FONTYS UNIVERSITY OF APPLIED SCIENCES

A comparative validity study of physical testing and diagnostic musculoskeletal ultrasound of subacromial impingement syndrome: A literature review

Axel Andreas Abrahamsen Hedman[a,*]

a *Department of Physiotherapy, Fontys University of Applied Sciences, Eindhoven, The Netherlands.*
******* *Corresponding Author. E-mail Address: AxelHedman@gmail.com*

Student number: 2145209
Date: 30.05.2013
Version number: 1
Thesis supervisor: Marc Schmitz, E-mail: m.schmitz@fontys.nl

“Always pass on what you have learned.”

-Yoda

Abstract

Background: Subacromial impingement syndrome is the most common diagnosis of shoulder complaints. It is diagnosed with history taking followed by physical examination. The evidence for many of the physical tests is low, and musculoskeletal ultrasound has become popular in the primary care practice for diagnosing shoulder complaints. Accurate diagnosing based on physical testing or musculoskeletal ultrasound could improve prognoses and reduce health care costs.

Objective: To review the available evidence for the Hawkins-Kennedy test, Neer sign, Jobe test and drop-arm test, and compare it with the available evidence for musculoskeletal ultrasound on the rotator cuff pathology of subacromial impingement syndrome.

Methods: Medline and PubMed Central were searched for literature. The studies that were deemed fit for inclusion, were assessed with the QUADAS for methodological quality, and the validity outcome was extracted.

Results: The review included 27 studies, whereas 16 investigated the ultrasound validity and 11 investigated either the Hawkins-Kennedy test, Neer sign, Jobe test or drop-arm test. Ultrasound showed high sensitivity and specificity on full-thickness tears, as well as high specificity on partial-thickness tears and tendinopathy. The Hawkins-Kennedy test and Neer sign showed high sensitivity, and the drop-arm test showed high specificity on rotator cuff pathology. The Jobe test presented a wide variety of outcome on different pathologies.

Conclusion: The ultrasound likelihood ratios presented a valid diagnostic tool for full-thickness tears, as well as to exclude partial-thickness tears. For rotator cuff tendinopathy, ultrasound could help in excluding the pathology.
None of the physical tests had significant likelihood ratios. By combining the Hawkins-Kennedy test and the Neer sign with the drop-arm test help to include or exclude the diagnosis of rotator cuff pathology. The evidence for the Jobe test was inconclusive.

Keywords: *Subacromial impingement syndrome, Rotator cuff pathology, Physical examination, Musculoskeletal ultrasound, Accuracy, Validity, Likelihood ratios.*

Table of content:

Introduction

In the primary care setting, patients presenting with shoulder pain is very common, with studies showing a yearly prevalence of between 5 and 47%, and a lifetime prevalence of 7 to 67% in the general population in the Netherlands. The prevalence of shoulder pain also increases with age.[1] Shoulder complaints often have a negative prognosis for recovery, with 41% of patients having persisting pain after 12months in one study[2], and 51% after 18months in another study.[3] Furthermore, a specific diagnosis such as bursitis, frozen shoulder or rotator cuff tear is reported being an important recovery predictor compared to non-specific diagnoses in patients with upper extremity complaints.[4] In a Dutch study of chronic musculoskeletal pain, the prevalence of chronic shoulder pain was the second highest (15.1%) after chronic low back pain (21.2%) in 2003.[5] Health care costs are also increasing because of the high number of patients suffering from persisting shoulder complaints. Focusing on the improvements of the diagnosis could in turn give a better prognosis for recovery, and a decrease in persisting chronic pain and high health care costs.

Shoulder complaints are often multi-faceted, and can originate from a variety of pathologies, such as labral lesions, bursitis, calcifying tendonitis, tendinopathy, partial- and full-thickness tears of the rotator cuff muscles, which in the clinical practice often presents with similar symptoms.[6–10] The complexity of the shoulder makes it difficult to isolate a single structure with one test, and therefore provide little evidence on the function level of the international classification of functioning, disability and health (ICF).[11] This low specificity and poor understanding of examining shoulder complaints often leads to a diagnosis of “impingement” or subacromial impingement syndrome (SIS).[6–8,12,13]

History-taking and physical examination has long had its place as being the first step in diagnosing patients presenting with musculoskeletal and shoulder complaints[7,12,13], and if indicated the patients are referred for further investigation to musculoskeletal ultrasound, magnetic resonance imaging (MRI) or invasive techniques such as arthroscopy, injection-test or surgery.[14] Surgery, MRI and arthroscopy are utilized as the ‘golden standards’ in diagnosing musculoskeletal complaints, and are often carried out as reference tests in validity studies.[14,15]

SIS is by far the most common (80%) diagnosis of the shoulder complaints.[16] The diagnosis of SIS is clarified by the Royal Dutch Society for Physical Therapy (KNGF)[12] who used the definition; “complaints arising from dysfunction of the structures in the subacromial space, most commonly caused by impinging of the rotator cuff tendons in lifting the arm”. Further divided there is an internal and internal impingement. The definitions of external and internal impingement refers to the location of the actual “impinging”. The internal impingement affects the rotator cuff tendons between the humeral head and the glenoid rim. Differentiation between the underlying pathologies is a difficult matter, and might ask for a different diagnostic approach for each of them. The subacromial bursa and the caput longum tendon of the m. biceps brachii can be involved in the external impingement as well.[7,12] This review will be focusing on the rotator cuff involvement of SIS. Neer was the first to connect SIS to the rotator cuff

muscles.[12] Although, Neer's idea of rotator cuff tendinitis has recently been disproved by a lack of evidence of an actual inflammatory response in histological research. Therefore, a more accurate description of the process of SIS would be rotator cuff tendinopathy which leads into partial- and full-thickness tears.[12,13,17]

For the physical examination of SIS, there is a vast number of tests being used, but with varying accuracy.[13,14,18,19] The physical tests recommended by the KNGF in their "Evidence statement subacromiale klachten"[12] are considered to be the most sensitive among the physical tests for SIS, although the specificity has shown to be generally low.[8,12] The Hawkins-Kennedy test and Neer sign are recommended as tests for SIS in general, as well as the Jobe test and the drop-arm test for testing the integrity of the m. supraspinatus.[12]

Diagnostic musculoskeletal ultrasound (MSU) has found its way into the medical field as a means to diagnose a large variety of pathology.[20,21] Traditionally carried out by radiologists in the secondary care setting, but has later had growing popularity amongst physical therapists and other disciplines in the primary care setting.[22] The benefits of this are that MSU helps diagnose on the structural level of the ICF, while the history-taking and physical examination can cover the activity and participation level complaints.[10] MSU is less invasive as well as more cost- and time- efficient than the other diagnostic imaging devices, one could consider MSU as a diagnostic tool before MRI, arthroscopy or surgery. Some further reasons that MSU should be considered are; the low-to-none risks and higher patient satisfaction when diagnosed with MSU compared to MRI[23], many also suffer from claustrophobia and are unable to undergo an MRI scan. Although, in the diagnosing of SIS, more specifically, rotator cuff tendinopathy, partial-thickness tear and full-thickness tear there is variety in the results of MSU.[8,24–26]

Accurate diagnosing based on physical testing or MSU could reduce time-consuming and costly magnetic resonance imaging or invasive surgical techniques. It would allow physical therapists to give more specific treatment, which would lead to a better prognosis for recovery and better use of health care funds. This leads to the objective of this study; to review the available evidence for the physical tests recommended by the KNGF, and compare it with the available evidence for musculoskeletal ultrasound on the pathology of SIS.

Method

Inclusion and exclusion criteria:

The criteria for this review were defined before the initial search was made. The criteria entailed; type of study, language, participants, outcomes, tests and instruments, year of publication and methodological quality. A summary of the inclusion and exclusion criteria is in the table (Table 1) below.

Type of studies

Only studies that examined the accuracy of musculoskeletal ultrasound or the selected physical tests against a reference test for rotator cuff pathology were included. The selected physical tests being; Hawkins-Kennedy test, Neer sign, Jobe test, and drop-arm test.

Language

The language was limited to English, excluding studies which were not translated to or written in English. This was due to the language limitations of the author.

Participants

The participating patients had to be suspected of having rotator cuff pathology, while being free of systemic, metabolic or inflammatory diseases to be included. Patients were otherwise included from any clinical setting and of any age and gender. Patients with a history of trauma were excluded, due to the possible bio-mechanical arthrogenic disturbances that they could present with.

Outcome

Studies examining reliability or reproducibility were not included. The accuracy studies needed to present tables or text that included enough raw data (true-positive, true-negative, false-positive and false-negative) for the sensitivity and specificity calculations to be made, or a summary of those.

Tests and Instruments

Ultrasound studies had to employ a frequency of 7,5Mhz or higher to be included, as the visual output is easier to analyse with higher frequencies, and is often recommended by ultrasound studies.[26] As recommended, surgery (open or arthroscopic) and MRI was utilized as reference tests.[14,15] Subacromial injection test, ultrasound or other reference tests resulted in exclusion.

Table 1. Summary of inclusion and exclusion criteria

Inclusion criteria:	I. English literature II. Patients with suspicion of rotator cuff pathology III. Accuracy studies examining; Neer sign, Hawkins Kennedy test, Jobe test, drop-arm test or musculoskeletal ultrasound. IV. The use of open surgery, arthroscopy or MRI as reference test.

Table 1. Continued

Exclusion criteria:	I. Patients with systemic, metabolic or inflammatory disease II. Patients with history of trauma III. Ultrasound frequencies <7,5Mhz IV. Lacking data for accuracy calculations V. Unavailable as full-text article

Sources and search terms

The sources utilized to conduct this review were the databases; PubMed (Medline) and PubMed Central (PMC), by applying the search string; "Shoulder Impingement Syndrome" AND "Accuracy" AND in combination with "Ultrasonography", "Physical Examination", "Neer sign", "Hawkins Kennedy", "Jobe test" or "Drop-arm test".

Selection of studies

A step by step approach was used in the selection of studies, by following the inclusion and exclusion criteria found in table 1. The resulting titles from Medline and PMC were first identified as related or unrelated. The abstracts of the related titles underwent another selection procedure with the criteria (Table 1). The last step involved assessing the full-text articles of the related abstracts. When these were deemed fitting, they underwent a methodological quality assessment and data extraction.

Methodological quality assessment

For assessing the methodological quality and clinical relevance of the studies included, the "Quality Assessment of Diagnostic Accuracy Studies" tool[27] (QUADAS tool) was used. As shown in the appendices (Appendix I), the QUADAS tool is a list of 14 questions, were one can answer "Yes", "No" or "Unclear". In a score of seven or higher the research is considered to be of high quality, six out of fourteen or less is considered low quality. The assessing was done by a single independent reviewer. The filled out QUADAS tables are included in the appendices (Appendix II).

Data extraction and presentation

The included studies are presented in tables (Table 2-6) showing the study characteristic and the validity outcome simultaneously. The study characteristics were presented to show the heterogeneity of the background of the included studies. The validity outcome was measured in sensitivity and specificity. The sensitivity and specificity was either duplicated or calculated from raw data using a two by two table (Appendix III).[28]

Results

Identification and selection of available studies

A literature search of the databases Medline and PMC was conducted by a single independent reviewer in the beginning of April, 2013. This search entailed several steps before the relevant studies were identified and selected for inclusion. This series of step is shown below (Figure 1). The preliminary search resulted in an overflow of unrelated studies, and the search terms were further specified. Four earlier reviews (one for ultrasound and three for physical examination) were identified and included as sources, and provided 67 relevant studies. Once these studies were collected, the initial search was conducted. From this search 137 potentially relevant studies were found. Duplicates and studies that did not meet the set inclusion and exclusion criteria were excluded. The remaining 59 studies were retrieved as full-text (five studies could not be retrieved), whereas 32 of these were excluded, leaving this review with 11 studies for the physical examination and 16 studies for ultrasound.

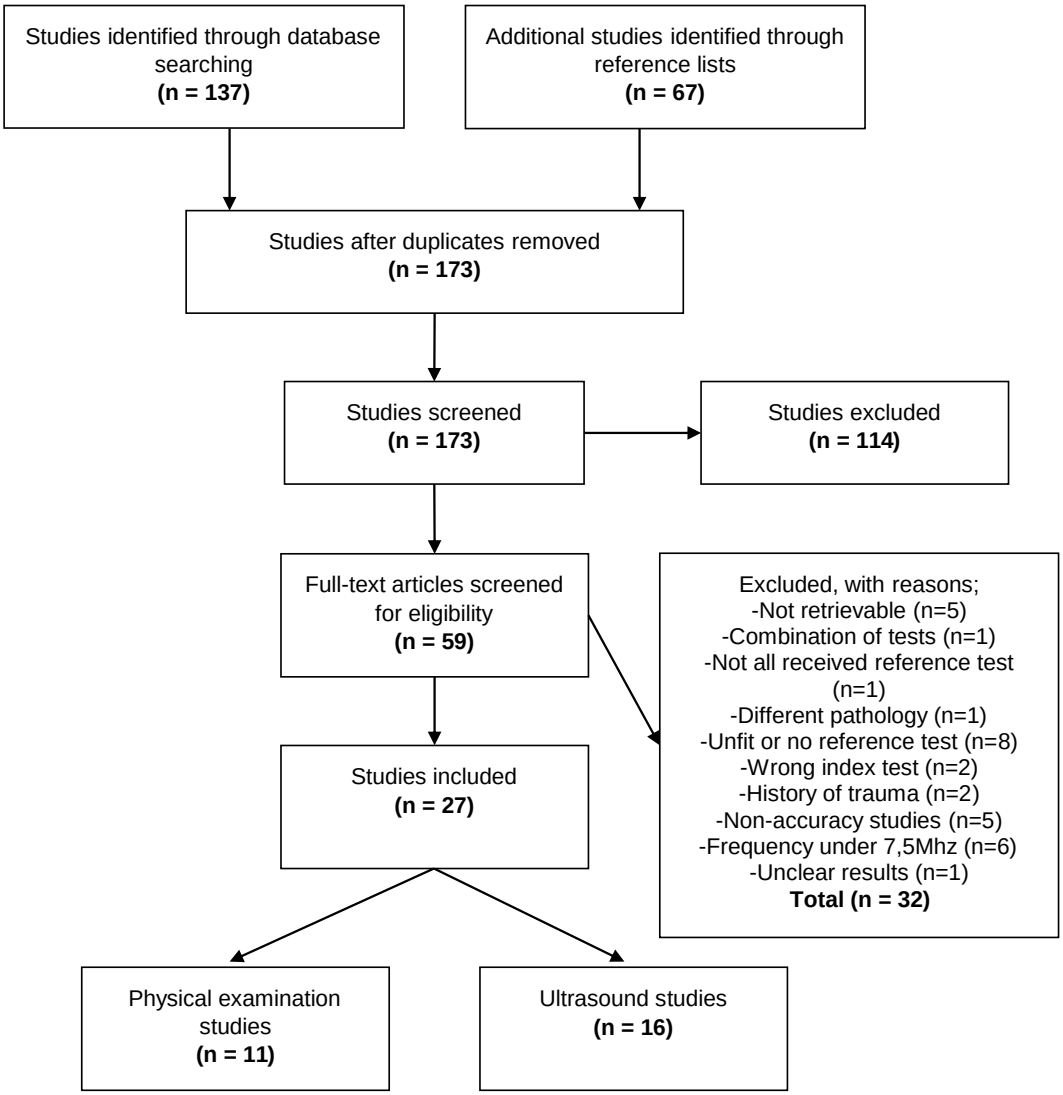

Figure 1. PRISMA flow diagram for identification and selection of studies.

Quality of available research

The methodological quality of the 27 included studies were assessed with the QUADAS tool[21], as previously described. Most studies showed a high QUADAS score by being over seven. In the physical examination studies, question eight was fulfilled by six studies. The 13th question about un-interpretable results were mostly answered "Unclear" in both US and PE studies. Ten out of the 27 studies got a negative point on the 14th question for not explaining the withdrawals from their studies. The structured overview of the QUADAS scores are attached to the appendices (Appendix II), and the total score of each study is included in the study outcome tables (Table 2-6) below.

Presentation of findings

The included study characteristics and their outcomes are presented in the tables below (Table 2-6). Structured alphabetically by the first authors' name, followed by year of publication. The tables include the country of origin, in which clinical settings were the studies and which type of study design they used. The sample size meant the number of shoulders that was involved, and was also reported in the outcome tables with the mean age of the involved participants and the age range. Information about the testing also had to be included in the tables, namely which reference test, which pathology and which diagnostic criteria the studies employed to accurately evaluate the index test. As stated above, the total score of the QUADAS was also added to these tables. The last two columns presents the sensitivity and specificity outcomes of the studies.

Study characteristics

The sample size of the included studies ranged from 20 to 1913 shoulders per study. The total sample size from the ultrasound studies were 1530 shoulders, and 3225 shoulders from the physical examination studies, leaving a total sample size of 4755 shoulders from the 27 included studies in this review.

The mean age of the participating patients ranged from 38 to 63, with the youngest patient being 13 years old and the oldest patient being 86 years old.

The studies were represented by 12 nationalities. The distribution was; seven from the U.S., three each from Japan, Spain and the U.K, two each from Australia and Taiwan (rep. of China), and one from Czech Republic, Finland, the Netherlands, Sweden and Turkey. The study setting was mostly hospital departments, with the exception of nine health care clinics, and all studies reported using open surgery, arthroscopy or MRI as the reference test.

The 16 ultrasound studies (Table 2) all had a QUADAS score of seven or over, the highest quality study[29] showing a score of 12. All the ultrasound studies targeted full-thickness tears, 11 studies targeted partial-thickness tears and two studies targeted tendinopathy of the rotator cuff. One study[30] reported outcome from two groups in diagnosing full-thickness tears, this was due to change in ultrasound operator, the first group was diagnosed by an operator with five years of experience, and the second group was diagnosed by an operator with 10 years of experience. The most frequently

Table 2. Musculoskeletal ultrasound outcome

Study and year of publication	Country, setting and study design	Ref. test†	Q.S.	Sample size (n=) and mean age (range)	Targeted pathology§	Diagnostic criteria*	SE	SP
Al-Shawi et al.[31], 2008	U.K., Orthopaedic centre, Prospective study	AS, MRI	8	n = 148 57 (31 - 82y)	FTT	A, B, C	96	95
Chang et al.[30], 2002	Taiwan, Orthopedic and radiology departments, Retrospective study	S	11	n = 75, NR (NR)	FTT (group 1) FTT (group 2)	A, C, D, E	52 92	92 100
Frei et al.[32], 2008	Czech republic, Orthopedic department, retrospective study	AS	10	n = 20, 56 (NR)	FTT	A, C	100	90
Iannotti et al.[33], 2005	U.S., Orthopedic surgery department, Prospective study	AS, S	8	n = 99, NR (NR)	PTT PTT, FTT	C, D. A, C, D	70 88	89 82
Martín-Hervás et al.[34], 2001	Spain, Orthopedic unit, Prospective study	AS, S	7	n = 61, NR (NR)	T(SS) PTT(SS) FTT(SS)	A, B, C, D, F, G.	67 13 58	88 68 100
Milosavljevic et al.[35], 2005	Sweden, Mixed departments, Unclear Design	AS	9	n = 190, 57 (22 – 78y)	PTT FTT	C. A, C, D	80 100	98 91
Naredo et al.[29], 1999	Spain, Rheumatology and radiology departments, Prospective study	MRI	12	n = 36, 62 (37 – 75y)	T(SS) PTT(SS) FTT(SS)	A, B, C, D, E	93 92 89	100 91 100
Paavolainen & Ahovuo[36], 1994	Finland, Orthopedic and radiology departments, Retrospective study	S	8	n = 49, 38 (24 – 76y)	FTT	A, D	74	95
Read & Perko[37], 1998	Australia, Orthopedic and radiology departments, Unclear design	AS, S	7	n = 42, 44 (19 – 70y)	PTT FTT	A, B, C, D	46 100	97 97
Rutten et al.[15], 2010	The Netherlands, Orthopedic and radiology departments, Retrospective study	AS, S	11	n = 68, 48 (24 – 81y)	PTT FTT	"Established criteria"	89 95	80 93
Sonnabend et al.[38], 1997	Australia, Orthopedic clinic, Retrospective study	AS, S	9	n = 117, 49 (14 – 79y)	PTT FTT	A, C, F	25 84	99 92
Takagishi et al.[39], 1996	Japan, Mixed departments, Unclear design	S	9	n = 122, 51 (26 – 81y)	PTT FTT	C, D, E	50 76	90 100
Teefey et al.[40], 2000	U.S., Institute of radiology and orthopedic department, Retrospective study	AS	9	n = 100, 56 (14 – 82y)	PTT FTT	A, B, C, E	67 100	85 85
Teefey et al.[41], 2004	U.S., Institute of radiology, Prospective study	AS	11	n = 71, 59 (34 – 80y)	PTT FTT	A, C, E	68 98	96 80
Yen et al.[42], 2004	Taiwan, Orthopedic department, Prospective study	S	10	n = 50, 63 (17 – 81y)	PTT, FTT	A, B, C, D, E	95	90
Ziegler et al.[43], 2004	U.S., Orthopedic clinic, Case series	S	10	n = 282, 50 (15 – 84y)	PTT FTT PTT, FTT	A, B, C, D, E	94 96 100	96 94 86

Abbreviations: Sample size - Number of shoulders included; Ref. test - Reference test; Q.S - QUADAS score; NR - Not reported; Se – Sensitivity; Sp – Specificity. .
†Ref. test: AS – arthroscopy; MRI – Magnetic resonance imaging; S – Surgery.
§Targeted pathology: T - Tendinopathy; PTT – Partial-thickness tear; FTT – Full-thickness tear; SS - Supraspinatus.
*Diagnostic criteria: A – non-visualisation of rotator cuff; B – loss of convexity; C – focal discontinuity; D – focal thinning; E – echogenic foci; F – echogenic band; G – presence of bursal fluid.

Table 3. Hawkins-Kennedy test outcome

Study and year of publication	Country, setting and study design	Ref. test†	Q.S.	Sample size (n=) and mean age (range)	Targeted pathology§	Diagnostic criteria	SE	SP
Calis et al.[44], 2000	Turkey, rheumatology and orthopaedic departments, Unclear design	MRI	7	n = 87, NR (18 – 70y)	T, PTT, FTT	Pain	92	25
Fowler et al.[45], 2010	U.K., Sports medicine clinic, Retrospective study	AS	10	n = 101, 40,8 (NR)	T, PTT, FTT	NR	58	72
Jia et al.[46], 2009	U.S., Orthopedic surgery department, Retrospective study	S	6	n = 1913, NR (NR)	T PTT FTT T, PTT, FTT (and MT)	NR NR NR NR	76 75 69 71	45 44 48 42
Macdonald et al.[47], 2000	Canada, University hospital, Unclear design	AS	7	n = 85, 40 (16 – 72y)	T, PTT, FTT	NR	88	43
Michener et al.[48], 2009	U.S., Orthopedic clinic, Prospective study	AS	13	n = 55, 40,6 (18 – 83y)	T, PTT, FTT	Pain	63	62
Nanda et al.[49], 2008	U.K., Orthopedic clinic, Unclear design	AS	10	n = 50, 52 (29 – 70y)	T, PTT, FTT	Pain	94	50
Park et al.[50], 2005	U.S., Orthopedic clinic, Unclear design	AS	11	n = 552, NR (NR)	T PTT FTT T, PTT, FTT (and SAB)	Pain Pain Pain Pain	76 75 69 72	45 44 48 66
Silva et al.[14], 2008	Spain, Rheumatology and radiology department, Prospective study	MRI	13	n = 29, 54 (24 – 82y)	T, PTT, FTT	Pain	74	40

Abbreviations: Sample size - Number of shoulders; Ref. test - Reference test; Q.S - QUADAS score; NR - Not reported; Se – Sensitivity; Sp – Specificity.
†Ref. test: AS – arthroscopy; MRI – Magnetic resonance imaging; S – Surgery.
§Targeted pathology: T - Tendinopathy; PTT – Partial-thickness tear; FTT – Full-thickness tear; MT - Massive tear; SAB - Subacromial bursitis.

Table 4. Neer sign outcome

Study and year of publication	Country, setting and study design	Ref. test†	Q.S.	Sample size (n=) and mean age (range)	Targeted pathology§	Diagnostic criteria	SE	SP
Calis et al.[44], 2000	Turkey, Rheumatology and orthopaedic departments, Unclear design	MRI	7	n = 87, NR (18 – 70y)	T, PTT, FTT	Pain	89	31
Jia et al.[46], 2009	U.S., Orthopedic surgery department, Retrospective study	S	6	n = 1913, NR (NR)	T PTT FTT T, PTT, FTT (and MT)	NR NR NR NR	86 75 59 64	49 48 48 43

Table 4. Continued

Macdonald et al.[47], 2000	Canada, University hospital, Unclear design	AS	7	n = 85, 40 (16 – 72y)	T, PTT, FTT	NR	83	51
Michener et al.[48], 2009	U.S., Orthopedic clinic, Prospective study	AS	13	n = 55, 40,6 (18 – 83y)	T, PTT, FTT	Pain	81	54
Nanda et al.[49], 2008	U.K., Orthopedic clinic, Unclear design	AS	10	n = 50, 52 (29 – 70y)	T, PTT, FTT	Pain	84	50
Park et al.[50], 2005	U.S., Orthopedic clinic, Unclear design	AS	11	n = 552, NR (NR)	T PTT FTT T, PTT, FTT (and SAB)	Pain Pain Pain Pain	84 75 59 68	49 48 47 69
Silva et al.[14], 2008	Spain, Rheumatology and radiology department, Prospective study	MRI	13	n = 29, 54 (24 – 82y)	T, PTT, FTT	Pain	68	30

Abbreviations: Sample size - Number of shoulders; Ref. test - Reference test; Q.S - QUADAS score; NR - Not reported; Se – Sensitivity; Sp – Specificity.
†Ref. test: AS – arthroscopy; MRI – Magnetic resonance imaging; S – Surgery.
§Targeted pathology: T - Tendinopathy; PTT – Partial-thickness tear; FTT – Full-thickness tear; MT - Massive tear; SAB - Subacromial bursitis.

Table 5. Jobe test outcome

Study and year of publication	**Country, setting and study design**	**Ref. test†**	**Q.S.**	**Sample size (n=) and mean age (range)**	**Targeted pathology§**	**Diagnostic criteria**	**SE**	**SP**
Fowler et al.[45], 2010	U.K., Sports medicine clinic, Retrospective study	AS	10	n = 101, 40,8 (NR)	T, PTT, FTT	NR	67	41
Holtby & Razmjou[51], 2004	Canada, Tertiary shoulder clinic, Prospective study	AS, S	9	n = 50, 50 (24 – 79y)	T and PTT. FTT.	Pain. Weakness +/- Pain.	62 41	54 70
Itoi et al.[52], 1999	Japan, Orthopedic surgery department, unclear design	MRI	8	n = 143, 43 (13 – 80y)	FTT (SS)	Pain Weakness Both	63 77 89	55 68 50
Itoi et al.[53], 2006	Japan, University clinic, Case series	S	7	n = 160, 53y (16 – 86y)	T, PTT, FTT (SS)	Pain Weakness	78 87	40 43
Michener et al.[48], 2009	U.S., Orthopedic clinic, prospective study	AS	13	n = 55, 40,6 (18 – 83y)	T, PTT, FTT	Weakness	50	87

Table 5. Continued

Nanda et al.[49], 2008	U.K., Orthopedic clinic, unclear design	AS	10	n = 50, 52 (29 – 70y)	T, PTT, FTT (SS)	Weakness	86	5
Park et al.[50], 2005	U.S., Orthopedic clinic, unclear design	AS	11	n = 552, NR (NR)	T PTT FTT T, PTT, FTT (and SAB)	Pain Pain Pain Pain	25 32 53 44	67 68 83 90
Silva et al.[14], 2008	Spain, Rheumatology and radiology department, prospective study	MRI	13	n = 29, 54 (24 – 82y)	T, PTT, FTT	Pain and weakness	74	30

Abbreviations: Sample size - Number of shoulders; Ref. test - Reference test; Q.S - QUADAS score; NR - Not reported; Se – Sensitivity; Sp – Specificity.
†Ref. test: AS – arthroscopy; MRI – Magnetic resonance imaging; S – Surgery.
§Targeted pathology: T - Tendinopathy; PTT – Partial-thickness tear; FTT – Full-thickness tear; MT - Massive tear; SAB - Subacromial bursitis; SS - Supraspinatus.

Table 6. Drop-arm test outcome

Study and year of publication	**Country, setting and study design**	**Ref. test†**	**Q.S.**	**Sample size (n=) and mean age (range)**	**Targeted pathology§**	**Diagnostic criteria**	**SE**	**SP**
Calis et al.[44], 2000	Turkey, Rheumatology and orthopaedic departments, Unclear design	MRI	7	n = 87, NR (18 – 70y)	T, PTT, FTT	Arm drops +/- severe pain	8	97
Nanda et al.[49], 2007	U.K., Orthopedic clinic, Unclear design	AS	10	n = 50, 52 (29 – 70y)	PTT, FTT	Arm drops +/- severe pain	31	75
Park et al.[50], 2005	U.S., Orthopedic clinic, Unclear design	AS	11	n = 552, NR (NR)	T PTT FTT T, PTT, FTT (and SAB)	Arm drops +/- severe pain	14 14 35 27	77 78 88 88

Abbreviations: Sample size - Number of shoulders; Ref. test - Reference test; Q.S - QUADAS score; NR - Not reported; Se – Sensitivity; Sp – Specificity.
†Ref. test: AS – arthroscopy; MRI – Magnetic resonance imaging; S – Surgery.
§Targeted pathology: T - Tendinopathy; PTT – Partial-thickness tear; FTT – Full-thickness tear; SAB - Subacromial bursitis.

"focal discontinuity" (16 of 16) and "non-visualisation of the rotator cuff" (13 of 16). It was not specified in most articles which were criteria was used for a full-thickness tear and which were for a partial-thickness tear or tendinopathy. The sensitivity and specificity for full-thickness tears ranged from 52% to 100%, and 80% to 100%, respectively. For partial-thickness tears sensitivity ranged from 13% to 94%, and specificity from 68% to 99%, while the tendinopathy focused solely on the tendon of supraspinatus, and had a sensitivity range was from 67% to 93%, and specificity from 88% to 100%. Both partial- and full- thickness tears together had a sensitivity and specificity range from 88% to 100%, and 82% to 90%, respectively.

Among the 11 physical examination studies (Table 3-6), there were eight studies which examined the Hawkins-Kennedy test, seven for the Neer sign, eight for the Jobe test and three for the drop-arm test. All except one study[46] had a QUADAS score of seven or higher. Two studies[14,48] had the score of 13 of 14 points. Out of these 11 studies investigating the physical examination, three studies[46,50,51] investigated the validity of the single parts of the rotator cuff pathology, and they also included other pathological definitions; massive tear[46] and subacromial bursitis.[50]

When targeting rotator cuff tendinopathy, partial- and full- thickness tears (incl. sub-acromial bursitis and massive tears) the sensitivity and specificity for the Hawkins-Kennedy test (Table 3) ranged from 58% to 94%, and 25% to 72%, respectively, while for the Neer sign (Table 4) the sensitivity ranged from 64% to 89%, and the specificity from 30% to 69%.

The diagnostic criteria for the Jobe test (Table 5) was not the same in all studies, six out of eight studies considered muscle weakness. Three studies[14,51,52] investigated weakness and pain together as one criteria. The study of Itoi et al.[53] focused solely on the results of the supraspinatus muscle and tendon, while all the other listed studies focused on the rotator cuff. The sensitivity of the Jobe test had a range from 44% to 86%, and the specificity ranged from 5% to 90% when targeting rotator cuff tendinopathy, partial- and full- thickness tears (incl. sub-acromial bursitis), while the Jobe test outcome for full-thickness tears shows a sensitivity range of 41% to 89%, and a specificity range of 50% to 83%.

The three studies[44,49,50] that investigated the drop-arm test (table 6) all stated the use of the same criteria, although the targeted pathology showed discrepancies. When focusing on rotator cuff tendinopathy, partial- and full- thickness tears (incl. sub-acromial bursitis), the sensitivity and specificity for the drop-arm test ranged from 8% to 27% and 88% to 97%, respectively. The sensitivity and specificity range for the drop-arm test in diagnosing partial- and full- thickness tears was 14% to 35%, and 75% to 88%, respectively.

Discussion

The main aim of this study was to review the available evidence for the physical tests recommended by the KNGF, and compare it with the available evidence for musculoskeletal ultrasound on diagnosing the rotator cuff pathology of SIS, with the purpose of providing an updated overview of the accuracy of these diagnostic tools for the clinical practice.

Musculoskeletal ultrasound findings

The ultrasound studies all reported a specificity over 80% for diagnosing full-thickness tears, and most studies reported sensitivity values over 74%, except the studies of Chang et al.[30] and Martín-Hervás et al.[34] These results suggests that MSU could be a good alternative to MRI or invasive techniques for ruling in or out full-thickness tears of the rotator cuff. The lowest result in the sensitivity was reported from Chang et al.[30], a study of high methodological quality (QUADAS score 11). The study stated dividing the participants retrospectively into two groups due to an alteration of the MSU operator. The first group had a less experienced operator (5 years) and the second had a more experienced operator (10 years). These results point to a significantly higher sensitivity with the more experienced operator. This is often the case when diagnosing with MSU.[20] Still most studies do not include information on the MSU operator's experience.

The ultrasound studies that investigated partial-thickness tears reported a wide range of outcome. Most of these studies reported a specificity over 80%, except for the study of Martín-Hervás et al.[34]. Five studies[29,33,35,37,43] had a sensitivity over 75%, while the six other studies[15,34,38–41] reported a sensitivity range from 13% to 68%. Martín-Hervás et al.[34] reported the lowest sensitivity outcome, although this result was only focused on the supraspinatus, and not on the rotator cuff as a group. The study of Martín-Hervás et al.[34] was also shown to be one of the studies with the lowest QUADAS result (7 of 14). The results from investigating partial-thickness tears suggests a more specific test, meaning MSU could be carried out to rule out this pathology when the diagnostic criteria is reported negative. The diagnostic criteria for partial-thickness tears are often similar to that of the full-thickness tears, and are interpreted as visual or non-visual (positive or negative). By those definitions, a positive criterion would actually constitute a full-thickness tear rather than a partial -thickness tear of the rotator cuff. The studies of Ziegler et al.[43], Yen et al.[42] and Iannotti et al.[33] reported results on both partial-thickness tears and full-thickness tears together. Investigating all rotator cuff tears justifies the use of same diagnostic criteria, because there is no differentiation of positive and negative results. The sensitivity and specificity from these studies ranged from 88% to 100%, and 82% to 90%, respectively. The studies[33,42,43] all showed a high methodological quality, with a QUADAS score range of eight to 10. The results from these studies suggests that MSU can be carried out to rule in and out all rotator cuff tears.

Martín-Hervás et al.[34] and Naredo et al.[29] were the only two studies investigating the accuracy of MSU on tendinopathy, and both studies focused solely on the supraspinatus tendon. Martín-Hervás et al.[34] reported a sensitivity of 67% and a specificity of 88%, while Naredo et al.[29] reported 93% sensitivity and

100% specificity. Further research is needed of MSU on the pathology of rotator cuff tendinopathy, although these results trend towards a higher specificity than sensitivity. The QUADAS score for Naredo et al.[29] was substantially higher (12 of 14) than that of Martín-Hervás et al.[34] (7 of 14). Both these studies were carried out at the university hospital of Madrid with only a few years difference (1999 and 2001), although none of the same authors were stated. Since both studies carried out a prospective study design, as well as different reference tests, the chance of replication of the sample size and participants was rejected. Once more the diagnostic criteria are stated the same for all rotator cuff pathologies in these studies.

Physical examination findings

The main findings for the physical tests are mostly focused on rotator cuff pathology as a general term. The studies of Park et al.[50] and Jia et al.[46] were among the only ones to differentiate between the targeted pathologies, as well as Holtby & Razmjou[51] on the Jobe test.

The sensitivity of the Hawkins-Kennedy test was over 75% in five[44,46,47,49,50] of the eight studies, and the specificity from all studies was below 72%. For the Neer sign, six of seven studies reported sensitivity over 75%, only Silva et al.[14] reported a lower sensitivity value. Silva et al.[14] included only 29 shoulders in their study, which is the second smallest study of this review. This could affect their outcome, although it had a QUADAS score of 13, which is the highest in this review. The differentiated pathologies, such as tendinopathy and partial-thickness tears had higher sensitivities from both Park et al.[50] and Jia et al.[46]. None of the studies reported a specificity over 75% for the Neer sign. This suggests a more sensitive tests than specific, which means that the Hawkins-Kennedy test and Neer sign would be better at ruling out the pathologies with a negative test.[28] Furthermore, a high sensitivity test could be combined with a high specificity test to produce a better opinion.[54] The studies of Park et al.[50] and Jia et al.[46] investigated these tests with remarkably similar results on all the targeted pathologies. Both of the studies also originated from Baltimore, U.S., one from “Johns Hopkins Bayview Medical Center”, while the other one from “The Johns Hopkins University”. Furthermore, the studies shared one author (E.G.M), which leads the author of this review to suspect a biased sample size. Although, the studies did not report any other connections. The author of this review can neither confirm nor exclude this suspicion. The reason for inclusion of both these studies in this review was the reported difference in reference tests performed.

The Jobe test results varied greatly, although sensitivities over 77% were reported from Itoi et al.[52] and Itoi et al.[53] who investigated with pain and weakness as separate diagnostic criteria solely for the supraspinatus pathology, and from Nanda et al.[49] who only viewed weakness as criteria. Michener et al.[48] and Park et al.[50] both with high QUADAS scores, reported specificity over 83%, although, with different criteria and pathology in their scopes. The KNGF[12] recommended that the Jobe test could be employed as a integrity test for the supraspinatus. There is neither heterogeneity in the criteria nor the targeted pathology for the Jobe test, and executing it should be carried out with caution. Furthermore, Nanda et al.[49] reported a specificity of only 5% for this test.

The drop-arm test results were relatively uniform, as well as the diagnostic criteria employed, the differentiating factor was the pathologies targeted. The specificity of all pathologies were reported being

over 75%, while sensitivities were all under 35%. Which suggests that this test could be better at ruling in the pathology when positive.

Likelihood ratios

Results reported in sensitivity and specificity or in 2 by 2 tables have a high statistical value, although, when reviewing evidence for the clinical practice, likelihood ratios could be of higher value, as they express the relationship between sensitivity and specificity.[28,55] Appendix IV offers an overview of the calculated positive and negative likelihood ratios for the included studies, as well as the mathematical formula employed.

A positive test result with a LR+ shift of 10 or over is significant for ruling in a pathology with a positive test, while a LR+ shift between 5 and 10 can be useful as post-test probability of a pathology. A LR- shift of 0.1 or below is considered a significant value, while a LR- shift of between 0.1 and 0.2 can be used in the same post-test context for the probability of the absence of pathology.[28,55]

The likelihood outcome for diagnosing full-thickness tears with MSU shows significant shifts in both positive and negative likelihood ratios in seven[15,30–32,35,37,43] of the 16 studies, as well as eight studies[29,34,36,39–42] showing either a significant LR+ or a significant LR-. These results tells that MSU can be carried out to rule full-thickness tears in as well as out. Ziegler et al.[43] had a calculated LR+ of 23.50 and LR- of 0.06 for partial-thickness tears, while Sonnabend et al.[38] showed a LR+ of 25.00, and Teefey et al.[41] 17.00 LR+. Furthermore, other studies presented LR+ shifts between 5 to 10, and LR- between 0.2 and 0.1, also pointing towards an accurate method of ruling out with a negative (non-visual criteria). The calculated LRs for tendinopathy also presented significance from Naredo et al.[29] Although, Martín-Hervás et al.[34] only had a LR+ shift of 5.58. For partial- and full- thickness tears together the studies of Yen et al.[42] and Ziegler et al.[43] showed significant LR- shifts, while LR+ shifts staying between 5 to 10 values.

None of the physical tests included in this review presented significant likelihood ratios. Although, Nanda et al.[49] showed a 0.12 shift in LR- on the Hawkins-Kennedy test, which means a negative test result could be executed in a sequence of tests, where the result of one test could have a pre-test impact on the next.[55] Three studies presented a near to neutral likelihood ratio, Silva et al.[14] on the Neer sign with a 0.97 LR+ to 1.07 LR-, as well as Nanda et al.[49] and Park et al.[50] on the Jobe test.

In relation to other reviews

To the knowledge of the author of this review, the only other review to investigate the accuracy for both physical examination and MSU on rotator cuff pathology is the review of Dinnes et al.[26]. Additionally it also investigated the accuracy of MRA and MRI. Compared to this review, more physical tests were included in the review of Dinnes et al.[26], and the Jobe test was not investigated. There was also a higher acceptability on reference tests included compared to this review, allowing both subacromial injections tests and arthrography to be utilized. There was also a lower threshold on the frequency used by the ultrasound studies, examining with frequencies under 7,5Mhz. Ottenheijm et al.[8] focused their review on updating the MSU part of Dinnes et al.[26], by extending the search from 2003 to 2012. It included

more specific pathologies such as calcifying tendonitis and subacromial bursitis. It concluded that MSU should be carried out to rule in or out full-thickness tears, as well as to rule in partial-thickness tears. Diagnosing tendinopathy with MSU were yet to be validated. This builds on this reviews outcome, although this reviews criteria raised the threshold for the included studies methodological quality.

In the review of Alqunaee et al.[19] only cohort studies were included, and the pathology was reported as three stages of SIS. It includes a meta-analysis of Neer test, Hawkins-Kennedy test, Jobe test and drop-arm test, as well as the lift-off test. It concludes that physical testing of SIS should be done in context in the overall patient assessment, and that the lift-off test has the highest value in ruling the pathology in.

Hegedus et al.[13] investigated a greater variety of shoulder pathologies compared to this review, and naturally also more physical tests. The review also included six studies that reported sensitivity only. The review of Hughes et al.[18] detected a typographical error in the review of Hegedus et al.[13], which stated that the specificity outcome of Itoi et al.[52] was 98%, although in the original study it was reported as 50%. Double-contrast arthrography was also accepted as a reference test. The main recommendations were that although there was no significant LRs, the Hawkins-Kennedy test could be carried out as a screening test for impingement, due to the high sensitivity, while the Jobe test could serve as a confirmatory test for rotator cuff tears, due to its high specificity. These recommendations are not all in line with the findings in this review, and could be accounted for due to the typographical error found by Hughes et al.[18].

Hughes et al.[18] reviewed 14 physical tests in diagnosing rotator cuff pathology, among these test were the Hawkins-Kennedy, Neer, Jobe and drop-arm test. The conclusions were that most test cannot accurately diagnose rotator cuff pathology, although a combination of tests can be carried out to create an opinion. This recommendation is also shared with Cools et al.[7], who presented an algorithm for clinical reasoning in early detection of shoulder pathology.

Limitations

Several limitations were identified in this review, as well as in the included studies. Five studies had to be excluded due to restricted access by the author, although, it is unknown in which way these results could have affected the outcome of this review. The review could have constructed a list of acceptable reference test, as well as the specific diagnostic criteria. The index tests were also not stated to the detail, as well as with which diagnostic criteria they were included. Another major limitation was that the review was carried out by a single author, which is reported as being biased as a mean of 8% of eligible studies are lost this way.[56] The experience of the author was limited as well, as only small lectures about research was the background for this review, and it being the first major research paper submitted.

The included studies have also shown critical aspects, the sample size being 50 or below in five[29,32,36,37,42] of the ultrasound studies and three[14,49,51] of the physical examination studies. The study designs were not clearly stated in three[35,37,39] of the ultrasound studies and four[44,47,49,50] of the physical examination studies, as well as two reported as case series studies[43,53]. None reported carrying out the study in a primary care setting, which means the results should be considered with care when applied in that setting. It could be speculated that the results would be different in primary care as in secondary

care, due to a lower prevalence of these patients, as the studies included in this review only examined patients that were suspected of having a rotator cuff pathology or who failed at improvement from conservative treatment. One ultrasound study[15] and three physical examination studies [45–47] did not clearly state which diagnostic criteria which were examined, although the physical examination studies referred to other studies when presenting the test, the ultrasound study simply stated using "established criteria". Another point of attention is the fact that most of the studies had a time delay between the index test and the reference test which was over what the author of this review considered adequate (within 6 weeks), and many studies did not even mention the time delay. This could allow for disease progression, and affect the outcome values. Only two physical examination studies[46,50] investigated the specific parts of rotator cuff pathology, the other physical examination studies targeted the pathology as a group. MRI was utilized as a reference test in two ultrasound studies[29,31] as well as in three physical examination studies[14,44,52], although it might not have enough accuracy to serve as a reference test for these pathologies.[26] The rest of the studies carried out open surgery or arthroscopy as reference tests.

Implications

This review adds all the available evidence for both musculoskeletal ultrasound as well as the physical tests recommended by the KNGF[12] together for comparison. It was built on requirements to filter out the studies not utilizing proper reference tests, as well as a higher minimum frequency for the ultrasound studies. Additional strong points are the variety in age of the participating patients, which implies that the results can be used on a wide range of age groups. The total sample size was substantial (4755 shoulders), which positively supports the outcome of this review.

General practitioners and physical therapists in the primary care should carry out the physical tests involved in this review with caution, although, a sequence of tests with high sensitivity and high specificity could give enough information to create an opinion. Combining a test with a high sensitivity with a test that has high specificity, can increase the chances of including or excluding pathology. These test mostly focused on the larger scope of rotator cuff pathology, and can therefore not diagnose anything more specific than rotator cuff pathology or subacromial impingement syndrome. MSU on the other hand has the ability to view the structures when applied by an experienced operator, and can give more information on the structure level of the ICF. MSU examination could be carried out to exclude full-thickness tears of the rotator cuff, as operative treatment is required in most of these cases, as well as to limit the number of patients with a lesser degree of rotator cuff pathology to undergo unnecessary invasive interventions.

Recommendations

There were only two trails[57,58] investigating MSU and physical examination together against a reference test. Therefore, this review recommends clinical trials to be carried out on the same patient group, performing the physical tests and the MSU blinded to each other. Comparing the outcome against a valid reference test for a more accurate comparison of the MSU and physical testing. Additionally, studies should be carried out to investigate the diagnostic criteria for partial-thickness tears, and how these differentiate from full-thickness tears, as well as criteria for rotator cuff tendinopathy.

The included studies are all executed from a secondary care setting, and further research on the accuracy of these methods should be carried out in the primary care setting, which could also include physical therapists with MSU education.
If this review is to be repeated in the future, it could be further expanded by executing a meta-analysis of the homogenous results to see the pooled outcome and to be able to make a clear statement about these tests.

Conclusion

The results show that MSU can include and exclude full-thickness tears, as well as exclude partial-thickness tears of the rotator cuff. The evidence for rotator cuff tendinopathy points towards a higher specificity than sensitivity, which could help in excluding the pathology.
The physical tests presented no significant likelihood ratios. The Hawkins-Kennedy test and Neer sign displayed high sensitivities, and the drop-arm test high specificity in diagnosing the general term of rotator cuff pathology. These tests combined could help to include or exclude the diagnosis of rotator cuff pathology. The evidence for the Jobe test was inconclusive in this review.

References

1. Luime JJ, Koes BW, Hendriksen IJM, Burdorf A, Verhagen AP, Miedema HS, et al. Prevalence and incidence of shoulder pain in the general population; a systematic review. Scand. J. Rheumatol. 2004;33(2):73–81.

2. Van der Windt DA, Koes BW, Boeke AJ, Devillé W, De Jong BA, Bouter LM. Shoulder disorders in general practice: prognostic indicators of outcome. Br. J. Gen. Pr. J. R. Coll. Gen. Pr. 1996 Sep;46(410):519–23.

3. Croft P, Pope D, Silman A. The clinical course of shoulder pain: prospective cohort study in primary care. Primary Care Rheumatology Society Shoulder Study Group. BMJ. 1996 Sep 7;313(7057):601–2.

4. Feleus A, Bierma-Zeinstra SMA, Miedema HS, Verhagen AP, Nauta AP, Burdorf A, et al. Prognostic indicators for non-recovery of non-traumatic complaints at arm, neck and shoulder in general practice--6 months follow-up. Rheumatol. Oxf. Engl. 2007 Jan;46(1):169–76.

5. Bala M, Bekkering T, Riemsmsa R, Harker J, Huygen F, Kleijnen J. “Epidemiology of chronic pain in the Netherlands”. Kleijnen Systematic Reviews Ltd – York – United Kingdom. Neth J Med. 2011;69:141–53.

6. Lewis JS. Rotator cuff tendinopathy/subacromial impingement syndrome: is it time for a new method of assessment? Br. J. Sports Med. 2009 Apr;43(4):259–64.

7. 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 Aug;42(8):628–35.

8. Ottenheijm RP, Jansen MJ, Staal JB, van den Bruel A, Weijers RE, de Bie RA, et al. Accuracy of diagnostic ultrasound in patients with suspected subacromial disorders: a systematic review and meta-analysis. Arch. Phys. Med. Rehabil. 2010 Oct;91(10):1616–25.

9. Gaitini D. Shoulder ultrasonography: performance and common findings. J. Clin. Imaging Sci. 2012;2:38.

10. Roe Y, Soberg HL, Bautz-Holter E, Ostensjo S. A systematic review of measures of shoulder pain and functioning using the International Classification of Functioning, Disability and Health (ICF). Bmc Musculoskelet. Disord. 2013 Feb 28;14(1):73.

11. WHO | International Classification of Functioning, Disability and Health (ICF) [Internet]. WHO. [cited 2013 May 29]. Available from: http://www.who.int/classifications/icf/en/

12. Jansen MJ, Brooijmans F, Geraets JJX., Lenssen A., Ottenheijm RP., Penning LI., et al. KNGF Evidence Statement Subacromiale klachten. Ned. Tijdschr. Voor Fysiother. Suppl. 2011;121(1).

13. Hegedus EJ, Goode A, Campbell S, Morin A, Tamaddoni M, Moorman CT 3rd, et al. Physical examination tests of the shoulder: a systematic review with meta-analysis of individual tests. Br. J. Sports Med. 2008 Feb;42(2):80–92; discussion 92.

14. Silva L, Andréu JL, Muñoz P, Pastrana M, Millán I, Sanz J, et al. Accuracy of physical examination in subacromial impingement syndrome. Rheumatol. Oxf. Engl. 2008 May;47(5):679–83.

15. Rutten MJCM, Spaargaren G-J, van Loon T, de Waal Malefijt MC, Kiemeney LALM, Jager GJ. Detection of rotator cuff tears: the value of MRI following ultrasound. Eur. Radiol. 2010 Feb;20(2):450–7.

16. Winters JC, Van der Windt DAWM, Spinnewijn WEM, De Jongh AC, Van der Heijden GJMG, Buis PAJ, et al. NHG-Standaard Schouderklachten - Tweede herziening. Huisarts Wet. 2008;51(11):555–65.

17. Seitz AL, McClure PW, Finucane S, Boardman ND 3rd, Michener LA. Mechanisms of rotator cuff tendinopathy: intrinsic, extrinsic, or both? Clin. Biomech. Bristol Avon. 2011 Jan;26(1):1–12.

18. Hughes PC, Taylor NF, Green RA. Most clinical tests cannot accurately diagnose rotator cuff pathology: a systematic review. Aust. J. Physiother. 2008;54(3):159–70.

19. Alqunaee M, Galvin R, Fahey T. Diagnostic accuracy of clinical tests for subacromial impingement syndrome: a systematic review and meta-analysis. Arch. Phys. Med. Rehabil. 2012 Feb;93(2):229–36.

20. Patil P, Dasgupta B. Role of diagnostic ultrasound in the assessment of musculoskeletal diseases. Ther. Adv. Musculoskelet. Dis. 2012 Oct;4(5):341–55.

21. Ozçakar L, Tok F, De Muynck M, Vanderstraeten G. Musculoskeletal ultrasonography in physical and rehabilitation medicine. J. Rehabil. Med. Off. J. Uems Eur. Board Phys. Rehabil. Med. 2012 Apr;44(4):310–8.

22. Karel YHJM, Scholten-Peeters WGM, Thoomes-de Graaf M, Duijn E, Ottenheijm RPG, van den Borne MPJ, et al. Current management and prognostic factors in physiotherapy practice for patients with shoulder pain: design of a prospective cohort study. Bmc Musculoskelet. Disord. 2013;14:62.

23. Middleton WD, Payne WT, Teefey SA, Hildebolt CF, Rubin DA, Yamaguchi K. Sonography and MRI of the shoulder: comparison of patient satisfaction. Ajr Am. J. Roentgenol. 2004 Nov;183(5):1449–52.

24. Girish G, Lobo LG, Jacobson JA, Morag Y, Miller B, Jamadar DA. Ultrasound of the shoulder: asymptomatic findings in men. Ajr Am. J. Roentgenol. 2011 Oct;197(4):W713–719.

25. Le Corroller T, Cohen M, Aswad R, Pauly V, Champsaur P. Sonography of the painful shoulder: role of the operator's experience. Skeletal Radiol. 2008 Nov;37(11):979–86.

26. Dinnes J, Loveman E, McIntyre L, Waugh N. The effectiveness of diagnostic tests for the assessment of shoulder pain due to soft tissue disorders: a systematic review. Heal. Technol. Assess. Winch. Engl. 2003;7(29):iii, 1–166.

27. Whiting P, Rutjes AWS, Reitsma JB, Bossuyt PMM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. Bmc Med. Res. Methodol. 2003 Nov 10;3:25.

28. Davidson M. The interpretation of diagnostic test: a primer for physiotherapists. Aust. J. Physiother. 2002;48(3):227–32.

29. Naredo AE, Aguado P, Padrón M, Bernad M, Uson J, Mayordomo L, et al. A comparative study of ultrasonography with magnetic resonance imaging in patients with painful shoulder. J. Clin. Rheumatol. Pr. Reports Rheum. Musculoskelet. Dis. 1999 Aug;5(4):184–92.

30. Chang C-Y, Wang S-F, Chiou H-J, Ma H-L, Sun Y-C, Wu H-D. Comparison of shoulder ultrasound and MR imaging in diagnosing full-thickness rotator cuff tears. Clin. Imaging. 2002 Feb;26(1):50–4.

31. Al-Shawi A, Badge R, Bunker T. The detection of full thickness rotator cuff tears using ultrasound. J. Bone Jt. Surg. - Br. Vol. 2008 Jul 1;90-B(7):889–92.

32. Frei R, Chládek P, Trc T, Kopecný Z, Kautzner J. Arthroscopic evaluation of ultrasonography and magnetic resonance imaging for diagnosis of rotator cuff tear. Ortop. Traumatol. Rehabil. 2008 Apr;10(2):111–4.

33. Iannotti JP, Ciccone J, Buss DD, Visotsky JL, Mascha E, Cotman K, et al. Accuracy of office-based ultrasonography of the shoulder for the diagnosis of rotator cuff tears. J. Bone Joint Surg. Am. 2005 Jun;87(6):1305–11.

34. Martín-Hervás C, Romero J, Navas-Acién A, Reboiras JJ, Munuera L. Ultrasonographic and magnetic resonance images of rotator cuff lesions compared with arthroscopy or open surgery findings. J. Should. Elb. Surg. Am. Should. Elb. Surg. Al. 2001 Oct;10(5):410–5.

35. Milosavljevic J, Elvin A, Rahme H. Ultrasonography of the rotator cuff: a comparison with arthroscopy in one-hundred-and-ninety consecutive cases. Acta Radiol. Stockh. Swed. 1987. 2005 Dec;46(8):858–65.

36. Paavolainen P, Ahovuo J. Ultrasonography and arthrography in the diagnosis of tears of the rotator cuff. J. Bone Joint Surg. Am. 1994 Mar;76(3):335–40.

37. Read JW, Perko M. Shoulder ultrasound: diagnostic accuracy for impingement syndrome, rotator cuff tear, and biceps tendon pathology. J. Should. Elb. Surg. Am. Should. Elb. Surg. Al. 1998 Jun;7(3):264–71.

38. Sonnabend DH, Hughes JS, Giuffre BM, Farrell R. The clinical role of shoulder ultrasound. Aust. N. Z. J. Surg. 1997 Sep;67(9):630–3.

39. Takagishi K, Makino K, Takahira N, Ikeda T, Tsuruno K, Itoman M. Ultrasonography for diagnosis of rotator cuff tear. Skeletal Radiol. 1996 Apr;25(3):221–4.

40. Teefey SA, Hasan SA, Middleton WD, Patel M, Wright RW, Yamaguchi K. Ultrasonography of the rotator cuff. A comparison of ultrasonographic and arthroscopic findings in one hundred consecutive cases. J. Bone Joint Surg. Am. 2000 Apr;82(4):498–504.

41. Teefey SA, Rubin DA, Middleton WD, Hildebolt CF, Leibold RA, Yamaguchi K. Detection and quantification of rotator cuff tears. Comparison of ultrasonographic, magnetic resonance imaging, and arthroscopic findings in seventy-one consecutive cases. J. Bone Joint Surg. Am. 2004 Apr;86-A(4):708–16.

42. Yen C-H, Chiou H-J, Chou Y-H, Hsu C-C, Wu J-J, Ma H-L, et al. Six surgery-correlated sonographic signs for rotator cuff tears: emphasis on partial-thickness tear. Clin. Imaging. 2004 Feb;28(1):69–76.

43. Ziegler DW. The use of in-office, orthopaedist-performed ultrasound of the shoulder to evaluate and manage rotator cuff disorders. J. Should. Elb. Surg. Am. Should. Elb. Surg. Al. 2004 Jun;13(3):291–7.

44. Caliş M, Akgün K, Birtane M, Karacan I, Caliş H, Tüzün F. Diagnostic values of clinical diagnostic tests in subacromial impingement syndrome. Ann. Rheum. Dis. 2000 Jan;59(1):44–7.

45. Fowler EM, Horsley IG, Rolf CG. Clinical and arthroscopic findings in recreationally active patients. Sports Med. Arthrosc. Rehabil. Ther. Technol. Smartt. 2010;2:2.

46. Jia X, Petersen SA, Khosravi AH, Almareddi V, Pannirselvam V, McFarland EG. Examination of the shoulder: the past, the present, and the future. J. Bone Joint Surg. Am. 2009 Nov;91 Suppl 6:10–8.

47. MacDonald PB, Clark P, Sutherland K. An analysis of the diagnostic accuracy of the Hawkins and Neer subacromial impingement signs. J. Should. Elb. Surg. Am. Should. Elb. Surg. Al. 2000 Aug;9(4):299–301.

48. Michener LA, Walsworth MK, Doukas WC, Murphy KP. Reliability and diagnostic accuracy of 5 physical examination tests and combination of tests for subacromial impingement. Arch. Phys. Med. Rehabil. 2009 Nov;90(11):1898–903.

49. Nanda R, Gupta S, Kanapathipillai P, Liow RYL, Rangan A. An assessment of the inter examiner reliability of clinical tests for subacromial impingement and rotator cuff integrity. Eur. J. Orthop. Surg. Traumatol. 2008 Jul 16;18(7):495–500.

50. Park HB, Yokota A, Gill HS, El Rassi G, McFarland EG. Diagnostic accuracy of clinical tests for the different degrees of subacromial impingement syndrome. J. Bone Joint Surg. Am. 2005 Jul;87(7):1446–55.

51. Holtby R, Razmjou H. Validity of the supraspinatus test as a single clinical test in diagnosing patients with rotator cuff pathology. J. Orthop. Sports Phys. Ther. 2004 Apr;34(4):194–200.

52. Itoi E, Kido T, Sano A, Urayama M, Sato K. Which is more useful, the “full can test” or the “empty can test,” in detecting the torn supraspinatus tendon? Am. J. Sports Med. 1999 Feb;27(1):65–8.

53. Itoi E, Minagawa H, Yamamoto N, Seki N, Abe H. Are pain location and physical examinations useful in locating a tear site of the rotator cuff? Am. J. Sports Med. 2006 Feb;34(2):256–64.

54. Lalkhen AG, McCluskey A. Clinical tests: sensitivity and specificity. Contin. Educ. Anaesth. Crit. Care Pain. 2008 Dec 1;8(6):221–3.

55. Grimes DA, Schulz KF. Refining clinical diagnosis with likelihood ratios. Lancet. 2005 Apr 23;365(9469):1500–5.

56. Edwards P, Clarke M, DiGuiseppi C, Pratap S, Roberts I, Wentz R. Identification of randomized controlled trials in systematic reviews: accuracy and reliability of screening records. Stat. Med. 2002 Jun 15;21(11):1635–40.

57. Ardic F, Kahraman Y, Kacar M, Kahraman MC, Findikoglu G, Yorgancioglu ZR. Shoulder impingement syndrome: relationships between clinical, functional, and radiologic findings. Am. J. Phys. Med. Rehabil. Assoc. Acad. Physiatr. 2006 Jan;85(1):53–60.

58. Naredo E, Aguado P, De Miguel E, Uson J, Mayordomo L, Gijon-Baños J, et al. Painful shoulder: comparison of physical examination and ultrasonographic findings. Ann. Rheum. Dis. 2002 Feb;61(2):132–6.

Appendices

Appendix I: The QUADAS tool

Appendix II: Overview table for ultrasound QUADAS scores

Appendix III: Overview table for physical examination QUADAS scores

Appendix IV: A 2x2 diagnostic “truth” table with the validity calculations

Appendix V: Ultrasound validity outcome with calculated likelihood ratios

Appendix VI: Physical examination validity outcome with calculated likelihood ratios

Appendix I: The QUADAS tool

The QUADAS tool

1. Was the spectrum of patients representative of the patients who will receive the test in practice?
2. Were selection criteria clearly described?
3. Is the reference standard likely to correctly classify the target condition?
4. Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?
5. Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?
6. Did patients receive the same reference standard regardless of the index test result?
7. Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?
8. Was the execution of the index test described in sufficient detail to permit replication of the test?
9. Was the execution of the reference standard described in sufficient detail to permit its replication?
10. Were the index test results interpreted without knowledge of the results of the reference standard?
11. Were the reference standard results interpreted without knowledge of the results of the index test?
12. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?
13. Were uninterpretable/ intermediate test results reported?
14. Were withdrawals from the study explained?

Extracted from: Penny Whiting et.al “The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews”, BMC Medical Research Methodology 2003, 3:25.

Appendix II: QUADAS scores for ultrasound studies

Study:	1. Patient spectrum represented ?	2. Criteria clearly stated ?	3. Reference test appropriate ?	4. Time period test and reference test ?	5. All patients had reference test ?	6. Same reference test for all patients?	7. Independent reference test ?	8. Index test well described ?	9. Reference test well described ?	10. US results without MRI/AS results ?	11. MRI/AS results without US results ?	12. Relevant clinical information ?	13. Un-interpretable results reported ?	14. Withdrawals explained ?	Total score:
Al-Shawi et al.[31]	**Y**	**N**	**Y**	**?**	**Y**	**N**	**Y**	**Y**	**Y**	**Y**	**N**	**?**	**?**	**Y**	**8**
Chang et al.[30]	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**N**	**Y**	**?**	**Y**	**11**
Frei et al.[32]	**Y**	**N**	**Y**	**?**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**?**	**Y**	**?**	**Y**	**10**
Ianotti et al.[33]	**Y**	**N**	**Y**	**?**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**?**	**N**	**?**	**?**	**8**
Martin-Hervas et al.[34]	**Y**	**N**	**Y**	**N**	**Y**	**Y**	**Y**	**Y**	**N**	**?**	**?**	**?**	**?**	**Y**	**7**
Milosavljevic et al.[35]	**Y**	**N**	**Y**	**N**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**?**	**Y**	**?**	**N**	**9**
Naredo et al.[29]	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**?**	**?**	**12**
Paavolainen & Ahovuo[36]	**Y**	**Y**	**Y**	**?**	**Y**	**Y**	**Y**	**Y**	**Y**	**?**	**?**	**Y**	**?**	**N**	**8**
Read & Perko[37]	**Y**	**N**	**Y**	**N**	**Y**	**Y**	**Y**	**Y**	**N**	**Y**	**N**	**?**	**?**	**N**	**7**
Rutten et al.[15]	**Y**	**Y**	**Y**	**N**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**N**	**Y**	**?**	**Y**	**11**
Sonnabend et al.[38]	**Y**	**N**	**Y**	**N**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**?**	**Y**	**?**	**N**	**9**
Takagishi et al.[39]	**Y**	**N**	**Y**	**?**	**Y**	**Y**	**Y**	**Y**	**Y**	**N**	**Y**	**Y**	**?**	**N**	**9**
Teefey et al.[40]	**Y**	**N**	**Y**	**N**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**?**	**Y**	**?**	**N**	**9**
Teefey et al.[41]	**Y**	**Y**	**Y**	**N**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**N**	**Y**	**?**	**Y**	**11**
Yen et al.[42]	**Y**	**N**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**?**	**Y**	**?**	**N**	**10**
Ziegler et al.[43]	**Y**	**N**	**Y**	**?**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**N**	**Y**	**?**	**Y**	**10**

Abbreviations: Y – Yes; N – No; ? – Unclear.

Appendix III: QUADAS scores for physical examination studies

Study:	1. Patient spectrum represented ?	2. Criteria clearly stated ?	3. Reference test appropriate ?	4. Time period test and reference test ?	5. All patients had reference test ?	6. Same reference test for all patients?	7. Independent reference test ?	8. Index test well described ?	9. Reference test well described ?	10. Index test results without MRI/AS results ?	11. MRI/AS results without index test results ?	12. Relevant clinical information ?	13. Un-interpretable results reported ?	14. Withdrawals explained ?	Total score:
Calis et al.[44]	**Y**	**Y**	**Y**	**?**	**Y**	**Y**	**Y**	**N**	**N**	**?**	**?**	**Y**	**?**	**?**	**7**
Fowler et al.[45]	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**N**	**N**	**Y**	**N**	**Y**	**?**	**Y**	**10**
Holtby & Razmjou[51]	**Y**	**Y**	**Y**	**?**	**Y**	**N**	**Y**	**N**	**Y**	**Y**	**Y**	**Y**	**?**	**N**	**9**
Itoi et al.[52]	**Y**	**N**	**Y**	**?**	**Y**	**Y**	**Y**	**N**	**Y**	**Y**	**?**	**Y**	**?**	**N**	**8**
Itoi et al.[53]	**Y**	**N**	**Y**	**?**	**Y**	**Y**	**Y**	**Y**	**N**	**?**	**?**	**Y**	**?**	**?**	**7**
Jia et al.[46]	**Y**	**N**	**Y**	**?**	**Y**	**N**	**Y**	**N**	**N**	**Y**	**?**	**Y**	**N**	**N**	**6**
MacDonald et al.[47]	**Y**	**N**	**Y**	**?**	**Y**	**Y**	**Y**	**Y**	**N**	**?**	**N**	**Y**	**?**	**?**	**7**
Michener et al.[48]	**Y**	**Y**	**Y**	**N**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**13**
Nanda et al.[49]	**Y**	**Y**	**Y**	**N**	**Y**	**N**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**?**	**?**	**10**
Park et al.[50]	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**N**	**Y**	**?**	**?**	**11**
Silva et al.[14]	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**?**	**Y**	**13**
Abbreviations: Y – Yes; N – No; ? – Unclear.															

Appendix IV: A 2x2 diagnostic "Truth" table with the validity calculations

Two by two table.

	Pathology *present*	Pathology *absent*
Diagnostic test *positive*	True positive = **A**	False positive = **B**
Diagnostic test *negative*	False negative = **C**	True negative = **D**

Sensitivity = **A/(A+C)**
Specificity = **D/(B+D)**
Positive predictive value (PPV) = **A/(A+B)**
Negative predictive value (NPV) = **D/(C+D)**
Likelihood ratio (test positive) = **sensitivity/(1-specificity)**
Likelihood ratio (test negative) = **(1-sensitivity)/specificity**

Validity calculations

Extracted from: Megan Davidson, "The interpretation of diagnostic tests: A primer for physiotherapists", Australian Journal of Physiotherapy 48: 227-233, 2002.

Appendix V: Ultrasound validity outcome with calculated likelihood ratios

Overview of **ultrasound** validity outcome

Study:	Diagnostic criteria§:	Target pathology*:	Se:	Sp:	LR+:	LR-:
Al-Shawi et al.[31]	A, B and C	FTT	96	95	**19.20**	**0.04**
Chang et al.[30]	A, C, D and E	FTT (group 1)	52	92	6.50	0.52
		FTT (group 2)	92	100	?	**0.08**
Frei et al.[32]	A and C	FTT	100	90	**10.00**	?
Iannotti et al.[33]	C and D	PTT	70	89	6.36	0.34
	A, C and D	PTT and FTT	88	82	4.89	0.15
Martin-Hervas et al.[34]	A, B, C, D, F and G	T	67	88	5.58	0.38
		PTT	13	68	0.41	1.28
		FTT	58	100	?	0.42
Milosavljevic et al.[35]	C	PTT	80	98	**40.0**	0.20
	A, C and D	FTT	100	91	**11.11**	?
Naredo et al.[29]	A, B, C, D, and E	T	93	100	?	**0.07**
		PTT	86	91	9.56	0.15
		FTT	89	100	?	0.11
Paavolainen & Ahovuo[36]	A and D	FTT	74	95	**14.80**	0.27
Read & Perko[37]	A, B, C and D	PTT	46	97	**15.33**	0.56
		FTT	100	97	**33.33**	?
Rutten et al.[15]	"Established	PTT	89	80	4.45	0.14
	Criteria"	FTT	95	93	**13.57**	**0.05**
Sonnabend et al.[38]	A, C and F	PTT	25	99	**25.00**	0.76
		FTT	84	92	**10.50**	0.17
Takagishi et al.[39]	C, D and E	PTT	50	90	5.00	0.56
		FTT	76	100	?	0.24
Teefey et al.[40]	A, B, C and E	PTT	67	85	4.47	0.39
		FTT	100	85	6.67	?
Teefey et al.[41]	A, C and E	PTT	68	96	**17.00**	0.33
		FTT	98	80	4.90	**0.03**
Yen et al.[42]	A, B, C, D and E	PTT and FTT	95	90	9.50	**0.06**
Ziegler et al.[43]	A, B, C, D and E	PTT	94	96	**23.50**	**0.06**
		FTT	96	94	**16.00**	**0.04**
		PTT and FTT	100	86	7.14	?

Abbreviations: ? – Unable to calculate; Se – Sensitivity; Sp – Specificity; LR+ – Positive likelihood ratio; LR- – Negative likelihood ratio.
*Targeted pathology: T – Tendinopathy; PTT – Partial-thickness tear; FTT – Full-thickness tear.
§Diagnostic criteria: A – non-visualisation of rotator cuff; B – loss of convexity; C – focal discontinuity; D – focal thinning; E – echogenic foci; F – echogenic band; G – presence of bursal fluid.

Appendix VI: Physical examination validity outcome with calculated likelihood ratios

Overview of the **Hawkins-Kennedy test** validity outcome

Study:	Target pathology*:	Diagnostic criteria:	Se:	Sp:	LR+:	LR-:
Calis et al.[44]	T, PTT and FTT	Pain	92	25	1.23	0.32
Fowler et al.[45]	T, PTT and FTT	NR	58	72	2.07	0.58
Jia et al.[46]	T	NR	76	45	1.38	0.53
	PTT	NR	75	44	1.34	0.57
	FTT	NR	69	48	1.33	0.65
	T, PTT and FTT (incl. MT)	NR	71	42	1.22	0.69
Macdonald et al.[47]	T, PTT and FTT	NR	88	43	1.54	0.28
Michener et al.[48]	T, PTT and FTT	Pain	63	62	1.66	0.60
Nanda et al.[49]	T, PTT and FTT	Pain	94	50	1.88	0.12
Park et al.[50]	T (incl. SAB)	Pain	76	45	1.38	0.53
	PTT	Pain	75	44	1.34	0.57
	FTT	Pain	69	48	1.33	0.65
	T, PTT and FTT (incl. SAB)	Pain	72	66	2.12	0.42
Silva et al.[14]	T, PTT and FTT	Pain	74	40	1.23	0.65

Abbreviations: NR – Not reported; Se – Sensitivity; Sp – Specificity; LR+ – Positive likelihood ratio; LR- – Negative likelihood ratio.
*Targeted pathology: T – Tendinopathy; PTT – Partial-thickness tear; FTT – Full-thickness tear; MT – Massive tear (<5cm); SAB - Subacromial bursitis.

Overview of the **Neer test** validity outcome

Study:	Target pathology*:	Diagnostic criteria:	Se:	Sp:	LR+:	LR-:
Calis et al.[44]	T, PTT and FTT	Pain	89	31	1.29	0.36
Jia et al.[46]	T	NR	86	49	1.69	0.29
	PTT	NR	75	48	1.44	0.52
	FTT	NR	59	48	1.14	0.85
	T, PTT and FTT (incl. MT)	NR	64	43	1.12	0.84
Macdonald et al.[47]	T, PTT and FTT	NR	83	51	1.69	0.33
Michener et al.[48]	T, PTT and FTT	Pain	81	54	1.76	0.35
Nanda et al.[49]	T, PTT and FTT	Pain	84	50	1.68	0.32
Park et al.[50]	T (incl. SAB)	Pain	84	49	1.65	0.33
	PTT	Pain	75	48	1.44	0.52
	FTT	Pain	59	47	1.11	0.87
	T, PTT and FTT (incl. SAB)	Pain	68	69	2.19	0.46
Silva et al.[14]	T, PTT and FTT	Pain	68	30	0.97	1.07

Abbreviations: NR – Not reported; Se – Sensitivity; Sp – Specificity; LR+ – Positive likelihood ratio; LR- – Negative likelihood ratio.
*Targeted pathology: T – Tendinopathy; PTT – Partial-thickness tear; FTT – Full-thickness tear; MT – Massive tear (<5cm); SAB - Subacromial bursitis.

Overview of the **Jobe test** validity outcome

Study:	Target pathology*:	Diagnostic criteria:	Se:	Sp:	LR+:	LR-:
Fowler et al.[45]	T, PTT and FTT	NR	67	41	1.14	0.80
Holtby & Razmjou[51]	T and PTT	Pain	62	54	1.35	0.70
	FTT	Weakness +/- Pain	41	70	1.37	0.84
Itoi et al.[52]	FTT	Pain	63	55	1.4	0.67
		Weakness	77	68	2.41	0.34
		P/W/Both	89	50	1.78	0.22
Itoi et al.[53]	T, PTT and FTT (SS)	Pain	78	40	1.3	0.55
		Weakness	87	43	1.53	0.30
Michener et al.[48]	T, PTT and FTT	Weakness	50	87	3.85	0.57
Nanda et al.[49]	PTT and FTT	Weakness	86	5	0.90	2.8
Park et al.[50]	T (incl. SAB)	Pain	25	67	0.76	1.12
	PTT	Pain	32	68	1.00	1.00
	FTT	Pain	53	83	3.12	0.57
	T, PTT and FTT (incl. SAB)	Pain	44	90	4.40	0.62
Silva et al.[14]	T, PTT and FTT	Pain and weakness	74	30	1.06	0.87

Abbreviations: NR – Not reported; Se – Sensitivity; Sp – Specificity; LR+ – Positive likelihood ratio; LR- – Negative likelihood ratio.
*Targeted pathology: T – Tendinopathy; PTT – Partial-thickness tear; FTT – Full-thickness tear; SAB - Subacromial bursitis.

Overview of the **drop-arm test** validity outcome

Study:	Target pathology*:	Diagnostic criteria:	Se:	Sp:	LR+:	LR-:
Calis et al.[44]	T, PTT and FTT	Arm drops and/or severe pain	8	97	2.67	0.95
Nanda et al.[49]	PTT and FTT	Arm drops and/or severe pain	31	75	1.24	0.92
Park et al.[50]	T (incl. SAB)	Arm drops and/or	14	77	0.61	1.12
	PTT	severe pain	14	78	0.64	1.10
	FTT		35	88	2.92	0.74
	T, PTT and FTT (incl. SAB)		27	88	2.25	0.83

Abbreviations: ? – Unclear; NR – Not reported; Se – Sensitivity; Sp – Specificity; LR+ – Positive likelihood ratio; LR- – Negative likelihood ratio.
*Targeted pathology: T – Tendinopathy; PTT – Partial-thickness tear; FTT – Full-thickness tear; ; SAB - Subacromial bursitis.

Appendix VII: Approval of project plan

Fontys Paramedische Hogeschool

B4 Assessment form project plan

Name: Axel Hedman Student no:

Date: 11/3/2013

Title: a comparative validity study of physical testing and MSU -----

General

- The project plan is according to format yes / no
- Spelling and language are correct yes / no

Problem description and problem definition (introduction)

- The problem description is sufficiently clearly formulated yes / no
- The problem description reflects social and paramedical relevance yes / no
- A concrete and relevant research question (or questions) can be formulated based on the problem definition, including possible sub questions yes / no

Objective

The objective is:

- Sufficiently clearly and concretely formulated yes / no
- Relevant for a selected target group within the (paramedical) professional practice yes / no
- Practically feasible yes / no
- Achievable within the set time yes / no

Project product

The project product:

- Is in line with the problem definition, research question and objective yes / no
- Is usable for the selected target group yes / no
- Is in line with the client's wishes yes / no
- The product requirements are accurately described yes / no

Activities/method

Sufficient insight is given into the type of activities and types of sources for the performance of the research and the realization of the product yes / no

Time schedule

- The time schedule gives a global phasing and time investment for the project as a whole and for the coming weeks an increasingly detailed schedule yes / no
- Important moments are recorded in the table (typographically noticeable) (e.g. contact moments, handing-in moments) yes / no
- The time schedule gives a global task division of the planned activities yes / no

Fontys Paramedische Hogeschool

Estimated costs

Clear insight is given in:

- The costs to be expected concerning money and hours — yes / no
- The division of these costs (project leader, student, programme) — yes / no

Literature

- Used and planned literature is specific and mentioned to a sufficient extent — yes / no
- Relevant and recent literature is referred to — yes / no
- Literature references, in the text and in the literature list, are made according to the Writer's Guide (Wouters 2012) — yes / no

Comments: Dear Axel. Congratulations. Only comments we can make are:
1) english is not always perfect
2) validity measurements: OK, best prevent that this will be too much (literature quantity) If it will be too much you'll have to make a choice.

All points under B3.1 up to and including B3.8 must be answered with a 'yes' in order to receive a GO for the project. The supervisor discusses with the student which points need adjustment.

GENERAL:	**GO**	**NO GO**

Name assessor: MARC SCHMITZ

Date + Signature: 11/3/2013