality and retear rate. For instance, Chan et al\textsuperscript{16} found no differences between groups for all aforementioned outcomes, which was similar to the founds from Littlewood et al.\textsuperscript{17} In contrast, the reviews from Chang et al,\textsuperscript{18} Huang et al,\textsuperscript{20} Riboh and Garrigues,\textsuperscript{21} and Shen et al\textsuperscript{15} found differences for ROM which favours the early management, especially of shoulder flexion.

Methodological appraisal
The $\kappa$ values of reviewers’ interagreement are available in the online supplementary material 3. The R-AMSTAR values ranged from 20 for Yi et al\textsuperscript{22} to 38 for Chan et al\textsuperscript{16} and Chang et al\textsuperscript{18}, of a possible total of 44; the individual scores are described in table 1. The item with lowest scores was 10, which considered publication bias and the use of statistical tests such as Egger regression to address this issue; only the studies by Chang et al\textsuperscript{18} and Chen et al\textsuperscript{23} fulfilled these criteria. The item with the highest score was 6, with all reviews apart from Shen et al\textsuperscript{15} and Chen et al\textsuperscript{23} scoring the maximum of 4; this item assess whether the characteristics of primary studies (ie, sample size, rehabilitation, outcomes) were described.

Regarding the level of evidence, the majority of the reviews contained only RCTs for qualitative and quantitative analysis. Table 2 shows the RCTs included in each systematic review. From the reviews assessed, only Cuff and Pupello\textsuperscript{26} were included in all reviews. The inclusion of other studies varies in a few systematic reviews. Other primary studies observed in the

### Table 1  R-AMSTAR score of systematic reviews

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Items</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chan et al (2014)\textsuperscript{16}</td>
<td></td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>38</td>
</tr>
<tr>
<td>Chang et al (2015)\textsuperscript{18}</td>
<td></td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>38</td>
</tr>
<tr>
<td>Chen et al (2015)\textsuperscript{23}</td>
<td></td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>Gallagher et al (2015)\textsuperscript{24}</td>
<td></td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>32</td>
</tr>
<tr>
<td>Huang et al (2013)\textsuperscript{10}</td>
<td></td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>29</td>
</tr>
<tr>
<td>Kluczynski et al (2014)\textsuperscript{19}</td>
<td></td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>21</td>
</tr>
<tr>
<td>Littlewood et al (2015)\textsuperscript{17}</td>
<td></td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>33</td>
</tr>
<tr>
<td>Riboh and Garrigues (2014)\textsuperscript{21}</td>
<td></td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>32</td>
</tr>
<tr>
<td>Shen et al (2014)\textsuperscript{15}</td>
<td></td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>29</td>
</tr>
<tr>
<td>Yi et al (2015)\textsuperscript{22}</td>
<td></td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>20</td>
</tr>
</tbody>
</table>

### Table 2  Randomised controlled trials included in the systematic reviews

<table>
<thead>
<tr>
<th>Randomised controlled trials</th>
<th>Systematic reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arndt et al\textsuperscript{25}</td>
<td>X X X X X X X X</td>
</tr>
<tr>
<td>Cuff and Pupello\textsuperscript{26}</td>
<td>X X X X X X X X</td>
</tr>
<tr>
<td>Cote and Mazzocca (unpublished abstract)</td>
<td>X</td>
</tr>
<tr>
<td>Deutsch et al\textsuperscript{27}</td>
<td>X</td>
</tr>
<tr>
<td>Duzgun et al\textsuperscript{28}</td>
<td>X X X X X X X</td>
</tr>
<tr>
<td>Keener et al\textsuperscript{29}</td>
<td>X X X X X X X X</td>
</tr>
<tr>
<td>Kim et al\textsuperscript{30}</td>
<td>X X X X X X X X</td>
</tr>
<tr>
<td>Klintberg et al\textsuperscript{31}</td>
<td>X</td>
</tr>
<tr>
<td>Koh et al\textsuperscript{32}</td>
<td>X X X X X X X X</td>
</tr>
<tr>
<td>Lee et al\textsuperscript{33}</td>
<td>X X X X X X X X</td>
</tr>
</tbody>
</table>

reviews, such as Garofalo et al.,34 Hayes et al.,35 Klintberg et al.,31 Lastayo et al.,36 Raab et al.,37 and Roddey et al.,38 were screened but further excluded, as they did not fulfil the inclusion criteria. Since the review of Kluczynski et al.19 used RCTs and studies with other levels of evidence, only the RCTs were added to Table 2; however, a detailed screening in the references demanded attention for the abstract from Deutsch et al.,27 which is indeed an RCT. For this reason, this abstract in addition to an unpublished abstract from Cote and Mazzocca were also included. To use the unpublished abstract, the permission from the authors was requested by email.

Systematic review update

We identified 1722 potentially eligible RCTs. For the final analysis, 11 full texts and 2 abstracts were assessed for inclusion. Seven studies were included in our meta-analysis. Two RCTs31 39 were excluded as they assessed patients with traumatic tears. The flow diagram (Figure 2) describes the selection process.

The agreements between the reviewers regarding the risk of bias items are available in the online supplementary material 3; the κ scores varied from 0.64 to 1. Figures 3 and 4 show the final risk of bias scores. The study with the lowest risk of bias was from Koh et al.32 and the studies with the highest risk of bias were from Cote and Mazzocca (unpublished abstract) and Deutsch et al.27

Participants

The mean age of participants varied between 55.3 and 65.1 (49.7% men). Three studies assessed only supraspinatus and the other eight did not use one of the rotator cuff muscles as inclusion criteria. The tear size varied, the majority included medium-sized tears and five studies included large tears in their groups. The surgery characteristics also varied: all used the arthroscopic technique; the footprint fixation was not homogeneous and multiple surgical techniques (single row, double row, suture bridge) were used. Additional procedures (including long head of biceps tenodesis or tenotomy, acromioplasty and capsular release) were also reported, but only Lee et al.33 excluded participants who had additional procedures in combination with the rotator cuff repair. More information about the main characteristics of the primary studies is detailed in the online supplementary material 4.

Orthoses and physiotherapy

The orthoses used to restrict shoulder movement were diverse. Four studies clearly stated the use of a sling: Cuff and Pupello26 described the use of a shoulder immobiliser; Deutsch et al.27 an Ultrasling and Kim et al.30 and De Roo et al.40 prescribed a brace. Kim et al.30 Lee et al.13 and De Roo et al.10 also made use of a pillow to maintain an abduction angle of 30° and Koh et al.32 to maintain an abduction angle of 20°. No further information about orthotic material or design were available.

A summary of rehabilitation programmes is available in the online supplementary file 5. There was variation in the initiation of early rehabilitation: four studies started passive ROM on the first postoperative day, two studies started after 2 days, four waited to complete 1-week postsurgery and Koh et al.32 had the latest start, after 5 weeks. Likewise, the conservative/delayed groups showed variations: one starting on the first postoperative day, five starting after 4 weeks, one starting after 5 weeks, three after 6 weeks and one after 9 weeks.

Despite the differences, the rationale for load increase was similar, starting with passive exercises, progressing to active ROM and then strengthening. The most common exercises in the first stage were the pendulum and active ROM for the hand, wrist and elbow. The most complete therapy description was from Duzgun et al.28 who described soft tissue mobilisation and cold packs in the first stage, and proprioceptive neuromuscular facilitation techniques in the strengthening stage.

Meta-analyses update and grading of evidence

Pain (VAS)

Meta-analysis was possible at 6 and 24 months follow-up (figures 5 and 6). Two studies were included, with 207 patients.
Figure 4  Risk of bias summary.

Figure 5  Meta-analysis of pain intensity at 6 months postoperative measured by visual analogue scale.

Figure 6  Meta-analysis of pain intensity at 24 months postoperative measured by visual analogue scale.
No statistical differences were found for 6 (p=0.26) or 24 months (p=0.49) follow-up.

Grading of evidence: there is moderate evidence that early rehabilitation does not improve pain compared with conservative rehabilitation.

American Shoulder and Elbow Surgeons questionnaire
Meta-analysis was possible at 6, 12 and 24 months follow-up (figures 7–9). For 12 and 24 months, 2 studies were included totalling 214 and 207 patients, respectively. For 6 months, 3 studies were included, with 312 patients. At all time points, there were no statistical differences in American Shoulder and Elbow Surgeons questionnaire score between the early rehabilitation and conservative rehabilitation groups (p=0.29, 0.49 and 0.15).

Constant-Murley score
Meta-analysis was possible for the Constant-Murley score at 6 and 12 months follow-up (figures 10 and 11). Three studies, with a total of 312 patients, were included for 6-month comparison and 2 studies, with a total 214 patients, for the 12-month comparison. At all time points, there were no statistical differences in the Constant-Murley score between the early rehabilitation and conservative rehabilitation groups (p=0.44 and p=0.79).

Simple Shoulder Test
Meta-analysis was possible for the SST at 6 and 12 months follow-up (figures 12 and 13) (2 studies; 214 patients). At both time points, there was no statistical difference in SST between the early rehabilitation and conservative rehabilitation groups (p=0.44 and 0.62, respectively).

Other functional scores
Meta-analysis was precluded for the Disabilities of the Arm, Shoulder and Hand (DASH), Single Assessment Numeric Evaluation (SANE), Shoulder Pain and Disability Index (SPADI), University of California at Los Angeles Shoulder Score (UCLA) and Western Ontario Rotator Cuff (WORC) scales due to the heterogeneity of measurement tools. These instruments have been reported across different studies: Cote and Mazzocca (unpublished abstract) in their abstract reported no difference for the WORC after 6 months follow-up; they did not describe any result for the SANE. Duzgun et al showed a lower DASH

**Figure 7** Meta-analysis of American Shoulder and Elbow Surgery questionnaire at 6 months.

**Figure 8** Meta-analysis of American Shoulder and Elbow Surgery questionnaire at 12 months.

**Figure 9** Meta-analysis of American Shoulder and Elbow Surgery questionnaire at 24 months.

**Figure 10** Meta-analysis of the Constant-Murley score at 6 months.
score for the early rehabilitation group in the 6 months follow-up, although this difference was not statistically significant. Lee et al\textsuperscript{33} who used the UCLA described that both groups improved their scores, but no statistical differences were found for 6 or 12 months follow-up. De Roo et al\textsuperscript{40} did not find any differences for the SPADI score for 4 months follow-up.

Grading of evidence: there is moderate evidence that early rehabilitation does not improve function status compared with conservative rehabilitation.

Range of motion
The meta-analyses were separated according to movements: flexion and external rotation, which were measured with a goniometer and expressed in degrees. Shoulder internal rotation was not evaluated using meta-analysis because most of the measurements were related to the hand positioning of the patient to their own back and not described as a glenohumeral joint angle. Only De Roo et al\textsuperscript{40} assessed the joint angle for internal rotation, but did not find significant differences at 6 weeks or 4 months postsurgery. Abduction was found only for Lee et al\textsuperscript{33} which showed no statistical difference for 6 or 12 months, and for De Roo et al\textsuperscript{40} which showed no statistical difference for 6 weeks and 4 months. Owing to the different follow-up periods, new meta-analyses for abduction were not performed.

Flexion
Meta-analysis was possible for 6 (5 studies; 468 patients) and 24 months follow-up (2 studies; 207 patients) (figures 14 and 15). There was no statistical difference in glenohumeral flexion ROM between the early rehabilitation and conservative rehabilitation groups for 6 (p=0.09) or 24 months (p=0.61).

External rotation
Meta-analysis was possible for 6 and 24 months (figures 16 and 17). Five studies with a total of 468 patients were included for the 6 months comparison, 2 studies with 207 participants were included for the 24 months comparison. There was no statistical difference in glenohumeral external rotation ROM between the

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Early Mean</th>
<th>SD</th>
<th>Total</th>
<th>Conservative Mean</th>
<th>SD</th>
<th>Total</th>
<th>Mean Difference IV, Fixed, 95% CI</th>
<th>Mean Difference IV, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keener et al, 2014</td>
<td>78.1</td>
<td>10</td>
<td>59</td>
<td>79.8</td>
<td>12.3</td>
<td>50</td>
<td>49.2% -0.00 [-9.08, 3.48]</td>
<td>115</td>
</tr>
<tr>
<td>Kim et al, 2012</td>
<td>68.91</td>
<td>7.48</td>
<td>56</td>
<td>69.63</td>
<td>13.43</td>
<td>48</td>
<td>50.2% -0.02 [-4.28, 4.22]</td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>99</td>
<td>100.0% -0.41 [-3.41, 2.60]</td>
<td>115</td>
<td>99</td>
<td>100.0% -0.00 [0.08, 0.00]</td>
<td>0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: Ch^2 = 0.00, df = 1 (P = 0.80), I^2 = 0%</td>
<td>Test for overall effect: Z = 0.27 (P = 0.79)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 11** Meta-analysis of the Constant-Murley score at 12 months.

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Early Mean</th>
<th>SD</th>
<th>Total</th>
<th>Conservative Mean</th>
<th>SD</th>
<th>Total</th>
<th>Mean Difference IV, Fixed, 95% CI</th>
<th>Mean Difference IV, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keener et al, 2014</td>
<td>9.1</td>
<td>2.7</td>
<td>59</td>
<td>9.2</td>
<td>2.9</td>
<td>50</td>
<td>59.9% -0.20 [-1.26, 0.88]</td>
<td>115</td>
</tr>
<tr>
<td>Kim et al, 2012</td>
<td>7.81</td>
<td>3.17</td>
<td>56</td>
<td>6.7</td>
<td>3.55</td>
<td>49</td>
<td>40.1% 1.11 [-0.19, 2.41]</td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>99</td>
<td>100.0% 0.32 [-0.49, 1.14]</td>
<td>115</td>
<td>99</td>
<td>100.0% -0.00 [0.00, 0.00]</td>
<td>0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: Ch^2 = 2.38, df = 1 (P = 0.12), I^2 = 58%</td>
<td>Test for overall effect: Z = 0.78 (P = 0.44)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 12** Meta-analysis of Simple Shoulder Test at 6 months.

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Early Mean</th>
<th>SD</th>
<th>Total</th>
<th>Conservative Mean</th>
<th>SD</th>
<th>Total</th>
<th>Mean Difference IV, Fixed, 95% CI</th>
<th>Mean Difference IV, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arndt et al, 2012</td>
<td>158.4</td>
<td>22.9</td>
<td>49</td>
<td>146.4</td>
<td>30</td>
<td>43</td>
<td>12.2% 12.00 [0.98, 23.02]</td>
<td>115</td>
</tr>
<tr>
<td>Keener et al, 2014</td>
<td>155</td>
<td>18.1</td>
<td>59</td>
<td>154</td>
<td>17.8</td>
<td>50</td>
<td>32.5% 1.00 [5.76, 7.76]</td>
<td></td>
</tr>
<tr>
<td>Kim et al, 2012</td>
<td>150.57</td>
<td>33.2</td>
<td>56</td>
<td>147.14</td>
<td>21.2</td>
<td>49</td>
<td>13.4% 3.43 [7.10, 13.98]</td>
<td></td>
</tr>
<tr>
<td>Koh et al, 2014</td>
<td>126.9</td>
<td>26.3</td>
<td>46</td>
<td>129.5</td>
<td>22.1</td>
<td>52</td>
<td>14.4% -2.60 [-12.76, 7.55]</td>
<td></td>
</tr>
<tr>
<td>Lee et al, 2012</td>
<td>157.3</td>
<td>11.4</td>
<td>30</td>
<td>151.9</td>
<td>18.2</td>
<td>34</td>
<td>27.5% 5.40 [-1.95, 12.75]</td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>240</td>
<td>228</td>
<td>100.0% 3.36 [0.49, 7.21]</td>
<td>240</td>
<td>228</td>
<td>100.0% -0.50 [0.00, 0.00]</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: Ch^2 = 4.45, df = 4 (P = 0.35), I^2 = 10%</td>
<td>Test for overall effect: Z = 1.71 (P = 0.09)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 13** Meta-analysis of Simple Shoulder Test at 12 months.

**Figure 14** Meta-analysis of range of motion for shoulder flexion at 6 months.
early and conservative rehabilitation groups for 6 (p=0.13) or 24 months (p=0.52).

Grading of evidence: there is weak evidence that early rehabilitation improves ROM compared with conservative rehabilitation.

Retear rate
Meta-analysis was possible for 12 months follow-up (figure 18) (5 studies, 410 participants). There was no statistical difference (p=0.31) in retear rate between the early and conservative rehabilitation groups.

Grading of evidence: there is moderate evidence that early rehabilitation does not cause higher retear rates.

Additional details regarding the grading of evidence can be found in online supplementary file 6.

DISCUSSION
We aimed to systematically analyse and determine the effectiveness of rehabilitation for patients who had a surgical repair of the rotator cuff, and we found no differences for ROM, functional status and retear rate between early and conservative rehabilitation. The rationale for choosing to perform an overview of systematic reviews is in supporting a faster and more reliable decision-making for the clinician, particularly with the large increase in published material in this field over the last 2 years.

Systematic reviews features
The agreement scores between reviewers about the systematic reviews quality were high, indicating good to excellent classification. While comparing the results from the 10 selected reviews, it is important to highlight some differences that may influence on their result and conclusion made. The first is related to methodological quality: The studies from Chang et al18 and Chan et al16 had the highest scores, and are considered to have the most reliable method. Between them, Chang et al18 was the only study to assess publication bias using specific statistical tests: for example, Egger test or funnel plot. This test shows

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Early</th>
<th>Conservative</th>
<th>Mean difference IV, Fixed, 95% CI</th>
<th>Mean difference IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keener et al, 2014</td>
<td>164 13.4</td>
<td>59 163 15.9</td>
<td>0.01 [4.56, 5.68]</td>
<td>1.00 [7.65, 11.65]</td>
</tr>
<tr>
<td>Koh et al, 2014</td>
<td>143.7 23.5</td>
<td>46 141.7 25.2</td>
<td>0.24 [7.65, 11.65]</td>
<td>-2.00 [7.65, 11.65]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td></td>
<td></td>
<td>1.00 [4.56, 5.68]</td>
<td>1.00 [7.65, 11.65]</td>
</tr>
<tr>
<td>Heterogeneity: CHI² = 0.03, df = 1 (P = 0.86); P = 0%</td>
<td>Test for overall effect Z = 0.51 (P = 0.61)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 15 Meta-analysis of range of motion for shoulder flexion a 24 months.

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Early</th>
<th>Conservative</th>
<th>Mean difference IV, Fixed, 95% CI</th>
<th>Mean difference IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arndt et al, 2012</td>
<td>54.3 12.5</td>
<td>49 44.3 19.4</td>
<td>10.00 [3.23, 16.77]</td>
<td>-2.30 [-8.48, 3.88]</td>
</tr>
<tr>
<td>Keener et al, 2014</td>
<td>61.6 17.8</td>
<td>59 63.9 15.1</td>
<td>21.98 [0.15, 42.81]</td>
<td>-2.30 [-8.48, 3.88]</td>
</tr>
<tr>
<td>Kim et al, 2012</td>
<td>77.2 20.1</td>
<td>56 73.8 29.7</td>
<td>49.14 [3.45, 95.84]</td>
<td>-2.30 [-8.48, 3.88]</td>
</tr>
<tr>
<td>Koh et al, 2014</td>
<td>18.9 16.5</td>
<td>46 19.9 15.4</td>
<td>21.98 [0.15, 42.81]</td>
<td>-2.30 [-8.48, 3.88]</td>
</tr>
<tr>
<td>Lee et al, 2012</td>
<td>50.3 11.2</td>
<td>30 41.6 14.9</td>
<td>21.98 [0.15, 42.81]</td>
<td>-2.30 [-8.48, 3.88]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td></td>
<td></td>
<td>21.98 [0.15, 42.81]</td>
<td>-2.30 [-8.48, 3.88]</td>
</tr>
<tr>
<td>Heterogeneity: Tau² = 21.04; CHI² = 10.02, df = 4 (P = 0.03); P = 63%</td>
<td>Test for overall effect Z = 1.53 (P = 0.13)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 16 Meta-analysis of range of motion for shoulder external rotation at 6 months.

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Early</th>
<th>Conservative</th>
<th>Mean difference IV, Fixed, 95% CI</th>
<th>Mean difference IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Koh et al, 2014</td>
<td>33.1 18.8</td>
<td>46 29.5 21.2</td>
<td>34.22 [4.32, 11.52]</td>
<td>3.60 [4.32, 11.52]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td></td>
<td></td>
<td>-4.20 [-9.91, 1.51]</td>
<td>3.60 [4.32, 11.52]</td>
</tr>
<tr>
<td>Heterogeneity: CHI² = 2.45, df = 1 (P = 0.12); P = 59%</td>
<td>Test for overall effect Z = 0.65 (P = 0.52)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 17 Meta-analysis of range of motion for shoulder external rotation at 24 months.

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Early Events</th>
<th>Total</th>
<th>Conservative Events</th>
<th>Total</th>
<th>Odds Ratio M-H, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cuff &amp; Pupello, 2012</td>
<td>5</td>
<td>33</td>
<td>3</td>
<td>35</td>
<td>1.40 [0.42, 4.87]</td>
</tr>
<tr>
<td>Keener et al, 2014</td>
<td>8</td>
<td>83</td>
<td>3</td>
<td>53</td>
<td>18.8 [4.2, 7.38]</td>
</tr>
<tr>
<td>Kim et al, 2012</td>
<td>7</td>
<td>56</td>
<td>9</td>
<td>49</td>
<td>47.8 [2.2, 1.88]</td>
</tr>
<tr>
<td>Koh et al, 2014</td>
<td>2</td>
<td>25</td>
<td>2</td>
<td>32</td>
<td>9.2 [0.17, 9.97]</td>
</tr>
<tr>
<td>Lee et al, 2012</td>
<td>7</td>
<td>30</td>
<td>3</td>
<td>34</td>
<td>12.3 [0.73, 13.49]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td></td>
<td></td>
<td>207</td>
<td></td>
<td>1.37 [0.74, 2.53]</td>
</tr>
<tr>
<td>Heterogeneity: CHI² = 3.52, df = 4 (P = 0.47); P = 0%</td>
<td>Test for overall effect Z = 1.01 (P = 0.31)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 18 Meta-analysis of OR for retears at 12 months.
whether small studies with unfavourable results due to no significant differences may impact on the final result of a meta-analysis when multiple and more powerful studies are compared.\textsuperscript{42} Although assessing publication bias would seem an important factor for a review, its use is not recommended with continuous data when the number of studies is fewer than 10; in this case, the regression test do not have enough power to show funnel asymmetry.\textsuperscript{41,42} While none of the selected systematic reviews used more than 10 primary studies for meta-analysis, the publication bias is not a major concern for the purposes of the review’s topic. Therefore, even if item 11 of the R-AMSTAR was discounted, the study of Chan \textit{et al}\textsuperscript{16} would stay with the best score.

Another disparity is related to primary studies and their respective levels of evidence. Although the publication dates among the multiple reviews were not >2 years, the variation of studies included was diverse. Only the RCTs from Cuff and Pupello\textsuperscript{26} were cited across all publications. However, one explanation for the difference could be the use of different search strategies and databases. For instance, Kluczynski \textit{et al}\textsuperscript{19} and Huang \textit{et al}\textsuperscript{26} used search terms and key words, but did not have a structured search strategy. In contrast, Chan \textit{et al}\textsuperscript{16} Chang \textit{et al}\textsuperscript{18} and Shen \textit{et al}\textsuperscript{15} organised different strategies for each database.

Systematic review update

Most studies were at high risk of bias and did not fulfil essential components such as a proper method of randomisation and allocation concealment. Furthermore, other items such as \textit{coinervention avoided} and \textit{compliance} have been neglected. It is well established that the first two aforementioned are crucial to ensuring that the results of the studies are reliable and valid.\textsuperscript{3} The component \textit{compliance} should be reported for rehabilitation trials testing different protocols as it contributes to making clear how many sessions each participant of each group attended, thus it is possible to know if the groups are truly comparable.\textsuperscript{9}

Meta-analyses update

Based on the information of the multiple studies, it was possible to separate the meta-analyses in relation to each functional questionnaire, pain, ROM of different movements and retears ratio. The results of new meta-analyses presented were similar to those from Chan \textit{et al}\textsuperscript{16} however, we presented new analyses for 6, 12 and 24 months, not just for the final follow-up. Moreover, the review from Chan \textit{et al}\textsuperscript{16} has some flaws; they used the data of the follow-up of 24 months from Keener \textit{et al}\textsuperscript{22} to compare with the 12 months follow-up from Kim \textit{et al}\textsuperscript{20}. They also combined results from Cuff and Pupello\textsuperscript{26} in the same analysis; however, the original article does not contain information of the SD, which makes the analyses challenging. Their further efforts were to input the p value, but it did not show any alterations to their results. Chang \textit{et al}\textsuperscript{18} also did meta-analyses for functional scores; however, they used the standardised mean difference using multiple questionnaires in the same analysis.

The screening of individual RCTs revealed the possibility of separated analyses for each questionnaire. Although the use of the standardised mean difference is not incorrect, it will not inform how much improvement for every questionnaire is necessary, as the standardised mean difference will report results as a general unit rather than specific. The separated report, using the mean difference, is more advantageous as it provides the therapist with a choice of which instrument they would like to use.

Assessment of movement

For ROM, the meta-analyses were separated for different movements. The results were similar to the review of Riboh and Garrigues.\textsuperscript{21} They presented meta-analyses for flexion and external rotation for 3, 6 and 12 months. They showed statistically significant differences in favour of early management for flexion at 3, 6 and 12 months, and external rotation only at 3 months. In the new meta-analysis for 6 months, the inclusion of the data from Koh \textit{et al}\textsuperscript{12} changed the previous result now to not statistically significant. As no other data were added to 3 and 12 months for flexion and external rotation, only new meta-analyses were performed for 6 and 24 months. It is important to highlight that the difference between early and conservative rehabilitation for 3 months from Riboh and Garrigues\textsuperscript{21} was 14.7, which is above the MCID of 14. However, this difference was not consistent for the other follow-ups or movements.

Rehabilitation aspects

Immobilisation

The type of orthoses varied among studies and there was no consensus regarding whether the shoulder should be angled in abduction or maintained besides the thorax. The most common reported method was the sling alone, but four studies\textsuperscript{30,32,33,40} described the use of an abduction pillow, with different angles.

The prescription of immobilisation posture should consider the characteristics of the repair. Mechanical stress in the surgical site must be avoided as much as possible, to facilitate safe healing. According to a recent survey with physiotherapist and surgeons in the UK about the current practice on rotator cuff rehabilitation, 86% indicate that their patients use a sling, only 18% use a brace and 2% stated other forms of immobilisation.\textsuperscript{43} Depending on factors such as muscle involved, tear size and surgical method applied, the positioning should be different, higher abduction angles for more severe stages.\textsuperscript{44}
Rehabilitation protocol from Lee et al. is highly aggressive. In the very first day postoperative, passive ROM of shoulder flexion and external rotation were already implemented, in addition to stretching of shoulder muscles. The frequency was also high, being performed twice a day and self-passive ROM up to three times per day, already in the first week. In comparison to the meta-analysis of Huang et al., two new available studies were published and one excluded in the updated retears meta-analyses. The reason for the exclusion of the Arndt et al. is due to lack of clear information about the absolute numbers of individuals who had retears in each group, as well as the absolute total number of patients in each group that were assessed for retears. Thus, the inputted data from Huang et al. of 41 patients in each arm are incorrect and are a conjecture. Other reviews from Chen et al. and Littlewood et al. that performed meta-analysis for retears also erroneously included the study from Arndt et al.

The agreement with the conclusion of Thomson et al. for more mild cases is ratified based on two parameters, that are calculated based on the results of the new meta-analysis: the absolute risk increase (ARI) and the number needed to treat or in the specific case of retears ratio, that is an unfavourable outcome, this is referred to as the number needed to harm (NNH).

The ARI indicates that the early group had 3.2% more retears cases. The NNH indicates that 32 patients treated with early rehabilitation are needed for one to have a retear. However, as stated previously, the study from Lee et al. had patients with larger tears and a very aggressive protocol. Hence, if the ARI and NNH are calculated without the inclusion of Lee et al., the results are: ARI=1.29% and NNH=77.5. This result reveals that early mobilisation for patients with smaller tears have just 1.29% more retears episodes, compared with those who had a more conservative approach. Furthermore, 78 patients needed to harm show that the chances of having a recurrence because of more permissive mobilisation are very low, as the retears may be caused by other factors rather than the mobilisation itself.

In contrast, for more severe stages, the recommendation from Thomson et al. must be considered carefully. Conclusions based only on the result from Lee et al. where values of ARI and NNH are 14% and 7.14, respectively, should not be taken further due to the presence of bias issues, in addition to the concerns related to their protocol previously described. Considering the other studies included in the updated review, it is still not possible to drive to definitive conclusions, as all studies that included patients with large tears also failed to fulfill fundamental methodological aspects, which indicates important risk of bias.

Limitations
Although strict methods were used for this systematic review, it presents some limitations. It was not possible to perform meta-analysis only with high-quality RCTs; however, the objective was to review and critically analyse the available evidence. It was not possible to stipulate what is the ideal frequency and intensity of the treatment for any stage of rotator cuff tears. The overlap of the first mobilisation/rehabilitation session for both groups is considerable, which may influence possible differences in outcomes between groups. Another relevant limitation is the lack of information about the tear characteristics in many studies.

The results from the meta-analysis of pain must be carefully interpreted. It is not clear whether other treatments for pain management (e.g., NSAID injection) were used in any of the trials as this information was not reported.

Based on the studies analysed, it is clear that mild stages permit an early approach to recover ROM, but it was not possible to formulate recommendations of when mobilisation should start for patients with more severe stages, because of the lack of studies focusing on this subgroup.

Clinical message
The optimal time to initiate more substantial shoulder mobilisation appears to be around 3–4 weeks postsurgery. If the mechanical stimuli are applied when the collagen fibres begin to develop their structural arrangement, the tendons may have their viscoelastic properties enhanced. This may help to avoid subsequent complications related to tissue adhesion and stiffness. Owing to limited high level evidence for patients with large tears, the reasoning to underpin an appropriate protocol should follow the mechanobiology of tendon healing and metabolic characteristics. Factors that can compromise patients’ progression must be considered and mutual consensus between therapist and surgeon is appropriate.

Implications for research
Further, high-quality RCTs are necessary to safely recommend the optimal to start the rehabilitation, especially for more severe subgroups, that may present higher risk of having complications regarding their ROM and function. Important factors that may affect the risk of bias for rehabilitation trials such as compliance and co-intervention avoidance must be considered. Future studies should also ensure adequate random sequence generation, allocation concealment and blinding of assessors. The risk for bias was the main reason for downgrading the GRADE outcomes classification in our review.

New studies should use methods that are able to provide a better description of movement control. It is worth exploring how muscle behaviour develops from presurgery through to the follow-up period and how physiotherapy influences muscular recovery.

What are the findings?
- There is little evidence to confirm whether or not early mobilisation improves outcomes.
- The results for pain must be interpreted carefully as information about additional therapies for pain management is lacking.
- Early mobilisation is not the direct cause of higher retear rates. Patients factors that compromise recovery must be considered.

How might it impact on clinical practice in the future?
- Rehabilitation plays an important role after rotator cuff repairs.
- Existing literature has shown a variety of outcomes from early rehabilitation.
- It is unclear if early rehabilitation benefits patients after surgery.
CONCLUSION
Early mobilisation does not improve functional outcomes, pain or ROM when compared with a conservative rehabilitation and does not cause a higher retear rate. However, there is still no consensus on which is the best approach due to the heterogeneity of protocols and methodological quality of reviews and primary studies.

Acknowledgements
Bruno Fles Mazuquin has a PhD scholarship sponsored by the CAPES foundation, through the Science without Borders program from the Brazilian government.

Competing interests
None declared.

Provenance and peer review
Not commissioned; externally peer reviewed.

Open Access
This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

REFERENCES

47 Sedgwick P. What is number needed to harm (NNH)? BMJ 2013;347:f4869.