Clinical effectiveness of tenotomy versus tenodesis for long head of biceps pathology: a systematic review and meta-analysis

Alexander William Hartland,1 Raisa Islam,1 Kar Hao Teoh,1 Mustafa Saad Rashid.2,3

ABSTRACT

Objectives The comparative clinical effectiveness of common surgical techniques to address long head of biceps (LHB) pathology is unclear. We synthesised the evidence to compare the clinical effectiveness of tenotomy versus tenodesis.

Design A systematic review and meta-analysis using the Grading of Recommendations Assessment, Development and Evaluation approach.

Data sources EMBASE, Medline, PsycINFO and the Cochrane Library of randomised controlled trials were searched through 31 October 2021.

Eligibility criteria We included randomised controlled trials, reporting patient reported outcome measures, comparing LHB tenotomy with tenodesis for LHB pathology, with or without concomitant rotator cuff pathology. Studies including patients treated for superior labral anterior–posterior tears were excluded. No language limits were employed. All publications from database inception to 31 October 2021 were included.

Data extraction and synthesis Screening was performed by two authors independently. A third author reviewed the article, where consensus for inclusion was required. Data were extracted by two authors. Data were synthesised using RevMan. Inverse variance statistics and a random effects model were used.

Results 860 patients from 11 RCTs (426 tenotomy vs 434 tenodesis) were included. Pooled analysis of patient-reported functional outcome measures data demonstrated comparable outcomes (n=10 studies; 403 tenotomy vs 416 tenodesis; standardised mean difference (SMD): 0.14, 95% CI −0.04 to 0.32, p=0.13). There was no significant difference for pain (Visual Analogue Scale) (n=8 studies; 345 tenotomy vs 350 tenodesis; MD: −0.11, 95% CI −0.28 to 0.06, p=0.21). Tenodesis resulted in a lower rate of Popeye deformity (n=10 studies; 401 tenotomy vs 410 tenodesis; OR: 0.29, 95% CI 0.19 to 0.45, p<0.00001). Tenotomy demonstrated shorter operative time (n=4 studies; 204 tenotomy vs 201 tenodesis; MD 15.2, 95% CI 1.06 to 29.36, p<0.00001).

Conclusions Aside from a lower rate of cosmetic deformity, tenodesis yielded no significant clinical benefit to tenotomy for addressing LHB pathology.

PROSPERO registration number CRD42020198658.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ The present systematic review and meta-analysis was undertaken to include only randomised controlled trials in an attempt to reduce risk of bias within studies.
⇒ We employed the Grading of Recommendations Assessment, Development and Evaluation approach to evaluate the certainty of our findings in a systematic manner.
⇒ A sensitivity analysis accounting for the influence of concomitant rotator cuff tear was undertaken.
⇒ Pooling of various patient reported outcome measures was used and reported using standardised mean difference, which increases power but may minimise small but significant effects between different outcome measures.

INTRODUCTION

Pathology in the long head of biceps was described as early as 1835.1 Initial descriptions relate to dislocation from the bicipital groove and traumatic rupture.2–12 Surgeons became more specialised, and diagnostic techniques improved, associations with rotator cuff tears were described.3–14 The long head of biceps was also noted to be an isolated source of pain in the anterior shoulder in some patients.5–6 In most cases, the long head of biceps tendon is thought to be an associated pathology. The range of pathology ranges from fraying and inflammation to partial tearing and rupture. Medial subluxation is commonly seen with upper border of subscapularis tendon tears.7 Lateral subluxation can be associated with leading edge of supraspinatus tears, which may involve the rotator cable. Occasionally, the long head of biceps may be the principal source of anterior shoulder pain. In younger patients, it can be associated with superior labral injuries.8–10 Long head of biceps tenotomy, often performed arthroscopically, involves
detaching the tendon from the superior labral anchor, typically with a radiofrequency ablation device. Long head of biceps tenodesis involves reattaching the detached long head of biceps, using a button device or suture anchor typically, to the proximal humerus. This can be either intra-articularly, within the bicipital groove, or in the subpectoral region of the proximal humerus. With significant advances in arthroscopic instrumentation and implants, a large increase in variations of both tenotomy and tenodesis have been reported. Some authors have demonstrated tenotomy at the biceps anchor can auto-tenodese in the intra-articular opening of the groove, as the stump base retracts. This implant-free technique may have cost and time-saving benefits; however, a larger proportion of techniques have reported utilisation of implants to secure the tendon. These include suture anchors, tenodesis screws and suspensory buttons. Having chosen to tenotomise or tenodese, and after the fixation type is chosen, the surgeon must then determine the location of the tenodesis.

There remains significant debate regarding optimal location of tenodesis to address long head of biceps pathology. Some authors have suggested the bicipital groove and transverse humeral ligament as causative factors for pain and advocated for subpectoral tenodesis location. Others have demonstrated intra-articular tenodesis at the proximal portion of the bicipital groove can yield reliable improvement in symptoms. Cramping pain in the muscle belly has been reported following tenodesis. It is thought to be related to mismatch of tension between the muscle bellies. O’Brien popularised the arthroscopic long head of biceps tendon transfer (to the short head) procedure as means of addressing differential traction between the muscle bellies.

The increasing number of variations of surgical techniques addressing long head of biceps tendon pathology in the last few decades has been accompanied by an increase in published studies reporting clinical outcomes. Most techniques have been reported with clinical outcomes from case series or cohort studies. Few studies have evaluated differences between surgical techniques using a randomised controlled trial (RCT) design. Most RCTs have enrolled less than 50 patients total. Studies of this magnitude are unlikely to find difference in clinical outcomes unless a large treatment effect exists between techniques. A systematic review and meta-analysis of multiple RCTs may provide sufficient statistical power to detect differences between various surgical techniques for managing long head of biceps tendon pathology.

The aim of this study was to systematically review the literature and quantitively synthesise data relating to surgical treatment of pathology in the long head of biceps tendon. The primary aim was to determine whether a difference in clinical effectiveness exists between tenotomy and tenodesis using data from RCTs.

METHODS
This study is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement. A protocol for the study was registered with International Prospective Register for Systematic Reviews (https://www.crd.york.ac.uk/prospero/), and published prior to study commencement.

Eligibility criteria
Only RCTs were included. All other trial designs were excluded. Studies with human patients of any age undergoing any type of surgery to the long head of biceps were included. This included both arthroscopic and open techniques. The intervention of interest was the surgical procedure of biceps tenodesis. The comparators included any alternative surgical procedures for biceps pathology, including biceps tenotomy. Studies relating to superior labral anterior–posterior (SLAP) tears predominantly were excluded. The primary outcome included patient reported functional outcome measures related to the shoulder. Secondary outcome measures included pain Visual Analogue Scale (VAS), rate of cosmetic deformity (‘Popeye’ sign) and operative time.

Information sources and search strategy
The following bibliographic databases were searched: MEDLINE, EMBASE, PsycINFO and the Cochrane Library of RCTs. No restrictions were placed on language. Eligible studies published from database inception up to 31 October 2021 were included. To increase sensitivity and heighten precision, the RCT filters developed by the Cochrane collaborative were used for each database in the search strategy. The used search terms are included in the online supplemental appendix. Search terms were generated using a population, intervention, comparator and outcome (PICO) approach, and combined with the Cochrane RCTs filters for each database searched. Most search terms were limited to abstract and title fields. References from published systematic reviews investigating the same or similar topic were manually searched for relevant included studies. The database search was conducted between 1 October 2021 and 31 October 2021. Two coauthors conducted the electronic search of databases (AWH and MSR).

Study selection and data extraction
All literature search results were combined in Endnote V.X9 (Clarivate Analytics). Duplicate articles were removed. Two independent reviewers (AWH and RI) screened the titles and abstracts, with consensus sought prior to full text review. Subsequent full text review of articles meeting all eligibility criteria determined the final inclusion.

Data extraction involved two independent reviewers (AWH and KT). A standardised proforma was used by one reviewer (AWH) to extract the required data. A second reviewer (KT) then checked the extracted data for any inaccuracies. Any differences found during the
The random effects approach for meta-analysis incorporates an assumption that different studies are estimating different, yet related, intervention effects. It uses the mean difference (SMD) to pool results as most of the patient-reported outcome measures (PROMs) pooled using standardised mean difference (SMD) and a random effects model. The SMD is a summary statistic in meta-analysis used when studies measure the same outcome (e.g., pain in function after surgery to the long head of biceps) but measures it in different ways (e.g., different PROMs). It allows the results of the studies to be presented on a uniform scale.\(^2\)

The random effects approach for meta-analysis incorporates an assumption that different studies are estimating different, yet related, intervention effects. It uses the inverse-variance statistical method.\(^2\) It allows for a degree of heterogeneity between studies that is not readily explained by other factors; in the context of surgical trials, the treatment effect may be influenced by subtle and difficult to detect variations in patient selection, surgical technique and other factors.\(^2\) This included the Constant-Murley Score (CMS), the American Shoulder and Elbow Surgeons Shoulder Score (ASES) and University of California at Los Angeles (UCLA) shoulder score. Pooling of PROMs allowed for increased statistical power and permitted a larger number of studies to be included in each meta-analysis. The decision was made to pool results as most of the patient-reported outcome measures are thought to measure similar aspects of the clinical outcome. That is, most outcome scores included employ questions to quantify degree of pain, and level functional limitation. The CMS also includes a physician-administered objective component using a spring balance and a goniometer to assess strength and range of motion, respectively. Studies have reported good correlation between the CMS, ASES and Oxford Shoulder Score (OSS).\(^2,20\) The limitation to pooling PROMs in this way is the possibility of minimising small but significant effects potentially seen between different outcome measures. Secondary outcomes included pain VAS, rate of cosmetic deformity (Popeye sign) and operative time.

Assessment of risk of bias

Two reviewers assessed risk of bias independently (AWH and KT). To assess for potential bias of individual studies, the Cochrane collaboration Risk of Bias tool V.2.0 was used.\(^2\) Within this tool, there are five domains of bias, with each domain being assigned a level of risk of bias (high risk, low risk or some concerns). Interpretation of the risk of bias for each domain is guided by preset signalling questions. The tool subsequently generates an overall risk of bias for each study.

To assess for risk of bias across studies, publication bias was assessed using a funnel plot of included studies investigating the primary outcome. We reviewed available trial protocols and registrations to compare predefined outcomes and statistical analysis plans with those reported in the published manuscripts, assessing for selective reporting within studies.

Data synthesis and statistical analysis

Data were synthesised and analysed as primary and secondary outcomes. The primary outcome included PROMs, pooled using an SMD scale. A random effects model was used for analysis due to expected heterogeneity across the studies. Heterogeneity was quantified using the \(I^2\) value and the \(\chi^2\) test for heterogeneity. Statistical heterogeneity is a consequence of methodological and clinical diversity, occurring when different studies are brought together in a meta-analysis. The \(I^2\) statistic describes the percentage of the variability in treatment effect estimates due to heterogeneity rather than chance.\(^2\) \(I^2\) values were interpreted as described in the Cochrane Handbook: 0%–40% might not be important, 30%–60% may represent moderate heterogeneity, 50%–90% may represent substantial heterogeneity and 75%–100% represents considerable heterogeneity. Data were only synthesised if the method in which the outcome was recorded was comparable. Outcomes with continuous variables, such as operative time, were summarised using mean differences and inverse variance statistical analysis. Outcomes with dichotomous data, such as rate of cosmetic deformity (Popeye sign), were summarised using ORs with 95% CIs. A sensitivity analysis was undertaken to compare studies involving patients with and without an intact rotator cuff.

Determining certainty of findings (using GRADE approach)

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework was used to describe the strength of the body of evidence provided and confidence in the findings.\(^2\) Each outcome was determined to be of very low, low, moderate or high certainty. Certainty can be adjusted due to risk of bias, inconsistency, indirectness, imprecision and publication bias. The level of certainty can be increased by a large magnitude of effect, a dose–response gradient or the effect of plausible residual confounding.

Patient and public involvement

No patients were involved.

RESULTS

Study selection

The literature search of EMBASE, MEDLINE, PsycINFO and Cochrane library databases identified 1034 records.
After removal of duplicates, 623 unique manuscripts were screened. Five hundred and ninety eight records were excluded after title and abstract screening; 25 records were selected for full text review. Of these 25 records, 15 were eligible for inclusion in the systematic review. Eleven were eligible for inclusion in the meta-analysis. The study selection process, with reasons for exclusion, is presented in the PRISMA flowchart (figure 1).

**Study characteristics**

Included trials were published between 2006 and 2021. Individual characteristics are reported in table 1. Only RCTs were included. Of these, three were conducted in Italy, three in the Republic of Korea, two in China, two in the USA, two in Iran, one in Canada, one in Spain and one in the Netherlands. Eleven studies compared biceps tenodesis with biceps tenotomy with sample sizes ranging from 34 to 151 and a total of 860 patients. These were included in the meta-analysis. One study compared tenodesis in the groove and tenotomy of the long head of biceps attachment on the supraglenoid tubercle with tenodesis only. One study compared interference screw tenodesis with suture anchor tenodesis, one compared

---

**Figure 1** PRISMA flow diagram for included studies. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.
Table 1  List of included studies and study characteristics

<table>
<thead>
<tr>
<th>First author+year</th>
<th>Journal</th>
<th>Study design</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Indication</th>
<th>Age (years) (range or SD)</th>
<th>Tenodesis location</th>
<th>Tenodesis method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Franceschi 2006</td>
<td>International orthopaedics</td>
<td>RCT</td>
<td>Biceps tenodesis without tenotomy</td>
<td>Biceps tenodesis with tenotomy</td>
<td>Full thickness rotator cuff lesion and retraction and biceps pathology</td>
<td>60.3 (41–79)</td>
<td>58.1 (40–81)</td>
<td>Just proximal to bicipital groove</td>
</tr>
<tr>
<td>De Carli 2012</td>
<td>KSSTA</td>
<td>RCT</td>
<td>Biceps tenotomy and tenodesis</td>
<td>Biceps tenotomy</td>
<td>Small–large rotator cuff tear and associated degenerative lesion of LHB</td>
<td>56.3 (3.9)</td>
<td>59.6 (8.7)</td>
<td>Not reported</td>
</tr>
<tr>
<td>Zhang 2013</td>
<td>KSSTA</td>
<td>RCT</td>
<td>Biceps tenodesis</td>
<td>Biceps tenotomy</td>
<td>Rotator cuff tears and LHB pathology, &gt;55 years</td>
<td>61 (55–71)</td>
<td>61 (55–67)</td>
<td>Intertubercular groove</td>
</tr>
<tr>
<td>Lee 2016</td>
<td>JSES</td>
<td>RCT</td>
<td>Biceps tenodesis</td>
<td>Biceps tenotomy</td>
<td>Symptomatic LHB partial tear and small–medium rotator cuff tear</td>
<td>62.9 (50–75)</td>
<td>62.8 (55–77)</td>
<td>Intertubercular groove, just lateral to insertion of subscapularis tendon</td>
</tr>
<tr>
<td>Lee 2016</td>
<td>Arthroscopy</td>
<td>RCT</td>
<td>Biceps tenodesis</td>
<td>Biceps tenotomy/ debridement</td>
<td>Rotator cuff tear and SLBC lesion</td>
<td>56.6 (42–76)</td>
<td>61 (53–69)</td>
<td>Approx. 3.5 cm distal to top of bicipital groove</td>
</tr>
<tr>
<td>Park 2016</td>
<td>AJSM</td>
<td>RCT</td>
<td>Interference screw biceps tenodesis</td>
<td>Suture anchor biceps tenodesis</td>
<td>Partial or full thickness rotator cuff tear and biceps lesion</td>
<td>61.2 (7.2)</td>
<td>62.4 (8.2)</td>
<td>Approx. 3 cm distal to humeral articular margin</td>
</tr>
<tr>
<td>Castricini 2017</td>
<td>KSSTA</td>
<td>RCT</td>
<td>Biceps tenodesis</td>
<td>Biceps tenotomy</td>
<td>Grade I or II full thickness supraspinatus tendon tear with LHB lesion, &gt;40 years</td>
<td>57.1 (40–70)</td>
<td>59.9 (40–71)</td>
<td>Bicipital groove</td>
</tr>
<tr>
<td>Mardani-Kivi 2018</td>
<td>Techniques in Shoulder &amp; Elbow Surgery</td>
<td>RCT</td>
<td>Biceps tenodesis</td>
<td>Biceps tenotomy</td>
<td>Rotator cuff tear with superior labral lesion and biceps pathology, 45–60 years</td>
<td>55.5 (5.2)</td>
<td>54.5 (5.3)</td>
<td>Intra-articular</td>
</tr>
<tr>
<td>Belay 2019</td>
<td>KSSTA</td>
<td>RCT</td>
<td>Biceps tenodesis</td>
<td>Biceps tenotomy</td>
<td>Partial tear or subluxation of LHB and repairable rotator cuff</td>
<td>52.9 (10.8)</td>
<td>57.7 (8.7)</td>
<td>Humerus</td>
</tr>
<tr>
<td>First author+year</td>
<td>Journal</td>
<td>Study design</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Indication</td>
<td>Age (years) (range or SD)</td>
<td>Tenodesis location</td>
<td>Tenodesis method</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------</td>
<td>--------------</td>
<td>--------------</td>
<td>------------</td>
<td>------------</td>
<td>--------------------------</td>
<td>-------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Mardani-Kivi 2019</td>
<td>Journal of Orthopaedics and Traumatology</td>
<td>RCT</td>
<td>Arthroscopic intra-articular tenodesis</td>
<td>Open subpectoral tenodesis</td>
<td>Rotator cuff tear and biceps subluxation/dislocation/partial tear, or SLAP lesion, 18–65 years</td>
<td>56.1 (6.2) 55.2 (7.7)</td>
<td>Greater tuberosity/bicipital groove</td>
<td>Interference screw</td>
</tr>
<tr>
<td>Zhang 2019</td>
<td>China Journal of Orthopaedics and Traumatology</td>
<td>RCT</td>
<td>Arthroscopic intra-articular tenodesis</td>
<td>Arthroscopic Biceps tenotomy</td>
<td>LHB tendinitis and rotator cuff disease/tear, 50–80 years</td>
<td>60.5 (6.3) 62.2 (6.1)</td>
<td>Bicipital groove midpoint</td>
<td>Suture anchor</td>
</tr>
<tr>
<td>Forsythe 2020</td>
<td>Arthroscopy</td>
<td>RCT</td>
<td>Arthroscopic suprapectoral biceps tenodesis</td>
<td>Open subpectoral biceps tenodesis</td>
<td>Biceps tendinopathy—anterior shoulder pain, bicipital groove tenderness, positive provocative manoeuvre</td>
<td>50.2 (10.5) 50.3 (10.4)</td>
<td>Zone 3 of bicipital tunnel (subpectoral)</td>
<td>Interference screw</td>
</tr>
<tr>
<td>MacDonald 2020</td>
<td>AJSM</td>
<td>RCT</td>
<td>Biceps tenodesis</td>
<td>Biceps tenotomy</td>
<td>LHB lesion±rotator cuff tear, &gt;18 years</td>
<td>58.7 (10.9) 56.3 (8.1)</td>
<td>Arthroscopic suprapectoral/Open subpectoral</td>
<td>Screw/button</td>
</tr>
<tr>
<td>Garcia 2020</td>
<td>Muscles, Ligaments and Tendons Journal</td>
<td>RCT</td>
<td>Biceps tenodesis</td>
<td>Biceps tenotomy</td>
<td>LHB pathology, men 40–65 years</td>
<td>50.7 (6.3) 54.7 (5.8)</td>
<td>2 cm inferior to the upper vertex of the greater tuberosity</td>
<td>Interference screw</td>
</tr>
<tr>
<td>van Deurzen 2021</td>
<td>Arthroscopy</td>
<td>RCT</td>
<td>Biceps tenodesis</td>
<td>Biceps tenotomy</td>
<td>Nontraumatic small-medium sized supraspinatus and/or infraspinatus lesion and inflamed/unstable LHB tendon or LHB tear &gt;30%, &gt;50 years</td>
<td>61 (51–76) 61 (51–79)</td>
<td>Proximal bicipital groove</td>
<td>Suture anchor</td>
</tr>
</tbody>
</table>

arthroscopic intra-articular tenodesis with open subpectoral tenodesis and one compared arthroscopic supraperiosteal tenodesis with open subpectoral tenodesis. In total, the 15 studies included 1084 patients, which were included in the systematic review.

**Risk of bias within studies**

Overall, 3 studies were assessed as low risk of bias; 10 studies had ‘some concerns’ regarding bias; 2 studies were assessed as high risk of bias (table 2). In the two studies assessed as high risk of bias, no final follow-up time was stated, or there was a wide range in the time of assessment. One of these studies was at risk of ‘selection of the reported result’ due to the absence of information on final follow-up. Where studies were assessed as ‘some concerns’ of bias, missing outcome data and possible selection of the reported result, was frequently found. Not

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Randomisation process</th>
<th>Deviations from the intended interventions</th>
<th>Missing outcome data</th>
<th>Measurement of the outcome</th>
<th>Selection of the reported result</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Franceschi et al 2006</td>
<td>!</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>!</td>
<td>Low risk</td>
</tr>
<tr>
<td>Sanders et al 2012</td>
<td>!</td>
<td>+</td>
<td>+</td>
<td>!</td>
<td>!</td>
<td>Some concerns</td>
</tr>
<tr>
<td>Zhang et al 2013</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>!</td>
<td>!</td>
<td>High risk</td>
</tr>
<tr>
<td>Lee et al 2016</td>
<td>+</td>
<td>+</td>
<td>!</td>
<td>+</td>
<td>!</td>
<td>!</td>
</tr>
<tr>
<td>Lee et al 2016</td>
<td>+</td>
<td>+</td>
<td>!</td>
<td>+</td>
<td>!</td>
<td>!</td>
</tr>
<tr>
<td>Park et al 2016</td>
<td>+</td>
<td>+</td>
<td>!</td>
<td>+</td>
<td>!</td>
<td>!</td>
</tr>
<tr>
<td>Castricini et al 2017</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>!</td>
</tr>
<tr>
<td>Mardani-Kivi et al 2018</td>
<td>+</td>
<td>+</td>
<td>!</td>
<td>!</td>
<td>!</td>
<td>!</td>
</tr>
<tr>
<td>Belay et al 2019</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>!</td>
</tr>
<tr>
<td>Mardani-Kivi et al 2019</td>
<td>+</td>
<td>+</td>
<td>!</td>
<td>!</td>
<td>!</td>
<td>!</td>
</tr>
<tr>
<td>Zhang et al 2019</td>
<td>!</td>
<td>+</td>
<td>+</td>
<td>!</td>
<td>+</td>
<td>!</td>
</tr>
<tr>
<td>Forsythe et al 2020</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>!</td>
</tr>
<tr>
<td>MacDonald et al 2020</td>
<td>+</td>
<td>+</td>
<td>!</td>
<td>+</td>
<td>+</td>
<td>!</td>
</tr>
<tr>
<td>Garcia et al 2020</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>!</td>
<td>!</td>
<td>!</td>
</tr>
<tr>
<td>van Deurzen et al 2021</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>!</td>
<td>!</td>
<td>!</td>
</tr>
</tbody>
</table>
all participants randomised had available data and there was often a lack of a clear, predefined, statistical analysis plan, finalised prior to commencement of the trial.

**Risk of bias across studies**

Low–moderate heterogeneity was found between studies when investigating pooled PROMs as the primary outcome ($I^2$: 37%). This was similar when excluding studies without a concurrent rotator cuff tear ($I^2$: 40%). Some heterogeneity was expected due to the variety of PROMs used. Heterogeneity between studies when looking at secondary outcomes was low: pain VAS ($I^2$: 0%), and ‘Popeye’ deformity ($I^2$: 5%). However, considerable heterogeneity was found for operative time ($I^2$: 96%). No clear evidence of publication bias was found (figure 2).

**Results of individual studies**

Of the 15 included studies in this systematic review, 10 were parallel-group RCTs comparing biceps tenodesis with biceps tenotomy.33–42 One three-arm study compared biceps tenodesis, tenotomy and debridement.43 The tenodesis attachment site and fixation method varied. Three studies compared differences in tenodesis technique.30–32 Of these, two RCTs compared tenodesis location (arthroscopic suprapectoral vs open subpectoral).31 32 The other RCT compared interference screw with suture anchor fixation for biceps tenodesis.30 One RCT evaluated the addition of tenotomy of the long head of biceps attachment on the supraglenoid tubercle after tenodesis within the proximal aspect of the groove.29

**Differences in function (data from patient-reported outcome measures)**

Most studies found no difference in function as reported by a variety of patient-reported outcome measures, including Constant, American Shoulder and Elbow Society (ASES) shoulder score, University of California Los Angeles (UCLA) score, Western Ontario Rotator Cuff (WORC) score, Single Assessment Numerical Evaluation (SANE) score, Simple Shoulder Test (SST), EuroQol 5 domain (EQ5D) score, Dutch Oxford Shoulder Score (OSS), Shoulder Strength Index (SSI), Korean Shoulder Score (KSS), 36-item Short Form (SF36), Visual Analogue Scale (VAS), Numerical Rating Scale (NRS), and Disabilities of the Arm, Shoulder and Hand (DASH). Mardani-Kivi et al (2018) demonstrated statistically significant difference in patient satisfaction as measured by VAS in favour of tenodesis (9.53 vs 9.07, $p=0.0001$); however, the difference reported is small and likely not clinically significant.37 Zhang et al34 found a difference in pain as measured by VAS at 2 weeks postoperatively (3.1 tenotomy vs 4.8 tenodesis, $p<0.001$), that disappeared by the 4-week postoperative timepoint (2.0 tenotomy vs 2.1 tenodesis, $p<0.001$).34

**Rate of ‘Popeye deformity’ and cramping pain**

Of 14 studies that reported rate of Popeye deformity, 8 demonstrated a higher frequency in the tenotomy group.33 35–41 Park et al demonstrated anatomic failure using ultrasonography or MRI in 7/33 in those receiving tenodesis with an interference screw versus 2/34 receiving tenodesis with a suture anchor.30 However, there was no

---

**Figure 2** Funnel plot of included studies reporting the primary outcome investigating publication bias. SMD, standardised mean difference.
significant difference in rate of reported Popeye deformity or biceps cramping between the two groups.\textsuperscript{30} Van Deurzen et al found high rates of Popeye deformity in both tenodesis (16/48) and tenotomy (24/52) groups, resulting in lack of difference statistically.\textsuperscript{31} Their technique for tenodesis included incorporating the biceps with the anterior suture anchor used in the concomitant rotator cuff repair. They also found no difference in rate of biceps cramping between the two groups.\textsuperscript{42} Castriconi et al found higher rates of Popeye deformity in the tenotomy group (5/24 vs 18/31, p=0.006) at 24 months postoperatively.\textsuperscript{36} They also reported higher rates of biceps cramping at 6 months in the tenodesis group (3/24 tenodesis vs 0/31 tenotomy). All 3 patients that reported cramping resolved by 24 months postoperatively.\textsuperscript{36} Ten studies in total reported biceps cramping as an outcome measure.\textsuperscript{30,32,34,36,37,40–43} Most demonstrated no difference between the two groups they evaluated with respect to cramping.\textsuperscript{30,32,34,40–43} Mardani-Kivi et al\textsuperscript{37} reported 9/29 in the tenotomy group and 0/33 in the tenodesis group experienced biceps cramping (p=0.0001).\textsuperscript{35} Garcia et al enrolled 70 male heavy manual workers into an RCT comparing tenotomy versus tenodesis. They demonstrated higher rates of Popeye deformity in the tenotomy group (13/25 vs 2/18) at 12 months (p=0.01); however, no difference in the rate of reported biceps cramping.\textsuperscript{41}

Range of motion and strength differences
Some studies measured shoulder and/or elbow range of motion using goniometers. Others included dynamometer testing using a variety of testing protocols to record strength measurements at the shoulder, elbow, and forearm. Of the 15 included studies, 11 reported these variables as outcome measures.\textsuperscript{29,30,32–36,38,40–43} All studies report no difference in shoulder range of motion between the two groups evaluated. Nine studies evaluated elbow flexion strength and reported no differences. De Carli et al also found no differences in shoulder flexion strength, shoulder extension strength and shoulder abduction strength (tested at 45° abduction) between tenotomy and tenodesis.\textsuperscript{38} Lee et al reported no difference in elbow flexion strength and forearm pronation strength with tenotomy compared with tenodesis. They did demonstrate greater supination strength in the tenodesis compared with the tenotomy group (99.8% compared with contralateral side vs 81.8% compared with the contralateral side).\textsuperscript{39} Oh et al conducted a three-arm trial comparing debridement, tenotomy and tenodesis. In this study, they found lower supination strength in the tenotomy group (88% compared with the contralateral side) compared with the tenodesis (112%) and debridement (129%) groups (p=0.039).\textsuperscript{43} Garcia et al reported increased fatigue time in the tenodesis group compared with tenotomy (122 s vs 95 s, p=0.03).\textsuperscript{44} Van Deurzen et al also reported lower rates of reported upper arm fatigue in the tenodesis group compared with tenotomy (31% vs 50%, p=0.05); however, they did not formally test fatigability.\textsuperscript{42}

One study compared biceps tenodesis with an interference screw to a suture anchor biceps tenodesis.\textsuperscript{30} This study demonstrated no significant difference in functional scores (ASES, Constant, SST and KSS), pain (VAS), range of motion (ROM) or ‘Popeye’ deformity between these groups. A greater number of failures were observed in the interference screw group (n=7) compared with the suture anchor group (n=2).

Regarding tenodesis location, one study compared arthroscopic intra-articular biceps tenodesis with open subpectoral biceps tenodesis.\textsuperscript{31} This study found no difference in functional scores (Constant and SST), pain (VAS), and patient satisfaction between these groups. Forsythe et al conducted an RCT comparing arthroscopic suprapectoral biceps tenodesis with open subpectoral biceps tenodesis.\textsuperscript{32} This study found no significant difference in functional scores (ASES, Constant and SANE), ROM, pain or complication rates between these groups. They reported a significantly greater surgical time for arthroscopic suprapectoral biceps tenodesis group (16.9±8.4 min vs 9.8±3.1 min, p<0.001).

Synthesis of results
A meta-analysis of 10 included studies was performed for the primary outcome, investigating differences in patient-reported outcome measures comparing tenodesis with tenotomy. Various PROMs measured by these individual studies (Constant, ASES and UCLA) were pooled and an SMD scale was used.

Primary outcome
Pooled PROMs
Pooled analysis of different PROMs (Constant, ASES and UCLA) measured by 10 individual studies (416 patients in tenodesis group and 403 patients in tenotomy group) demonstrated comparable outcomes between tenotomy and tenodesis (SMD: 0.14, 95% CI −0.04 to 0.32; p=0.13; I²: 37%) (figure 3).

Pooled PROMs (with rotator cuff tear)
A sensitivity analysis comparing tenodesis with tenotomy, excluding the study involving patients without a rotator cuff tear, demonstrated comparable outcomes between tenotomy and tenodesis (n=9 studies; 385 tenotomy vs 394 tenodesis; SMD: 0.12, 95% CI −0.07 to 0.31; p=0.22; I²: 40%).

Secondary outcomes
Pain (VAS)
Postoperative pain was measured using the VAS in 8 studies (350 patients in tenodesis group and 345 patients in tenotomy group). The pooled mean difference was −0.11 (n=8 studies; 345 tenotomy vs 350 tenodesis; MD: −0.11, 95% CI −0.28 to 0.06, p=0.21; I²: 0%), revealing no significant difference between groups (figure 4A).

Rate of ‘Popeye’ deformity
Ten studies reported the incidence of ‘Popeye’ deformity (including 410 patients in tenodesis group, and 401...
patients in tenotomy group). Long head of biceps tenodesis significantly reduced the rate of ‘Popeye’ deformity (n=10 studies; 401 tenotomy vs 410 tenodesis; OR: 0.29, 95% CI 0.19 to 0.45, p<0.00001; I²: 5%) (figure 4B).

### Operative time

Four studies measured operative time in minutes in 405 patients (201 patients in tenodesis group and 204 patients in tenotomy group). Long head of biceps tenodesis significantly reduced the rate of ‘Popeye’ deformity (n=10 studies; 401 tenotomy vs 410 tenodesis; OR: 0.29, 95% CI 0.19 to 0.45, p<0.00001; I²: 5%) (figure 4B).

### Figure 3

Forest plot investigating differences between tenotomy and tenodesis for primary outcome (function as assessed by a patient-reported outcome measure), pooled using SMD. SMD, standardised mean difference.

### Figure 4

Forest plots investigating differences between tenotomy and tenodesis for secondary outcomes: (A) pain as measured by VAS, (B) rate of Popeye deformity and (C) operative time. VAS, Visual Analogue Scale.
in tenotomy group). Long head of biceps tenotomy demonstrated a significantly shorter operative time (n=4 studies; 204 tenotomy vs 201 tenodesis; MD: 15.21, 95% CI 1.06 to 29.36, p<0.00001; I²: 96%) (figure 4C).

**Certainty of meta-analysis findings using GRADE framework**

The quality of evidence was rated for all four outcomes using the GRADE framework (table 3). The finding of no significant difference between tenotomy and tenodesis for the primary outcome (PROMs) was determined as ‘very low’ certainty. The findings relating to secondary outcomes were assessed as ‘moderate’ certainty (no difference in pain as measured by VAS and lower rate of ‘Popeye’ deformity favouring tenodesis) and ‘low’ certainty (shorter operative time favouring tenotomy). The GRADE evaluation for rate of ‘Popeye’ deformity and operative time were upgraded due to a large effect size. This was indicated by a test for overall effect Z-score of >2. A Z-score of 1 is equivalent to 1 SD; a Z-score of >2 was chosen as 2 SDs away from the mean is widely accepted as significant. The primary outcome was assessed as higher importance than the secondary outcomes. Detailed explanations for adjustments to certainty within the GRADE framework as outlined in the footnote of table 3.

**DISCUSSION**

This systematic review identified 15 RCTs comparing various techniques for addressing long head of biceps pathology. Of these, 11 performed a comparison of tenotomy versus tenodesis. These studies were included in the meta-analysis. Pooled results from patient reported outcome measures from 10 of these studies (403 patients received tenotomy compared with 416 who received tenodesis) demonstrated no clinical difference between tenotomy and tenodesis. Nine out of 10 studies included patients with concomitant rotator cuff tear. Sensitivity analysis, including only studies with concomitant rotator cuff tear, did not affect this finding. No difference was observed regarding pain as measured by VAS. Rate of ‘Popeye’ deformity was significantly lower in the tenodesis group compared with the tenotomy group (OR: 0.29, 95% CI 0.19 to 0.45). Tenotomy demonstrated a shorter operative compared with tenodesis (MD: 15.21, 95% CI 1.06 to 29.36). Using the GRADE framework, there was a very low certainty of finding relating to function (PROMs), low certainty for operative time, moderate certainty for pain (VAS) and Popeye deformity.

The strengths of this study are that a broad and systematic search was performed, not limited by language. All included studies were assessed for risk of bias using version 2.0 of the Cochrane risk of bias tool. Risk of bias across studies was also assessed using a funnel plot and demonstrated no evidence for publication bias. A sensitivity analysis was employed to determine change to the primary outcome when uncoupling associated pathology. Certainty of findings were reported in accordance with the GRADE framework, allowing for multiple variables.

There are limitations to this study. The systematic review and meta-analysis is restricted by the quality and size of the included RCTs. Twelve studies were assessed to have either ‘some concerns’ or ‘high risk’ of bias. Six of 10 studies used to compare functional PROMs enrolled 40 or fewer patients in each group. Despite included data from over 800 patients, β error may exist for detecting a clinical difference between tenotomy and tenodesis. Larger low risk of bias studies are needed to determine whether modest clinical differences exist between these two techniques. A large number of variations on the surgical techniques were used across the included studies. While this provides some external validity, subtle clinical differences may exist between techniques employed for tenotomy or tenodesis. There was some heterogeneity detected for the primary outcome measure (function as assessed by PROMs, I²: 37%). Due to the variety of functional outcome measures used across studies, we opted to pool measures using SMD. This analysis may potentially dilute subtle clinical difference detected by some measures but not others.

Recently, several meta-analyses have been published on the topic of surgical treatments for the long head of biceps.34-50 These studies are all limited to some extent, with some disagreement in their findings. Dekker et al reviewed cadaveric studies and demonstrated no difference in ultimate load to failure in various fixation types including screws and anchors.51 They also demonstrated no difference in load to failure when suprapectoral and subpectoral tenodesis locations were compared. Another meta-analysis of similar cadaveric studies demonstrated superior ultimate load of failure with screws compared with suture anchors (86N greater load 95% CI 34 to 138, p=0.002).52 They did find similar results to Dekker et al when comparing tenodesis location.

Meta-analyses investigating clinical differences between tenodesis and tenotomy of the long head of biceps have included different studies despite similar objectives. Belk et al included five RCTs yielded from a limited search strategy.47 They found no differences in patient-reported outcome measures. Tenodesis was found to have superior forearm supination strength. The limitations of this study were the limited search strategy, and inclusion of both long head of biceps tendon and superior labral pathology. Zhu et al performed a similar limited search strategy to Belk et al and yielded similar results.48 Both studies limited the language to English, and used V.1.0 of the Cochrane risk of bias tool to assess bias within studies. Na et al performed a meta-analysis comparing tenotomy to tenodesis to treat long head of biceps pathology in the context of repairable rotator cuff tears.44 They included two RCTs and five cohort studies in their data synthesis. Tenodesis demonstrated lower rates of ‘Popeye’ deformity, and lower Constant score; however, this was less than the minimum clinically important difference.44 This study was again limited by a limited search strategy, and the inclusion of lower quality study designs, with greater risk of bias within studies. The most robust meta-analysis on this topic published to date was by Zhou et al. They performed a more systematic search of multiple electronic databases and included nine RCTs in their meta-analysis.50 They reported...
<table>
<thead>
<tr>
<th>Certainty assessment</th>
<th>No. of studies</th>
<th>Study design</th>
<th>Risk of bias</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
<th>Biceps tenodesis</th>
<th>Biceps tenotomy</th>
<th>Effect</th>
<th>Certainty</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Functional PROMs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Randomised</td>
<td>Serious</td>
<td>Serious†</td>
<td>Not serious</td>
<td>Serious‡</td>
<td>none</td>
<td>416</td>
<td>403</td>
<td>–</td>
<td>SMD 0.14</td>
<td>SD higher</td>
</tr>
<tr>
<td></td>
<td></td>
<td>trials</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.04 lower to 0.32 higher)</td>
<td>Very low</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Randomised</td>
<td>Serious§</td>
<td>Not serious</td>
<td>Not serious</td>
<td>Not serious</td>
<td>None</td>
<td>350</td>
<td>345</td>
<td>–</td>
<td>MD 0.11</td>
<td>lower</td>
</tr>
<tr>
<td></td>
<td></td>
<td>trials</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.28 lower to 0.06 higher)</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Popeye deformity</strong></td>
<td></td>
<td></td>
<td>Serious¶</td>
<td>Not serious</td>
<td>Not serious</td>
<td>Serious**</td>
<td>Strong association††</td>
<td>41/410 (10.0%)</td>
<td>109/401 (27.2%)</td>
<td>OR 0.29</td>
<td>(0.19 to 0.45)</td>
<td>174 fewer per 1000 (from 206 fewer to 128 fewer)</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Randomised</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*One out of 10 studies included were at high risk of bias using the Cochrane ROB V.2.0 tool. Only 2 out of 10 studies included were at low risk of bias using the Cochrane ROB V.2.0 tool.
†Moderate heterogeneity (I²=37%) with variable point estimates and CIs.
‡Small sample sizes with no power calculation and an underpowered sample size in 4 out of 10 studies included.
§One out of 8 studies included were at high risk of bias using the Cochrane ROB V2.0 tool. Only 2 out of 8 studies included were at low risk of bias using the Cochrane ROB V2.0 tool.
¶One out of 10 studies included were at high risk of bias using the Cochrane ROB V2.0 tool. Only 2 out of 10 studies included were at low risk of bias using the Cochrane ROB V2.0 tool.
††Effect size Z=5.04 (p<0.00001). A Z-score of >2 used as indication of large effect size.
GRADE, Grading of Recommendations Assessment, Development and Evaluation; MD, mean difference; PROMs, patient-reported functional outcome measures; ROB, Risk of Bias; SMD, standardised mean difference; VAS, Visual Analogue Scale.
that Constant scores and rate of ‘Popeye’ deformity favoured tenodesis over tenotomy. They also used the GRADE framework and a trial sequential analysis approach to determine certainty of findings. The main limitation of this study was the use of a fixed effects model, which does not account for significant heterogeneity between studies.

Despite the several recent meta-analyses comparing tenotomy and tenodesis, the current study allows for more precise findings. The systematic search using broad search terms and a PICO approach permitted a more complete inclusion of all relevant studies. Studies that only included patients treated with long head of biceps tendon pathology, either with or without rotator cuff tear were included. Another advantage in the current study is sensitivity analysis accounting for the presence of concomitant rotator cuff tear. SLAP tears were excluded as it likely represent a significantly different pathology. Included studies were limited to level I evidence and assessed for risk of bias within studies using the latest version of the Cochrane risk of bias tool (Version 2.0). Bias between studies was assessed using a funnel plot and demonstrated no publication bias. Given heterogeneity among studies, pooled PROMs were favoured and a random effects model was chosen. Certainty of findings were reported using the GRADE framework.

The clinical implications of this study are to report the most up to date and robust meta-analysis of the highest quality evidence on tenotomy versus tenodesis. There is reasonable certainty that tenotomy reduced the risk of postoperative Popeye deformity. Whether this confers a clinical advantage remains unclear. There were no significant clinical differences between tenotomy and tenodesis; however, there may be a beta error present and most studies are underpowered to detect a modest difference. Meta-analysis of published studies does not demonstrate a superior clinical improvement with tenodesis over tenotomy; however, a large well-designed RCT is required to investigate whether clinical differences exist between these techniques. In current practice, orthopaedic surgeons often favour tenodesis to address long head of biceps pathology; however, this study demonstrates that this approach unlikely yields superior clinical benefit over tenotomy.

CONCLUSION
Tenodesis produced lower rates of ‘Popeye’ deformity while tenotomy required shorter operative time. No functional differences were detected using a variety of patient-reported outcome measures between these techniques. A large well-designed RCT is needed to investigate differences in clinical effectiveness between tenotomy and tenodesis.

Contributors AWH was involved in performing the systematic electronic literature search, screening articles, extracting data, assessing risk of bias, conducting the statistical analysis, preparing the manuscript and proof reading. KHT was involved in performing the systematic electronic literature search, screening articles, extracting data, assessing risk of bias, preparing the manuscript and proof reading. MSR was involved in analyzing the data, preparing the manuscript and proof reading. MSR acted as the guarantor for this study.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available. Not applicable.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

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, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD Mustafa Saad Rashid http://orcid.org/0000-0002-6236-4120

REFERENCES
2 Soden J. Two cases of dislocation of the tendon of the long head of the biceps humeri from its groove. Med Chir Trans 1841;24:212–20.


29 Fricke K, Largo UG, Ruzini L, et al. To detach the long head of the biceps tendon after tenodesis or not: outcome analysis at the 4-year follow-up of two different techniques. *Int Orthop* 2007;31:337–45.


