Diagnosis, prevention and treatment of common shoulder injuries in sport: grading the evidence – a statement paper commissioned by the Danish Society of Sports Physical Therapy (DSSF)

Behnam Liaghat 1,2, Julie Rønne Pedersen 1, Rasmus Skov Husted 3,4,5,6,7, Lisbeth Lund Pedersen 1,8, Kristian Thorborg 1,9, Carsten B Juhl 1,10

ABSTRACT
This statement paper sumarisres and appraises the evidence on diagnosis, prevention, and treatment of common shoulder injuries in sports. We systematically searched Medline and Embase. The Grading of Recommendations Assessment, Development and Evaluation tool was applied to evaluate the overall quality of evidence. For diagnosis, we included 19 clinical tests from mixed populations. Tests for anterior instability, biceps-labrum complex injuries and full subscapularis rupture had high diagnostic accuracy (low to moderate quality of evidence). For prevention, we included 19 clinical tests from mixed populations. Tests for anterior instability, biceps-labrum complex injuries and full subscapularis rupture had high diagnostic accuracy (low to moderate quality of evidence).

WHAT IS ALREADY KNOWN
⇒ Shoulder injuries are very common in overhead athletes.
⇒ The quality of evidence related to diagnosis, prevention and treatment of the most common shoulder injuries has not been investigated.

WHAT ARE THE NEW FINDINGS
⇒ High diagnostic accuracy was observed for 10 tests covering anterior instability (apprehension test (relocation)), SLAP injuries (Biceps Load II), biceps-complex injuries (three-pack and Yergason), and rupture of the subscapularis (internal rotation lag).
⇒ Shoulder injury prevention programmes showed a moderate to large effect size in reducing the risk of shoulder injuries compared with no intervention.
⇒ Choosing an active rehabilitation programme seems to be beneficial compared with passive modalities in reducing pain and disability.
⇒ Most outcomes were graded as very low to moderate quality of evidence, indicating that future high-quality research may alter our findings.

INTRODUCTION
Shoulder injuries occur frequently in sports such as swimming, volleyball and handball, where movements are performed repeatedly over the head at high speed or in extreme positions.1-4 Athletes from these sports are often referred to as ‘overhead athletes’. It is widely accepted that these athletes are at significant risk of sustaining shoulder injuries, with several attempts made to identify risk factors and optimal management.5-8 The risk of shoulder injuries varies across sports and depends on the definition (time lost in sports, medical attention, severity etc.). In swimming, 23%–38% of athletes sustain shoulder injuries within 1 year,5-6 and 23% of volleyball players experience dominant shoulder pain during the ongoing season.7 In a large study of elite handball, 44%–75% have previously had shoulder pain, 20%–52% report current shoulder pain and the prevalence of weekly shoulder pain and substantial shoulder injuries is 28% and 12%, respectively.8 These high rates of shoulder injuries are also seen in other overhead sports such as baseball9 and waterpolo10 and may vary depending on age, sex and level of competition.

Shoulder injuries may result from an instantaneous transfer of large quantities of kinetic energy (ie, acute with sudden onset), from the gradual accumulation of low-energy transfer over time (ie, repetitive with gradual onset) or a combination of both mechanisms (ie, repetitive with sudden
onset).12 Although the International Olympic Committee recently provided consensus-based guidelines for recording and reporting epidemiological data on injuries in sport,12 defining and classifying mechanisms and mode of onset of shoulder injuries is challenged by the lack of consistent terminology in the existing literature.

The clinical management of shoulder injuries is of interest to a broad audience within the sports medicine community, including physiotherapists, sports physicians and orthopaedic surgeons with interest in the diagnosis, prevention and treatment of such injuries.3 Therefore, to improve the management of shoulder injuries, this statement paper, commissioned by the Danish Society of Sports Physical Therapy, aimed at providing an overview of the present literature. We graded the quality of evidence concerning the diagnostic accuracy of commonly used clinical tests and the effect of preventive and treatment strategies for the most common shoulder injuries.

METHODS

This statement paper is divided into three domains: (1) diagnosis, (2) prevention and (3) treatment of five common shoulder injuries, including subacromial impingement syndrome (SIS), internal posterosuperior impingement, biceps-labrum complex injuries, rotator cuff injuries, and anterior instability, but also unspecified shoulder pain or injuries. The SIS is not well defined but includes structural damage in the subacromial space, including tendinopathies, partial rotator cuff ruptures, bursitis—without distinction to the mechanism behind it.13 Internal posterosuperior impingement is a frequent cause of pain among overhead athletes, characterised by an increased or repeated contact between the rotator cuff tendons attachment on the tuberculum majus humeri and the posterosuperior aspect of the glenoid cavity when the arm is in an maximally abducted and externally rotated position.13 Clinical studies indicate that its cause is multifactorial, and several anatomical structures are involved in the condition. Biceps-labrum injuries (including SLAP) involve the biceps’ long head and its attachment to the glenoid labrum and can involve the labrum in varying degrees as an actual SLAP lesion.15 Rotator cuff injuries include tendinopathies and partial or complete tendon ruptures, most related to the supraspinatus tendon.16 Anterior shoulder instability may be caused by previous traumatic shoulder dislocation or repeated stress on the anterior structures with prolonged microtraumas and anterior subluxations.17 As such, the selected diagnoses and common shoulder injuries include athletes with any physical shoulder complaint irrespective of the need for medical attention or time loss,18 and represent the spectrum from acute, sudden onset injuries to repetitive, gradual onset injuries, including unspecified shoulder pain or injuries. However, we excluded diagnoses likely to be solely of traumatic origin (ie, dislocations and fractures) because athletes presenting with such injuries are expected to have a different clinical profile and management needs.

Search

Three systematic searches covering diagnosis, prevention, and treatment of the five selected shoulder injuries were performed in MEDLINE via PubMed and EMBASE via Ovid in April 2021 and updated August 2022, with no restrictions in the year of publication. Only publications in English were included. We searched individual text words in the title/abstract and supplemented them with Medical Subject Headings terms (MeSH). We combined anatomical region of interest (eg, ‘Shoulder (MeSH)’ OR ’rotator cuff’ OR ’shoulder’) AND the type of shoulder problem (eg, ‘Athletic injury (MeSH)’ OR ‘injur*’ OR ’strain*’ OR ’impingement’) AND outcome for diagnosis and prevention (eg, ’Physical examination (MeSH)’ OR ’test’ OR diagnostic*’ AND ’Primary prevention (MeSH)’ OR prevent*, respectively) or intervention for treatment (eg, ’Physical therapy modalities (MeSH)’ OR ’exercise’ OR ’education’ OR ’Shockwave’). For treatment and prevention, the search strategy also included research design (eg, ’Systematic review (MeSH)’ OR ’meta-analysis’ OR ’randomised controlled trial’). In addition, reference lists of the included studies were screened for potential eligible references. The complete search strategy for all searches is available in online supplemental file 1.

Selection of studies

Two authors (BL and JRP) independently screened studies using Covidence systematic review software (Veritas Health Innovation, Melbourne, Australia). Articles were initially screened by title and abstract for eligibility. Full-text articles were then screened for inclusion. Disagreements were solved by consensus. For all domains, studies were included based on the highest level of available evidence.19 For the diagnostic domain, systematic reviews and diagnostic cohort studies were preferred as these represent the highest starting point in the Grading of Recommendations Assessment, Development and Evaluation (GRADE) tool (3). We aimed to include studies comparing clinical tests to either arthroscopy, MRI, or ultrasound as the diagnostic reference standard. For the prevention and treatment domains, systematic reviews or randomised controlled trials (RCTs) were preferred as these represent the highest starting point in the GRADE tool (3). However, observational studies were also eligible for inclusion. For the prevention domain, all shoulder problem definitions were accepted (eg, time-loss, medical attention, severity scores, patient-reported outcomes (eg, pain)). Similarly, for the treatment domain, all effect measures were accepted (eg, risk of reinjuries, time to return to sport, number of criteria passed or patient-reported outcomes (eg, pain, physical function in sport and recreational activities)).

Appraisal

Two authors (BL and JRP) independently assessed the risk of bias in the included studies. The Cochrane Risk of Bias assessment tool 2.0 was used for RCTs,20 QUADAS-2 for diagnostic studies,21 ROBINS-I for non-randomised studies of interventions,22 and ROBIS for systematic reviews.23 Details of each tool can be found in online supplemental file 2. Disagreements were solved by consensus. If a systematic review included a risk of bias assessment of individual studies using one of the aforementioned assessment tools, no additional risk of bias assessment was conducted for these individual studies. However, if these tools were not used, or there was conflicting risk of bias assessments between two or more systematic reviews, risk of bias was reassessed. Two authors (BL and JRP) evaluated the overall quality of evidence for each outcome relating to diagnosis, prevention and treatment, following the GRADE approach.24 Disagreements were solved by consensus. The quality of evidence was graded as (1) high, indicating that further research is unlikely to alter the confidence in the estimates, (2) moderate, indicating that further research is likely to have an important impact on the confidence in the estimates and may change the estimates,24 low, indicating that further research is very likely to have an important impact on the confidence in the estimates and is likely to change the estimates or (4) very low, indicating high uncertainty about the estimates. For the prevention and treatment domains, the
starting quality of evidence was 'high' for data based on RCTs and 'low' for data based on observational studies. For the diagnostic domain, the starting quality of evidence was 'high' when based on cohort studies. Subsequently, the quality of evidence could be downgraded one or two levels for each of five domains, including study limitation (ie, serious risk of bias),23 imprecision (ie, wide CIs around estimates or few studies included),26 inconsistency (ie, heterogeneity of results cross studies if more than one study was included for a specific outcome/interval),27 indirectness (ie, lack of generalisability of the findings to the target population)28 and publication bias (ie, results depend on the sample size).29 The quality of evidence could be upgraded due to a large magnitude of effect, dose-response gradient, and effect of plausible residual confounding. The GRADE assessments can be found in online supplemental file 3.

Data synthesis
Diagnostic tests
The diagnostic accuracy of clinical tests was estimated as positive (LR+) and negative (LR−) likelihood ratios and expressed the change in odds of the patient having the problem.30 An LR+ >1 increases the post-test probability of a diagnosis following a positive test, while an LR− <1 decreases the post-test probability of a diagnosis following a negative test. The diagnostic accuracy was classified based on current guidelines as: very low (LR+ 1 to 2; LR− 0.5 to 1), low (LR+ 2 to 5; LR− 0.2 to 0.5), moderate (LR+ 5 to 10; LR− 0.1 to 0.2) and large (LR+ >10; LR− <0.1).31 When available, positive predictive values (PPV) and negative predictive values (NPV) were extracted or calculated from data presented in the individual studies. The PPV expresses the probability that a patient with a positive test had the problem in question, while NPV expresses the probability that a patient with a negative test did not have the problem. The diagnostic accuracy was classified based on current guidelines as: high (≥0.85), moderate (0.70–0.84) and low (≤0.69).32

Prevention and treatment
Due to substantial clinical heterogeneity in types of shoulder injuries, and intervention and population characteristics, pooling of results were not possible and a narrative synthesis was adopted. For dichotomous prevention outcomes (ie, problem/no problem), the effect of the intervention was presented as ORs, HRs, or incidence rate ratios. For continuous prevention outcomes (eg, pain measured using patient-reported outcomes, number of pain episodes), the effect of the intervention was presented as between-group differences. For the treatment domain, the effect was presented as between-group differences or, if between-group differences were not available, as within-group changes.

RESULTS
The main findings are summarised in figure 1.

Domain 1: diagnostic tests
A total of 1286 studies were identified in the literature search. No studies were identified on the diagnostics of the selected shoulder injuries in athletes. Data on diagnostic tests were therefore obtained from studies including both athlete and non-athlete populations. Five studies (two systematic reviews33,34 and three prospective studies35–37) were included (table 1). No meta-analysis of diagnostic accuracy was available in the literature or conducted as a part of this statement paper.

Subacromial impingement syndrome
Several tests for SIS were examined in a systematic review by Hegedus et al,33 of which one test battery, the Composite test, used five of the most commonly used tests in clinical practice (Hawkins-Kennedy, Neer, Painful arc, Empty can/Jobe and external rotation against resistance) and was included in this statement paper. Based on one prospective study (n=55), the composite test showed low diagnostic accuracy, with a low agreement between clinical testing and arthroscopic examination (LR+=2.93, LR−=0.34; low quality of evidence).38 Based on one prospective study (n=69), the posterior impingement test showed low to moderate diagnostic accuracy in predicting a positive or negative arthroscopic examination (LR+=5.0, LR−=0.29; not possible to assess the quality of evidence as the full-text version of the article was not accessible).37

Anterior instability
Eight studies, including four clinical tests, were included in the systematic review by Hegedus et al.33 Based on two studies (n=409), the apprehension test showed high diagnostic accuracy when positive (LR+=17.21) and low diagnostic accuracy when negative (LR−=0.39) in predicting a positive or negative arthroscopic examination, respectively (low to moderate quality of evidence). The diagnostic accuracy of the relocation test was investigated in three studies (n=509) and showed moderate diagnostic accuracy when positive (LR+=5.48) and low diagnostic accuracy when negative (LR−=0.55) in predicting a positive and negative arthroscopic examination, respectively (very low quality of evidence). A combination of the apprehension test and relocation test showed high diagnostic accuracy (LR+=39.68, LR−=0.19) in predicting a positive and negative arthroscopic examination (one study, n=46) (moderate quality of evidence). Based on two studies (n=128), the surprise test showed moderate diagnostic accuracy when positive (LR+=5.42) and low diagnostic accuracy when negative (LR−=0.25) in

Figure 1 Summary of main findings about diagnostic tests, prevention and treatment of common shoulder injuries in sport.
## Table 1  Diagnostic accuracy of clinical tests and grading the quality of evidence

<table>
<thead>
<tr>
<th>Clinical test</th>
<th>Likelihood ratio</th>
<th>Positive and negative predictive values</th>
<th>Diagnostic accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>High*</td>
</tr>
<tr>
<td><strong>Arthroscopy used as reference standard</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composit test (combination of Hawkins-Kennedy, Neer, Painful arc, Empty can/Jobe, external rotation against resistance)</td>
<td>LR+=2.93</td>
<td>LR−=0.34</td>
<td>Low quality of evidence</td>
</tr>
<tr>
<td>Internal posterosuperior impingement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posterior impingement test</td>
<td>LR+=3.0</td>
<td>LR−=0.29</td>
<td>Low quality of evidence</td>
</tr>
<tr>
<td>Anterior instability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apprehension</td>
<td>LR+=17.21</td>
<td>LR−=0.39</td>
<td>Moderate quality of evidence</td>
</tr>
<tr>
<td>Relocation</td>
<td>LR+=5.48</td>
<td>LR−=0.55</td>
<td>Very low quality of evidence</td>
</tr>
<tr>
<td>Surprise</td>
<td>LR+=5.42</td>
<td>LR−=0.25</td>
<td>Very low quality of evidence</td>
</tr>
<tr>
<td>Apprehension + relocation</td>
<td>LR+=39.68</td>
<td>LR−=0.19</td>
<td>Moderate quality of evidence</td>
</tr>
<tr>
<td><strong>SLAP</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biceps load II</td>
<td>LR+=26.38</td>
<td>LR−=0.11</td>
<td>Moderate quality of evidence</td>
</tr>
<tr>
<td>Biceps-Labrum complex injuries</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O’Brien’s active compression; Inside</td>
<td>LR+=1.62</td>
<td>LR−=0.27</td>
<td>Low quality of evidence</td>
</tr>
<tr>
<td>O’Brien’s active compression; Junctional</td>
<td>LR+=2.48</td>
<td>LR−=0.15</td>
<td>Low quality of evidence</td>
</tr>
<tr>
<td>O’Brien’s active compression; Bicipital tunnel</td>
<td>LR+=2.00</td>
<td>LR−=0.08</td>
<td>Low quality of evidence</td>
</tr>
<tr>
<td>Throwing test; Inside</td>
<td>LR+=2.32</td>
<td>LR−=0.36</td>
<td>Low quality of evidence</td>
</tr>
<tr>
<td>Throwing test; Junctional</td>
<td>LR+=3.42</td>
<td>LR−=0.35</td>
<td>Low quality of evidence</td>
</tr>
<tr>
<td>Throwing test; Bicipital tunnel</td>
<td>LR+=2.09</td>
<td>LR−=0.40</td>
<td>Low quality of evidence</td>
</tr>
<tr>
<td>Bicipital tunnel palpation; Inside</td>
<td>LR+=1.92</td>
<td>LR−=0.16</td>
<td>Low quality of evidence</td>
</tr>
<tr>
<td>Bicipital tunnel palpation; Junctional</td>
<td>LR+=3.43</td>
<td>LR−=0.09</td>
<td>Low quality of evidence</td>
</tr>
<tr>
<td>Bicipital tunnel palpation; Bicipital tunnel</td>
<td>LR+=2.24</td>
<td>LR−=0.04</td>
<td>Low quality of evidence</td>
</tr>
<tr>
<td>Yergason’s test; Inside</td>
<td>LR+=2.13</td>
<td>LR−=0.04</td>
<td>Low quality of evidence</td>
</tr>
<tr>
<td>Yergason’s test; Junctional</td>
<td>LR+=6.57</td>
<td>LR−=0.83</td>
<td>Low quality of evidence</td>
</tr>
<tr>
<td>MRI or ultrasound used as reference standard</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rotator cuff injury</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Painful arc</td>
<td>LR+=3.70</td>
<td>LR−=0.36</td>
<td>Low quality of evidence</td>
</tr>
<tr>
<td>Gerber/Lift-off test</td>
<td>LR+=1.40–1.50</td>
<td>LR−=0.63–0.85</td>
<td>Very low quality of evidence</td>
</tr>
<tr>
<td>External rotation against resistance</td>
<td>LR+=2.60</td>
<td>LR−=0.49</td>
<td>Low quality of evidence</td>
</tr>
<tr>
<td>Full can</td>
<td>LR+=2.40</td>
<td>LR−=0.37</td>
<td>Low quality of evidence</td>
</tr>
<tr>
<td>Empty can/Jobe</td>
<td>LR+=1.30</td>
<td>LR−=0.64</td>
<td>Very low quality of evidence</td>
</tr>
</tbody>
</table>
predicting a positive or negative arthroscopic examination, respectively (very low quality of evidence).

**Biceps labral complex injuries**

Based on one study 117 (n=127), the Biceps Load II test showed high diagnostic accuracy when positive (LR+=26.38) and negative (LR−=0.11) in predicting a positive and negative arthroscopic examination, respectively (moderate quality of evidence). The diagnostic accuracy of the three-pack test (active compression, throwing test and bicipital tunnel palpation) was investigated for three anatomically different locations in the shoulder joint (inside, junctional and bicipital tunnel) in 116 patients. 116 Active compression in the bicipital tunnel showed high diagnostic accuracy when negative in ruling out injury on arthroscopic examination (LR−=0.08) (moderate quality of evidence). Bicipital tunnel palpation junctional showed high diagnostic accuracy when negative in ruling out injury on arthroscopic examination (LR−=0.09) (moderate quality of evidence). Bicipital tunnel palpation on the bicipital tunnel showed high diagnostic accuracy when negative in ruling out injury on arthroscopic examination (LR−=0.04) (moderate quality of evidence). For the remaining locations (inside and junctional), low to moderate diagnostic accuracy (LR+=1.62 to 3.43, LR−=...
Table 3  Treatment of shoulder injuries: effect and grading the quality of evidence

<table>
<thead>
<tr>
<th>Injury</th>
<th>Effect</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Large</td>
</tr>
<tr>
<td>Subacromial impingement</td>
<td>Within-group baseline and follow-up scores (SD) of pain (NRS) Intervention 7.88 (1.02) to 3.56 (1.31) Control 7.71 (0.83) to 8.00 (0.88)</td>
<td>Moderate quality of evidence</td>
</tr>
<tr>
<td>Supraspinatus tendinopathy</td>
<td>Within-group baseline and follow-up scores (SD) of pain (VAS) Hyperthermia 5.96 (0.83) to 1.2 (0.63) Ultrasound 6.3 (0.86) to 5.15 (0.87) Passive stretches 6.1 (0.89) to 4.9 (0.88)</td>
<td>Moderate quality of evidence</td>
</tr>
<tr>
<td>Shoulder pain</td>
<td>Within-group changes (95% CI) of pain (NRS) Mobilisation 0.6 (0.1 to 1.1) Manual treatment 0.0 (0.0 to 0.3) Attention 0.2 (−0.2 to 0.7)</td>
<td>Moderate quality of evidence</td>
</tr>
<tr>
<td>Rotator cuff injury</td>
<td>Within-group baseline and follow-up scores (SD) of physical function (Constant and Murley score) Hyperthermia 58.5 (3.6) to 82.0 (5.7) Ultrasound 58.92 (8.1) to 61.8 (4.2) Passive stretches 59.5 (2.7) to 63.3 (5.6)</td>
<td>Moderate quality of evidence</td>
</tr>
<tr>
<td>Rotator cuff injury</td>
<td>With in-group baseline and follow-up scores (SD) of pain (VAS) Intervention 7.72 (1.2) to 2.3 (1.9) Control 7.68 (1.3) to 2.9 (1.8)</td>
<td>Very low quality of evidence</td>
</tr>
<tr>
<td>Supraspinatus tendinopathy</td>
<td>Within-group baseline and follow-up scores (SD) of pain (VAS) Hyperthermia 5.64 (0.81) to 2.1 (0.6) Ultrasound 5.9 (0.8) to 5.3 (0.8) Passive stretches 6.1 (0.89) to 5.1 (0.8)</td>
<td>Very low quality of evidence</td>
</tr>
<tr>
<td>Subacromial impingement</td>
<td>Within-group baseline and follow-up scores (SD) of pain (VAS) Intervention 7.88 (1.02) to 3.56 (1.31) Control 7.71 (0.83) to 8.00 (0.88)</td>
<td>Moderate quality of evidence</td>
</tr>
</tbody>
</table>

0.16 to 0.40) were observed for all tests (low to moderate quality of evidence).

Rotator cuff injury
Four studies, including five clinical tests for rotator cuff injury (painful arc, gerber/lift-off test, external rotation against resistance, full can, empty can/Jobe), were included. All tests showed low diagnostic accuracy when positive (LR+ = 1.30 to 3.70) and negative (LR− = 0.36 to 0.85) in predicting positive or negative rotator cuff injury on ultrasound or MRI, respectively (very low to low quality of evidence). One study including three clinical tests for full rotator cuff rupture was included. External lag rotation for infraspinatus and supraspinatus showed moderate diagnostic accuracy when positive (LR+ = 7.2) and low diagnostic accuracy when negative (LR− = 0.57) in predicting positive and negative rotator cuff rupture on ultrasound, respectively (low quality of evidence). Internal rotation lag...
test for subscapularis showed moderate diagnostic accuracy when positive (LR+=5.6) and high diagnostic accuracy when negative (LR=−0.04) in predicting positive and negative rotator cuff rupture on ultrasound, respectively (moderate quality of evidence). Lastly, the drop-sign for infraspinatus and supraspinatus showed low diagnostic accuracy (LR+=3.2, LR=−0.35) in predicting positive and negative rotator cuff rupture on ultrasound (low quality of evidence).

Domain 2: prevention
A total of 932 studies were identified in the literature search. No systematic reviews were identified. Eight RCTs43-49 and one prospective study50 were included (table 2). The highest quality of evidence for a preventative effect was observed for Oslo Sports Trauma Research Center (OSTRC) Shoulder Injury Prevention programme,1 the Shoulder Control programme,47 a throwing injury prevention programme,43 and the FIFA 11+ shoulder injury prevention programme (moderate quality of evidence).44 Performing the OSTRC Shoulder Injury Prevention programme during warm-up three times per week for 7 months resulted in 28% lower odds of shoulder injuries (OR=0.72 (95% CI 0.52 to 0.98)) and 22% reduced odds of substantial shoulder injuries (OR=0.78 (95% CI 0.53 to 1.16)) as measured by the OSTRC Overuse Injury Questionnaire, which covers all shoulder injuries regardless of time-loss from sport or medical attention. The Shoulder Control programme focusing on shoulder and trunk strength and control, trunk mobility and handball throwing load (velocity and frequency) performed three times weekly resulted in a 56% lower shoulder injury rate compared with usual care (HRR=0.44 (95% CI 0.29 to 0.68)).47 A throwing injury prevention programme consisting of stretching, mobility exercises and balance exercises performed at least once a week during warm-up was superior to a 12-month control period in preventing shoulder injuries in young baseball players (HR=0.48 (95% CI 0.21 to 1.08)).43 Performing the FIFA 11+ shoulder injury prevention programme during warm-up in all training sessions for 6 months compared with usual care resulted in a lower incidence rate of shoulder injuries among soccer goalkeepers (IRR=0.28 (95% CI 0.13 to 0.60)).44

Domain 3: treatment
A total of 4040 studies were identified in the literature search. No systematic reviews were identified. Four RCTs51-54 and three prospective studies55-57 including athletes were included (table 3).

Subacromial impingement syndrome
Limited evidence was found to guide the treatment of SIS. A single RCT54 in male baseball players observed a larger within-group reduction in shoulder pain after a programme combining warm-up modalities (laser, ultrasound, cycling, and stretching), shoulder strengthening exercises, and cool down (stretching, ice packs, electrotherapy, compression) performed three times per week for 12 weeks compared with no intervention (0–10 Numeric Rating Scale (NRS) mean (SD): intervention group changed from 7.88 (1.02) to 3.56 (1.31); control group from 7.71 (0.83) to 8.00 (0.88) (moderate quality of evidence). The between-group difference and corresponding 95% CI were not reported. A prospective study57 of 47 young athletes showed reductions in pain and disability following daily shoulder strengthening exercises (Shoulder Pain and Disability Index (mean (SD): from 29.86 (17.03) to 11.70 (13.78)) (very low quality of evidence).

Supraspinatus tendinopathy
One three-arm RCT54 allocated athletes randomly to hyperthermia (434 MHz, heating the skin to 38°C–40°C), continuous ultrasound (1 MHz, intensity 2.0 w/cm²), or exercise intervention. Hyperthermia and ultrasound were performed three times per week for 30 min for 4 weeks, while the exercise intervention was performed twice daily for 4 weeks. The effect of the intervention on pain and physical function was measured at end-of-the intervention using the mean of three Visual Analogue Scale scores (VAS) (at night, at rest and with movement, respectively), and the Constant and Murley score, respectively. The group receiving hyperthermia experienced significant improvements in pain (within-group change (SD): VAS 5.96 (0.83) to 1.2 (0.63), p=0.03) and physical function (within-group change (SD): from Constant and Murley score 58.57 (3.92) to 82 (5.73). The remaining groups experienced smaller and non-significant improvements in pain and physical function (moderate quality of evidence).

Shoulder pain
Two RCTs52 53 and two prospective studies56 57 were included. The highest quality of evidence was found for an intervention of one session of anteroposterior mobilisation of the glenohumeral joint compared with massage and attention controls in improving pain and physical function in 31 athletes from different sports.52 The group receiving anteroposterior mobilisation experienced a statistically significant reduction in shoulder pain measured using an 11-point NRS (within-group change (95% CI): 0.6 (0.1 to 1.1)) at end-of-intervention (moderate quality of evidence). The remaining groups experienced smaller and non-significant reductions in shoulder pain (manual control: 0.0 (0.0 to 0.3), attention control: 0.2 (−0.2 to 0.7)). None of the group experienced a statistically significant improvement in physical function measured using the Disabilities of the Arm, Shoulder, and Hand (mobilisation: 0.3 (−2.7 to 3.4), manual control: 0.5 (−0.3 to 1.3), attention control: 0.7 (−0.6 to 2.0)) (moderate quality of evidence). The remaining studies provided low to very low quality of evidence for the effect of posture correcting exercises, strengthening exercises, and a combination of strengthening and stretching exercises on pain and physical function.53 56 57

DISCUSSION
The following paragraphs discuss the statements for diagnosis, prevention, and treatment, including the limitations in the existing literature.

Diagnosis
Diagnostic accuracy was based on studies with mixed populations with heterogeneous sex, age and activity levels but with clinical presentations likely comparable with athletes with shoulder injuries. High diagnostic accuracy was observed for 10 tests covering anterior instability, SLAP injuries and biceps-complex injuries. The positive apprehension test had the largest LR+ (17.21) of the three most commonly used anterior instability tests, which means that a positive test is highly effective at ruling in anterior instability. Conversely, there is uncertainty about whether a negative test rebuts anterior instability in patients who do not have anterior instability verified by arthroscopy. The combination of the apprehension and relocation tests was evaluated...
in traumatic anterior shoulder instability and showed an even higher LR+. Overall, the use of the apprehension test has good clinical utility for diagnosing anterior shoulder instability. The three-pack test for biceps-labrum complex injuries has been presented with test performances related to the exact point of interest (bicipital tunnel, inside, junctional), with varying diagnostic accuracy. Regrettably, no data is available concerning the combination of the three tests, although the name could indicate that it is a test battery and intended to be used as such. However, from a clinical perspective, the data supports the relevance of the three-pack tests to rule out an injury, while we propose that a positive test must be interpreted with caution. The Biceps Load II test had a high LR+ and low LR− for ruling in and out a SLAP injury in patients with shoulder pain and can be either performed alone or after a positive test in one or more of the three-pack tests. The internal rotation lag test had a low LR− to rule out rupture of the subscapularis muscle, and moderate diagnostic accuracy was observed for the internal rotation lag and external rotation lag for supraspinatus and infraspinatus. The diagnostic tests available for subacromial impingement, internal postero-superior impingement, and rotator cuff injury (not rupture) showed low to very low diagnostic accuracy and low quality of evidence, and their utility in correctly identifying injury is therefore questionable. Concerning subacromial impingement, the diagnosis describes the clinical entity of a painful and functionally impaired shoulder usually experienced when combining shoulder elevation and rotation. Current data do not support the use of the Composite test, and this corresponds well with the unknown pathogenesis of the injury. Diagnostic ultrasound may be a valuable supportive tool in the clinical assessment of shoulder injuries in sport to evaluate structural integrity.

Prevention
Eight studies investigating the prevention of the selected shoulder injuries were identified, with six including overhead athletes (three handball and three baseball). Four studies had a moderate quality of evidence: (1) The OSTRC Shoulder Injury Prevention programme (moderate effect size) aims to increase internal glenohumeral rotation, external rotation strength and scapular muscle strength, as well as improve kinetic chain and thoracic mobility. However, the exercise programme did not affect the risk factors external rotation strength and internal rotation range of motion. The preventive effect of the programme must therefore be due to other factors; (2) The Shoulder Control programme was tested in 15–19 years handball-players with the aim of increasing shoulder and trunk strength and control, trunk mobility, and handball throwing load by using five principal exercises with four progressive levels performed as part of the warm-up routine. The absolute risk reduction was 11% (95% CI 4% to 18%). (3) The throwing injury prevention programme (large effect size) was tested in 9–11 years old baseball players and consisted of stretching, dynamic mobility and balance training during warm-up and showed a 52% injury risk reduction; (4) The FIFA 11+ shoulder prevention programme has a large effect size and was developed for goalkeepers and consisted of general warming-up exercises, exercises to improve strength and balance of the shoulder, elbow, wrist and finger muscles, and advanced exercises for core stability and muscle control. The evidence of effect of exercise in preventing sports-related shoulder injuries is evolving. Despite few identified studies of moderate evidence quality, the effect sizes are comparable to other body regions. Our findings provide preliminary data to support the implementation of shoulder prevention programmes—in line with the recommendations from a recent consensus statement which concluded that ‘injury prevention programmes/exercises are appropriate to prescribe for athletes of all levels to prevent shoulder injury’. However, three studies of low quality showed no preventative effect of exercise-based interventions, suggesting that investigations into differences in effects of exercise intervention characteristics (ie, frequency, type, dose, volume) is necessary to guide clinical practice. Further, we recognise that effective shoulder injury prevention programmes do not always directly elicit improvements in modifiable risk factors (eg, glenohumeral internal rotation deficit, muscular imbalances, or scapula dyskinesia), and that exercise progressions and exercises individualised to sports-specific movements throughout the whole range of motion may be preferable.

Treatment
The low number of studies investigating the treatment of the selected shoulder injuries in the athletic population is detrimental and calls for action. There was very low to moderate quality of evidence for a large effect of exercise-based treatment for subacromial impingement and shoulder pain. There was moderate quality of evidence of a small effect with manual treatment in overhead athletes on reducing shoulder pain on short term. These findings, although based on little evidence, are in line with current weak recommendations for exercise-based interventions in the general population to reduce pain and increase shoulder function. For the treatment of supraspinatus tendinopathy, hyperthermia treatment resulted in a large effect size in reducing pain and disability compared with ultrasound or pendular swinging and stretching exercises (low quality of evidence). However, this treatment is rarely used in a clinical sports setting and cannot be recommended based on the current study.

Limitations and methodological considerations
This paper has some limitations. First, inconsistent terminology and lack of consensus on definition and classification of shoulder injuries in the existing literature challenged the selection of diagnoses relevant for inclusion in this position statement. As such, we included diagnoses with different mechanisms and presentations that may represent the spectrum from acute, sudden-onset injuries to repetitive, gradual-onset injuries, including shoulder pain. Although we were unable to identify whether shoulder injuries in the individual studies may be the result of trauma or (in)direct contact mechanisms, we excluded diagnoses likely to be solely of traumatic origin (ie, dislocations and fractures). Furthermore, we emphasise that shoulder injuries are often not restricted to single structures, but may involve multiple structures, which may confound a test specific to a single tissue or dysfunction. Chronic shoulder injuries, SIS, supraspinatus tendinopathy, and shoulder pain may be considered within the same spectrum, challenging separate categorisation.

Most studies included a majority of young male athletes. As such, our results may have limited application to female athletes and athletes above the age of 25. In addition, none of the included studies reported data on race, socioeconomic status or other social determinants, thus hindering an evaluation of the applicability of the findings according to these characteristics. Future studies should prioritise the inclusion of more diverse samples and report more detailed information on athlete characteristics. We selected a sample of clinical tests that reflected current best practice for the included shoulder injuries. However, we acknowledge the vast number of tests available in the literature.
Furthermore, as limited evidence was available to guide the diagnosis of shoulder injuries in athletes, we included diagnostic studies of non-athlete populations, but we believe it is unlikely that these tests perform very differently in athlete than in non-athlete populations.

Finally, we did not perform risk of bias assessments of the individual studies if an assessment was already available as part of a systematic review. This was chosen a priori to acknowledge the original study findings but may introduce risk of bias, and potential rating discrepancies may exist due to differences in interpretation of risk of bias domains. However, since GRADE represents an overall assessment encompassing several domains and thus does not rely only on risk of bias, potential discrepancies are unlikely to change the quality of evidence provided in the present statement paper. As we were unable to pool individual study results, inconsistency was not reported in the GRADE assessments. As such, the quality of evidence may change when future studies are published and compared with the current body of literature.

**CONCLUSION**

We have graded the quality of evidence concerning diagnosis, prevention and treatment of common shoulder injuries and provide a comprehensive and up-to-date summary of the best available evidence. High diagnostic accuracy was observed for 10 tests covering anterior instability, SLAP injuries and biceps-complex injuries, and these can be used confidently in a clinical setting. For prevention, based on a few high-quality studies, implementing a shoulder injury prevention programme showed a moderate to large effect size in reducing the risk of shoulder injuries compared with no intervention. For the treatment of the selected shoulder injuries, choosing an active rehabilitation programme seems to be beneficial compared with passive modalities in reducing pain and disability. However, the evidence for most estimates were graded very low to moderate, indicating that future high-quality research may alter our findings.

**Author affiliations**

1Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark
2Centre for Evidence-Based Orthopaedics (CEBO), Department of Orthopaedic Surgery, Zealand University Hospital, Køge, Denmark
3Physical Medicine & Rehabilitation Research - Copenhagen (PMR-C), Department of Orthopedic Surgery and Physical Therapy, Copenhagen University Hospital, Hvidovre, Denmark
4Department of Clinical Research, Copenhagen University Hospital, Hvidovre, Denmark
5Center for General Practice, Aalborg University, Aalborg, Denmark
6Clinical Orthopedic Research Hvidovre (CCHR), Department of Orthopedic Surgery, Copenhagen University Hospital Amager-Hvidovre, Hvidovre, Denmark
7The Research Unit PROgrez, Department of Physiotherapy and Occupational Therapy, Næstved-Slagelse-Ringsted Hospitals, Slagelse, Denmark
8Danish Society of Sports Physical Therapy, Odense, Denmark
9Department of Orthopaedic Surgery, Sports Orthopedic Research Center – Copenhagen (SORC-C), Amager-Hvidovre Hospital, Faculty of Health Sciences, Copenhagen University Hospital, Copenhagen, Denmark
10Department of Physiotherapy and Occupational Therapy, Copenhagen University Hospital Herlev and Gentofte, Denmark

**Twitter** Behnam Liaghat @Behnam_Liaghat, Julie Ranne Pedersen @JulieRPedersen, Rasmus Skov Husted @Husted_RS, Lisbeth Lund Pedersen @Lisliunped, Kristian Thorgbor @KThorgbor and Carsten B Juhl @BoghJuHl

**Acknowledgements**

We would like to acknowledge the Danish Society of Sports Physical Therapy (DSSF) for supporting this project.

**Contributors** BL, JRP, RSH, KT, LLP and CBJ conceived the study idea. BL and JRP performed the systematic searches with input from RSH, KT, LLP and CBJ. BL and JRP conducted risk of bias assessments and GRADE assessments. BL and JRP wrote the initial draft. All authors revised the draft critically and agreed on the final version.

**Funding** The Danish Society of Sports Physical Therapy initiated the project and provided financial support to authors (BL, JRP, RSH, LLP and KT).

**Competing interests** KT is Deputy Editor in BJSM. KT and JRP have received grants from the Danish Society of Sports Physical Therapy unrelated to the present study.

**Patient consent for publication** Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**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 iDs**

Behnam Liaghat http://orcid.org/0000-0002-3050-0028
Julie Ranne Pedersen http://orcid.org/0000-0002-7924-5312
Rasmus Skov Husted http://orcid.org/0000-0003-2020-8792
Lisbeth Lund Pedersen http://orcid.org/0000-0002-3283-622X
Kristian Thorgbor http://orcid.org/0000-0001-9102-4515
Carsten B Juhl http://orcid.org/0000-0001-8456-5364

**References**

Review


31 Jaeschke R, Guyatt GH, Sackett DL. Users’ guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The evidence-based medicine working group. *JAMA* 1994;271:703–7.


33 Hegerud EJ, Goode AP, Cook CE, et al. Which physical examination tests provide clinicians with the most value when examining the shoulder? Update of a systematic review with meta-analysis of individual tests. *Br J Sports Med* 2012;46:964–78.


