BMJ Open

accessible medical research

Inter-examiner reproducibility of clinical tests and criteria used to identify subacromial impingement syndrome

Mikkel Vind,[1] Søren Bie Bogh,[1] Camilla Marie Larsen,[2] Hans Kromann Knudsen,[1] Karen Søgaard,[2] Birgit Juul-Kristensen[2]

ABSTRACT

Introduction: A specific algorithm has been proposed for classifying impingement related shoulder pain in athletes with overhead activity. Data on the inter-examiner reproducibility of the suggested clinical tests and criteria and their mutual dependencies for identifying subacromial impingement symptoms (SIS) are not available.

Objective: To test the inter-examiner reproducibility of selected tests and criteria suggested for classifying SIS and the mutual dependencies of each of the individual tests and SIS.

Method: A standardised three-phase protocol for clinical reproducibility studies was followed, consisting of a training, an overall agreement and a study phase. To proceed to the study phase, an overall agreement of 0.80 was required. In total 10, 20 and 44 subjects were included in the three phases, respectively. The case prevalence in the study phase was 50%. The inclusion criterion for cases was ≥3, and for controls ≤1 positive test out of four. Cohen's κ statistics were used for calculating agreement.

Results: In the overall agreement phase, an agreement of 0.90 was obtained, while in the study phase it was 0.98 with a κ of 0.95 for SIS. κ Values for the individual tests varied between 0.60 and 0.95. Mutual dependencies between each test and SIS showed Neer's test with anterior pain to be most often used to determine SIS.

Conclusions: Inter-examiner reproducibility was moderate to almost perfect for the selected tests and criteria for SIS. The next challenge will be to establish reproducibility in clinical practice, as well as the validity of the tests and criteria for SIS.

[1]Department of Physiotherapy, University College Lillebaelt, Odense, Denmark
[2]Research Unit for Musculoskeletal Function and Physiotherapy, Institute of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark

Correspondence to
Mikkel Vind; vindbie@live.dk

ARTICLE SUMMARY

Article focus

- This study examines the inter-tester reproducibility of Hawkins', Neer's, Jobe's and Apprehension tests, included in a recent clinical reasoning algorithm, both as individual tests and as performed as a battery of tests for screening for subacromial impingement symptoms (SIS) in overhead athletes.
- Furthermore, the mutual dependency of each individual test and SIS was determined.

Key messages

- Overall, the reproducibility was almost perfect for tests and criteria for SIS when following a standardised test protocol for reliability and validity studies.
- Mutual dependency was highest for Neer's test with anterior pain and SIS, implying that this was the test most often used to determine SIS.

Strengths and limitations of this study

- The strength of the study is the strict adherence to a standardised study protocol for reproducibility and validity studies, with a training and an overall agreement phase, including the requirement of a 0.50 prevalence of SIS, before performing the study phase. Since the study population consisted of overhead athletes, for whom the algorithm is intended, data can be generalised to the target population.
- The limitation of the study is that the current mutual dependencies may have been deceptively high because of a possible bias due to the inclusion criteria used in this algorithm, since subjects with two out of four positive test responses are not included. However, this limitation has not biased the high reproducibility of the tests and criteria for SIS.

INTRODUCTION

Fifty per cent of the general population experience shoulder pain every year.[1] Subacromial impingement syndrome (SIS) is the most common shoulder disorder in the population, representing 44−65% of all registered shoulder complaints in the clinical setting.[2] Its prevalence is especially high in sports with overhead activity, such as swimming, volleyball, handball and badminton. These overhead athletes have a high demand for optimal

shoulder performance and dynamic stability is required in order to prevent injury.[3−7]

Until now, historical information and clinical signs have been included in the clinical examination as the most common tools for diagnosing shoulder impingement, even though radiological examinations have

improved the accuracy of the diagnosis.[8] MRI has recently been recommended as the best method to identify shoulder lesions, while ultrasonography is also recognised as highly accurate in detecting SIS.[9] [10] However, neither method is able to identify subtle instability and scapula dyskinesis, while MRI findings may be unable to differentiate between symptomatic and asymptomatic shoulders or to provide a structural diagnosis.[1] Structural changes can be present without pain, thus giving a high number of false positive and negative results.[1] In the clinical setting, manual diagnostic tests are fast and easy to perform compared with imaging techniques, which are both time-consuming and expensive. A recent systematic review concluded that several clinical shoulder tests have sufficient sensitivity, but inadequate specificity.[11] Neer's and Hawkins' impingement tests, in particular, have been found useful in confirming SIS, but poor at ruling out pathology.[9] [12–14] Furthermore, Jobe's test and passive abduction had a relatively high sensitivity of 74% compared with MRI, and were described as good diagnostic tools.[9] A systematic review found Neer's and Hawkins' tests to have a sensitivity and specificity of 79% and 53%, and 79% and 59%, respectively, and states that no shoulder test can by itself confirm impingement.[11] It is further suggested that more large, prospective, well-designed studies should be carried out.[11] A recent prospective study of patients admitted for surgery found Neer's test, the Painful Arc test and the External Rotation Resistance test to be excellent screening tools to rule out SIS.[15] Furthermore, Painful Arc, the External Rotation Resistance test in the neutral position and Jobe's test had the highest diagnostic accuracy compared with surgical findings, and were best at confirming/ruling in SIS. Inter-examiner reliability showed κ values ranging from 0.39 to 0.67, but the protocol lacked an overall agreement phase, and this may have influenced κ values negatively.[15] A stronger design is therefore needed to establish more accurate κ values for inter-examiner reliability.

Recently, a clinical reasoning algorithm based on clinical tests has been developed as a screening tool for the detection of different shoulder pathologies in overhead athletes (figure 1).[16] The type of impingement is decided upon, and then the underlying pathologies are examined by a battery of commonly accepted shoulder tests. For classification of SIS and 'internal posterio-superior glenoid impingement', a group of four tests is performed (Jobe's test, Neer's test, Hawkins' test and the Apprehension test). The use of a group of tests as a diagnostic tool is supported by the results in a recent study, where three or more positive tests out of five confirmed SIS, and less than three positive tests out of five ruled out SIS.[15]

However, the algorithm has not been tested for reliability or validity, and there are no guidelines regarding how many positive test responses are required in order to differentiate between SIS and internal posterio-superior glenoid impingement. In order to validate this algorithm, it is necessary to establish the reproducibility. Therefore, the aim of this study was to determine the inter-examiner reproducibility of selected tests and criteria for classifying SIS based on the algorithm, and second, to test the mutual dependency of the individual tests and SIS.

METHODS
Study design
The study was carried out using the protocol for diagnostic procedures in reproducibility studies presented by the International Federation of Manual/Musculoskeletal Medicine (FIMM).[17] The protocol for a reproducibility study is a three-phase study consisting of a training phase, an overall agreement phase (>80%) and an actual study phase with a case prevalence of 50%. Two examiners (examiners A and B), with a maximum of 6 months' clinical experience, blinded to their mutual test results, performed all tests. The subjects were

Figure 1 Algorithm for clinical reasoning in the examination of impingement related shoulder pain (from Cools *et al*[16]). IR, internal rotation; GIRD, glenohumeral internal rotation deficit; ROM, range of motion; SLAP, superior labrum from anterior to posterior tear; SAT, scapular assistance test; SRT, scapular retraction test.
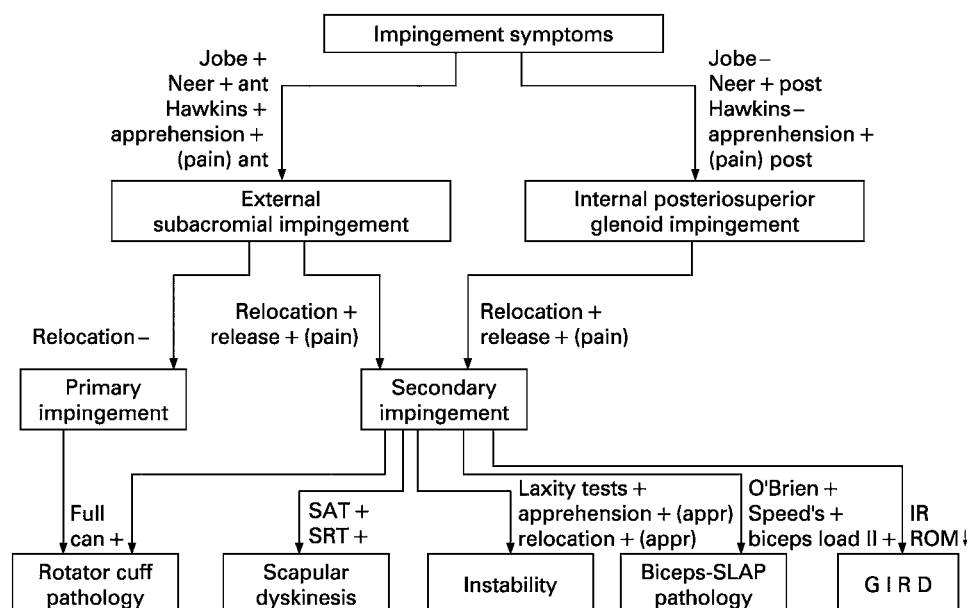
instructed not to give information about results from previous tests.

In the training phase, the test manual was refined on the basis of feedback, and the performance and interpretation of the tests were evaluated in order to remove potential errors. The four tests (Jobe, Neer, Hawkins and Apprehension) were used as a battery of tests to detect SIS. In the overall agreement phase, examiner A tested one sample first (11 players) and examiner B tested another sample first (nine players), and then the examiners switched samples and agreements were calculated. In the study phase, the 0.50 prevalence method was applied in order to obtain an accurate κ coefficient. In this phase, one examiner selected and tested a minimum of 20 subjects (10 cases and 10 controls) and sent them to the other examiner, and vice versa. The criterion for inclusion of subjects in the study phase was therefore either SIS (≥3 positive tests) for cases or no-SIS (≤1 positive tests out of four) for controls. A prevalence of 50% was obtained through this method (figure 2). This method is recommended when one diagnostic test is performed, and therefore all four tests were reduced to one diagnostic criterion classification, either SIS or no-SIS.

Materials
According to the study protocol, the training phase was carried out on 10 subjects (three male and seven female physiotherapy students with a mean±SD age of 25±4.1 years), and the overall agreement phase was carried out on 20 subjects (active male handball players with a mean±SD age of 19.9±3.7 years) with the domi-

nant shoulder being tested, comprising 15 right and five left shoulders. In total, 30 males and 14 females with a mean±SD age of 19.6±5.4 years, comprising 35 right and nine left shoulders, respectively (n=44), met the criteria for inclusion in the study phase. To include these 44 subjects, 134 shoulders of active overhead athletes were tested (figure 3). The study was approved by the Ethics Committee in Science of the Region of Southern Denmark.

Clinical tests
The tests were performed according to the recommendations of Cools et al[16] (table 1).

Jobe's test/empty can test
Jobe's test/empty can test was performed with the subject standing, and the arm to be examined in 90° elevation in the scapular plane in maximal internal rotation. The angle of the scapular plane was set to 40° as mentioned elsewhere in the literature.[18][19] This angle was marked on the floor to guide subject movement. Manual resistance prevented further elevation of the arm. The test was positive if the subject experienced pain.[16] Jobe's test, as a single test for impingement, has been found to have almost perfect levels of agreement.[12]

Neer's test
Neer's test was performed with the subject sitting on a chair. The examiner performed a maximal forward flexion of the subject's arm, while elevation of the scapula was prevented. The test was positive when pain was located in the anterior aspect of the shoulder ($Neer_{ant}$). Pain to the posterior aspect was also recorded. Combined with reports of anterior pain, these reports were called 'Neer general' ($Neer_{gen}$).

Hawkins' test
Hawkins' test was performed with the subject sitting on the examination table with feet on the ground. The arm was held in 90° forward flexion, elbow flexed to 90°. Passive internal rotation of the shoulder was performed. The test was positive if pain was reported. Hawkins' test
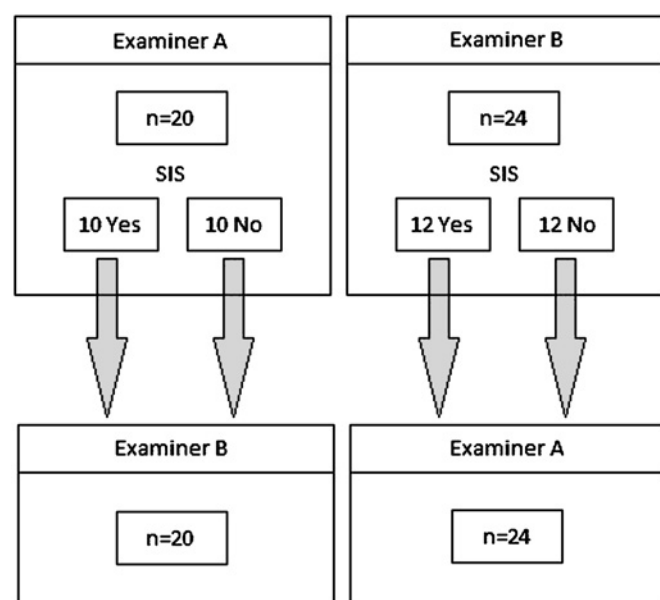
Figure 2 Flow diagram of the 0.50 prevalence method (study phase). Examiners A and B perform the battery of tests on their own groups of subjects. Both examiners sent the same number of positive and negative tests to each other (minimum of 10 positive and 10 negative).
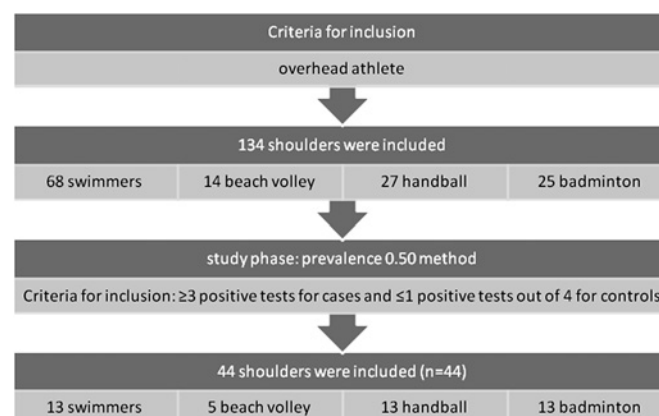
Figure 3 Flow diagram of the subjects included in the study phase.

Table 1 Performance of the clinical tests for SIS (from Cools et al[16]), described as if performed on the right shoulder

Introduction	1. I am going to perform four shoulder tests on you.
	2. For each different test I will ask if you experience any pain, and you must answer either yes or no.
	3. Sometimes I will ask if the pain is in the front or the back of your shoulder, and you use your judgement. Pain down to the insertion of the deltoid muscle is accepted as a positive response (show!).
Preparation	For Jobe's test a line is marked on the floor, on which the subject is asked to stand. On this line an angle of 40° is marked, which is used to point out the plane of the scapula.

	Description	Placing of hands, etc	Evaluation
Jobe's 	Subject is standing. The right shoulder is elevated 90° in the plane of the scapula and maximally internally rotated (empty can position). The examiner resists further elevation of the arm (isometric hold).	The left hand of the examiner is placed on the arm of the subject proximally to the wrist and downward pressure is applied.	Does this hurt your shoulder? Yes/no
Neer's 	Subject is sitting. The scapula is fixated/stabilised in a depressed position, while the shoulder is maximally forwardly flexed.	The left hand of the examiner is placed over the scapula. The right hand grasps distally to the elbow, and subsequently passively flexes the right shoulder of the subject.	Does this hurt your shoulder? Yes/no Does it hurt in the front of the shoulder or in the back?
Hawkins' 	Subject is sitting on the examination table. The shoulder is placed in 90° of forward flexion and passively internally rotated as far as possible (elbow flexed 90°).	The examiner stands in front of the subject. The right arm of the subject is placed in 90° of flexion. The left hand of the examiner is placed on the left shoulder of the subject, and serves to support the right arm of the subject. The examiner grasps the right arm of the subject proximally to the wrist and performs internal rotation of the right shoulder of the subject as far as possible.	Does this hurt your shoulder? Yes/no

Continued

Table 1 Continued

	Description	Placing of hands, etc	Evaluation
Apprehension	Subject is supine on the examination table. The subject is asked to move to the edge of the table. The shoulder is placed in 90° of horizontal abduction and is externally rotated to the end of range.	The examiner grasps the lower arm and performs a passive external rotation. The right hand of the examiner grasps the humerus for support.	Does this hurt your shoulder? Yes/no Does it hurt in the front of the shoulder or in the back?

Results are noted on the score sheet by the examiner.
SIS, subacromial impingement syndrome.

was found to have almost perfect levels of agreement,[12] and the test has been recommended for use in screening for impingement.[11]

The Apprehension test

The Apprehension test was carried out with the subject lying supine on the examination table. The shoulder (90° abduction, the elbow in 90° flexion) was passively placed in maximal external rotation. The test was considered positive if the subject reported pain from the anterior aspect of the shoulder.[16]

Statistics

For the dichotomous data, Cohen's κ statistics were used to calculate levels of agreement between the two examiners (inter-examiner reproducibility) with 95% CIs, and the relationship between the individual tests and the classification of SIS (mutual dependency). Furthermore, values for observed agreement, prevalence and expected agreement were calculated. Data were analysed in Microsoft Office Excel 2007, using 2×2 contingency tables. The κ values were interpreted at five levels according to Landis and Koch.[20]

RESULTS

The groups of cases and controls were comparable on age with mean (±SD) ages of 19.1±5.8 years and 20.0±5.1 years, respectively. The overall agreement phase was completed with an observed agreement of 90% (table 2), meaning that the criterion of a minimum of agreement of 80% was met.

In the study phase, an 'almost perfect' κ value of 0.95 was calculated for SIS (table 3), while κ values for the specific tests varied between 0.60 and 0.95. κ for Neer's test was calculated as 0.86 for anterior pain (Neer$_{ant}$) and as 0.95 for pain in general (Neer$_{gen}$).

The κ values for mutual dependency indicate that Neer's test with anterior pain (Neer$_{ant}$) was the most frequently used test for classifying SIS, with Jobe's test being the least frequently used (table 4). The highest κ values were obtained on the Neer$_{ant}$ with κ values for examiner A and B of 0.82 and 0.91, respectively, followed by Apprehension$_{ant}$ with κ=0.77 and κ=0.73, Hawkins' with κ=0.73 and κ=0.68 and Jobe's with κ=0.64 and κ=0.60.

DISCUSSION

The results from this study showed an almost perfect κ value for SIS (κ=0.95) and for the individual tests (κ values between 0.60 and 0.90). Calculations of mutual dependencies found the highest κ values for the Neer$_{ant}$ test followed by the Apprehension$_{ant}$ test, Hawkins' test and Jobe's test. Our study included an overall agreement phase and a 50% case prevalence, which may explain the

Table 2 Agreement regarding number (%) of positive and negative tests in the overall agreement and study phases for SIS, with calculations of observed agreement and κ values

SIS	Total agreement in positive tests (n, %)	Total agreement in negative tests (n, %)	Total agreement (n, %)	κ Value
Overall agreement phase (n=20)	6 (30%)	12 (60%)	18/20 (90%)	–
Study phase (n=44)	22 (50%)	21 (48%)	43/44 (98%)	0.95

SIS, subacromial impingement syndrome.

Table 3 κ Statistics and 95% CIs for SIS and each of the specific tests in the study phase

	κ Value (95% CI)	Prevalence of the index condition	Observed agreement	Expected agreement by chance
SIS	0.95 (0.87 to 1.00)	0.51	0.98	0.50
Test				
Jobe's	0.90 (0.76 to 1.00)	0.34	0.95	0.55
Neer's (ant pain)	0.86 (0.71 to 1.00)	0.53	0.93	0.50
Neer's (general pain)	0.95 (0.86 to 1.00)	0.58	0.58	0.51
Hawkins'	0.60 (0.34 to 0.85)	0.66	0.82	0.55
Apprehension (ant pain)	0.71 (0.59 to 0.98)	0.39	0.86	0.53

ant, anterior; SIS, subacromial impingement syndrome.

higher strength/validity of the present κ values compared to those found in earlier studies examining the reproducibility of clinical shoulder tests.[12][15][21] Without including an overall agreement phase, the risk of a systematic bias cannot be excluded.

Only one study has examined the reproducibility of several tests as a group. Michener et al found that three positive tests out of five were most likely to confirm SIS.[15] The conflicting results in the three studies mentioned above, however, together with the almost perfect levels of agreement in the current study, indicate a need for methodologically stronger study designs in order to minimise errors due to systematic bias. Implementation of a training phase, an overall agreement phase (≥80%) and a study phase including the 0.50 prevalence method in future studies is therefore recommended.

In the current study, Jobe's test produced the highest κ value (κ=0.90) of all the individual tests (table 3), even though the prevalence of 0.34 weakens the credibility of the κ value compared to a prevalence of 0.50.

For Neer's test for anterior pain (Neer$_{ant}$) and Neer's test for pain in general (Neer$_{gen}$), where pain is located either anteriorly and/or posteriorly in the shoulder, the κ values were 0.86 and 0.95, respectively, with a prevalence close to 0.50 (table 3). These results show that a high level of agreement was still obtained even when differentiating between anterior and posterior pain.

Dromerick et al examined Neer's test and presented a substantial κ value of 0.78.[21] Johansson et al[12] examined, among others, Jobe's test, Neer's test and Hawkins' test. High κ values were obtained for all three tests (Jobe's κ=0.94, Neer's test κ=1.0, Hawkins' test κ=0.91), but only Jobe's test was performed with a prevalence close to 0.50.[12] Michener et al also examined Jobe's, Neer's and Hawkins' tests for reproducibility and found κ values ranging from 0.39 to 0.47, with a prevalence of SIS of 0.29,[15] which may explain the relatively low κ.

The Apprehension test obtained a κ value of 0.71, with a prevalence of 0.39 making the result slightly weaker compared with a prevalence of 0.50. The Apprehension test was carried out in accordance with the classical protocol described by Cools et al (90° abduction and maximal external rotation).[16] However, disagreement was found, since Cools et al[16] refer to Meister et al[22] where the test is described and performed differently (with the shoulder abducted 90−110° and in slight extension). Meister et al conclude that posterior pain is a sign of posterio-superior glenoid impingement.[22] However, there are no references in the literature suggesting that the Apprehension test with anterior pain can be used for the detection of SIS. The validity of this modified test, therefore, needs to be verified, that is, the use of a modification of this test, originally designed for internal posterio-superior glenoid impingement, for the

Table 4 κ Statistics for mutual dependency of the individual tests and the diagnosis of SIS

Test	κ Value	Prevalence of the index condition	Observed agreement	Expected agreement by chance
Jobe's				
Examiner A	0.64	0.41	0.82	0.50
Examiner B	0.60	0.44	0.80	0.49
Neer's (anterior pain)				
Examiner A	0.82	0.50	0.91	0.50
Examiner B	0.91	0.54	0.95	0.50
Hawkins'				
Examiner A	0.73	0.57	0.86	0.50
Examiner B	0.68	0.60	0.84	0.50
Apprehension (anterior pain)				
Examiner A	0.77	0.44	0.89	0.50
Examiner B	0.73	0.45	0.86	0.50

SIS, subacromial impingement syndrome.

detection of SIS, which is regarded as an external anterio-superior glenoid impingement syndrome.

According to our results, there is clear coherence between a positive Apprehension$_{ant}$ test and SIS, shown by the substantial κ values for mutual dependency (κ=0.77 for examiner A, κ=0.73 for examiner B). The high κ values could be attributed to an actual relationship between anterior pain/instability and impingement symptoms, as argued by Meister et al.[22] This κ value should therefore be considered with some caution, since this was not a study on the validity of the tests. Whether there is a clear relationship between pain from the anterior aspect of the shoulder in the Apprehension test and SIS, can only be verified by imaging techniques.

Also Neer$_{ant}$ and Hawkins' tests had a relatively high mutual dependency, showing these tests to be closely related to SIS, in contrast to Jobe's test. However, we have no further explanation for this difference in mutual dependency between tests.

The current mutual dependencies could have been influenced by our inclusion criteria, since subjects with two out of four positive test responses were excluded, meaning that our results for mutual dependency are deceptively high. However, the Neer$_{ant}$ and the Apprehension tests have the highest values of mutual dependency, which indicates that these two tests should carry more weight than Jobe's test and Hawkins's test.

No prior studies were found that addressed the mutual dependencies of these tests.

The limitation of the study is that the current mutual dependencies may have been deceptively high because of a possible bias due to the inclusion criteria used in this algorithm, since subjects with two out of four positive test responses are not included. However, this limitation has not biased the high reproducibility of the tests and criteria for SIS. Another limitation is the unknown reproducibility of the current tests carried out by untrained clinicians, which of course might be different from the reproducibility of trained examiners. This means that reproducibility also needs to be tested in a normal clinical environment. In case inexperienced examiners have a low inter-examiner test reproducibility, our study has shown that it should be possible through education and training to obtain high enough skills to perform the tests in a reproducible way.

Reproducibility is only the first step on the path to establishing the diagnostic value of these tests. Therefore, the validity (ie, concurrent, discriminative, predictive, prescriptive) and the diagnostic accuracy (ie, sensitivity and specificity) of this clinical reasoning algorithm must be determined. With respect to internal impingement, exactly the same procedures (tests for reliability and validity) need to be performed.

The strength of the study lies in its design, which followed a standardised protocol, originally presented by FIMM.[17] The two examiners undertook a training phase in order to minimise the risk of bias in the performance of the tests, and an overall agreement of 90% was achieved before carrying out the study phase. The study phase was performed using the 0.50 prevalence method, which strengthens the κ values. Finally, the study was carried out in overhead athletes, for whom the algorithm is intended, making the results relevant for screening purposes within this group.

CONCLUSIONS
The current study on overhead athletes showed almost perfect inter-examiner reproducibility of the tests and criteria for SIS, based on a clinical reasoning algorithm. Also each of the selected tests (Jobe's, Neer's, Hawkins' and Apprehension) presented high levels of agreement and reproducibility. Based on the mutual dependency of each of the individual tests and SIS, Neer's test with anterior pain (Neer$_{ant}$) had the highest level of agreement, and is therefore important in the classification of SIS. Although showing excellent levels of agreement, it is recommended that further research be carried out in order to establish reproducibility in clinical practice, and also the validity of this clinical algorithm.

Competing interests None.

Patient consent Obtained.

Ethics approval The study was approved by the Ethics Committee in Science of the Region of Southern Denmark.

Contributors MV contributed to the conception, design of the study and writing of the article, the description of the tests and practical performance of the study, conducted the statistical analysis, and drafted and approved the final manuscript. SBB contributed to the conception, design of the study and writing of the article, the description of the test manual and practical performance of the study, conducted the statistical analysis, and drafted and approved the final manuscript. CML contributed to the description of the test manual and revised and approved the final manuscript. HKK contributed to the conception and design of the study, contributed to the description of the test manual, and revised and approved the final manuscript. KS contributed to the conception and design of the study, and revised and approved the final manuscript. BJK contributed to the conception, design of the study and writing of the article, and the description of the test manual, and revised and approved the final manuscript. All the authors will act as guarantors for the paper.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Data in 2×2 contingency tables, and the dataset are available from the corresponding author at vindbie@live.dk. Participants gave informed consent for data sharing.

REFERENCES
1. Lewis JS. Rotator cuff tendinopathy/subacromial impingement syndrome: is it time for a new method of assessment? *Br J Sports Med* 2009;43:259–64.
2. Michener LA, McClure PW, Karduna AR. Anatomical and biomechanical mechanisms of subacromial impingement syndrome. *Clin Biomech (Bristol, Avon)* 2003;18:369–79.
3. Belling Sorensen AK, Jorgensen U. Secondary impingement in the shoulder. An improved terminology in impingement. *Scand J Med Sci Sports* 2000;10:266–78.

4. Cools AM, Witvrouw EE, Declercq GA, *et al.* Scapular muscle recruitment patterns: trapezius muscle latency with and without impingement symptoms. *Am J Sports Med* 2003;31:542–9.
5. Bak K, Fauno P. Clinical findings in competitive swimmers with shoulder pain. *Am J Sports Med* 1997;25:254–60.
6. Pink MM, Tibone JE. The painful shoulder in the swimming athlete. *Orthop Clin North Am* 2000;31:247–61.
7. Fu FH, Harner CD, Klein AH. Shoulder impingement syndrome. A critical review. *Clin Orthop Relat Res* 1991;269:162–73.
8. Ardic F, Kahraman Y, Kacar M, *et al.* Shoulder impingement syndrome: relationships between clinical, functional, and radiologic findings. *Am J Phys Med Rehabil* 2006;85:53–60.
9. Silva L, Andreu JL, Munoz P, *et al.* Accuracy of physical examination in subacromial impingement syndrome. *Rheumatology (Oxford)* 2008;47:679–83.
10. Fodor D, Poanta L, Felea I, *et al.* Shoulder impingement syndrome: correlations between clinical tests and ultrasonographic findings. *Ortop Traumatol Rehabil* 2009;11:120–6.
11. Hegedus EJ, Goode A, Campbell S, *et al.* Physical examination tests of the shoulder: a systematic review with meta-analysis of individual tests. *Br J Sports Med* 2008;42:80–92. discussion 92.
12. Johansson K, Ivarson S. Intra- and interexaminer reliability of four manual shoulder maneuvers used to identify subacromial pain. *Man Ther* 2009;14:231–9.
13. Calis M, Akgun K, Birtane M, *et al.* Diagnostic values of clinical diagnostic tests in subacromial impingement syndrome. *Ann Rheum Dis* 2000;59:44–7.
14. Tennent TD, Beach WR, Meyers JF. A review of the special tests associated with shoulder examination. Part I: the rotator cuff tests. *Am J Sports Med* 2003;31:154–60.
15. Michener LA, Walsworth MK, Doukas WC, *et al.* Reliability and diagnostic accuracy of 5 physical examination tests and combination of tests for subacromial impingement. *Arch Phys Med Rehabil* 2009;90:1898–903.
16. Cools AM, Cambier D, Witvrouw EE. Screening the athlete's shoulder for impingement symptoms: a clinical reasoning algorithm for early detection of shoulder pathology. *Br J Sports Med* 2008;42:628–35.
17. Patijn J, Remvig L. *Reproducibility And Validity 2007-Protocol Formats For Diagnostics Procedures In Manual/Musculoskeletal Medicine.* 2007.
18. Dayanidhi S, Orlin M, Kozin S, *et al.* Scapular kinematics during humeral elevation in adults and children. *Clin Biomech (Bristol, Avon)* 2005;20:600–6.
19. McFarland EG, Kim TK. *Examination Of The Shoulder: The Complete Edition.* United States of America, New York: Thieme Medical Publishers, 2006.
20. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biom* 1977;33:159–74.
21. Dromerick AW, Kumar A, Volshteyn O, *et al.* Hemiplegic shoulder pain syndrome: interrater reliability of physical diagnosis signs. *Arch Phys Med Rehabil* 2006;87:294–5.
22. Meister K, Buckley B, Batts J. The posterior impingement sign: diagnosis of rotator cuff and posterior labral tears secondary to internal impingement in overhand athletes. *Am J Orthop* 2004;33:412–15.