Diagnostic ultrasound in patients with shoulder complaints: An inter-examiner agreement study among Dutch physical therapists.

Duijn, E.A.H.D.[1,2], Pouliart, N.[3], Verhagen, A.P.[1,4], Karel, Y.H.J.M. [1], Thoomes-de Graaf, M.[1], Koes, B.W.[1,6], Scholten-Peeters, G.G.M.[5]

[1] Department of General Practice, Erasmus Medical Center, Rotterdam, The Netherlands

[2] Department of Basic (bio) Medical Sciences (BMWE), Vrije Universiteit Brussel, Brussels, Belgium

[3] Department of Orthopaedics and Traumatology, Universitair Ziekenhuis Brussel (UZ Brussel) Laarbeeklaan 101, B-1090 Brussel, Belgium

[4] Discipline of Physiotherapy, Graduate School of Health, University of Technology Sydney, Sydney, Australia.

[5] Department of Human Movement Sciences; Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam Movement Sciences, The Netherlands

6 Center for Muscle and Joint Health, University of Southern Denmark, Odense, Denmark

Corresponding author:

Edwin Duijn

Department of Bio-Medical Sciences (BMWE), Vrije Universiteit Brussel

e-mail: edwin.duijn@vub.be

Abstract

Study design A cross-sectional inter-examiner agreement and reliability study among physical therapists in primary care.

Background Musculoskeletal ultrasound (MSU) is frequently used by physical therapists to improve specific diagnosis in patients with shoulder pain, especially for the diagnosis rotator cuff tendinopathy (RCT) including tears.

Objectives To estimate the inter-examiner agreement and reliability in physical therapists using MSU for patients with shoulder pain.

Methods Physical therapists performed diagnostic MSU in 62 patients with shoulder pain. Both physical therapists were blinded to each other's results and patients were not informed about the test results. We calculated the overall inter-examiner agreement, specific positive and negative inter-examiner agreement, and inter-examiner reliability (Cohen's Kappa's).

Results Overall agreement for detecting RC ruptures ranged from 61.7% to 85.5% and from 43.9% to 91.4% for specific positive agreement. The specific negative agreement was lower with values ranging from 44.4% to 79.1% for RC ruptures. Overall agreement for other pathology than ruptures related to SAPS, ranged from 72.6% to 93.6% and from 77.3% to 96% for specific positive agreement. The specific negative agreement was lower with values ranging from 44.4% to 79.1% for RC ruptures and 52.5%–83.3% for other pathology than ruptures related to SAPS. Reliability values varied from substantial for any thickness ruptures to moderate for partial thickness ruptures and fair for full thickness tears. Moreover, reliability was fair for cuff tendinopathy. The reliability for AC arthritis and no pathology found was fair and moderate. There was substantial agreement for the calcifying tendinopathy.

Conclusions Physical therapists using MSU agree on the diagnosis of cuff tendinopathy and on the presence of RCT in primary care but agree less on the absence of pathology.

Keywords: diagnostic ultrasound, shoulder, subacromial pain syndrome, reliability, physical therapy, primary care, rehabilitation

Background

Shoulder pain is the second most reported musculoskeletal complaint (Greving et al., 2012). A common diagnosis for shoulder pain is subacromial pain syndrome (SAPS) (Diercks et al., 2014, Karel et al 2017). The clinical diagnosis is mainly based on history taking and physical tests (Hegedus et al., 2008, Michener et al 2009). The term SAPS include pathologies such as: bursitis, tendinosis calcarean, supra-spinatus tendinopathy, tear(s) of the rotator cuff, biceps tendinitis and tendon cuff degeneration (Diercks et al 2014) which can be observed with ultrasound (Singh 2012). We have to realize that shoulder pain cannot always be explained by pathologies in anatomical shoulder structures (Noten et al, 2017).

Recently, Musculoskeletal Ultrasound (MSU) is also considered as a useful diagnostic tool for physicians in detecting rotator cuff disorders and long head of the biceps tendon pathology (Belanger et al 2019, Nazarian et al., 2013; Ottenheijm et al., 2010; Roy et al 2015, Rutten, Jager, & Blickman, 2006). Reported advantages of MSU are: portability, non-invasive, cheap, lack of contraindications and quick to perform (Nazarian et al., 2013; Rutten et al., 2006). Traditionally, MSU is performed by physicians (e.g. radiologists, orthopaedic surgeons and rheumatologists) who are educated in physical medicine and rehabilitation (Chen, Lin, Hsu, Chen, & Kang, 2011; Silva et al., 2008). The increasing technical developments, increased experience of operators and protocol driven approaches have improved the reliability and accuracy in finding rotator cuff pathology, in the last years (Okoroha, Fidai, Tramer, Davis, & Kolowich, 2018; Rutten et al., 2006; Smith, Back, Toms, & Hing, 2011).

Nowadays, MSU is used more and more by physical therapists (PTs) to improve their specific diagnosis in patients with shoulder paints (Karel et al 2017, Scholten-Peeters, Franken, Beumer, & Verhagen, 2014). However, the results of a survey showed that orthopaedic surgeons and radiologists show low trust in diagnostic MSU knowledge and skills of physical therapists and general practitioners in primary care (Scholten-Peeters et al., 2014). Primary care patients, however appreciate the use of diagnostic ultrasound performed by physiotherapists, to help them better understand their shoulder pain (Lumsden, Lucas-Garner, Sutherland, & Dodenhoff, 2018).

In The Netherlands 1 out of 6 physical therapy practices in primary care is now using MSU as a diagnostic tool for patients with shoulder pain and for determining the choice of physical therapy treatment (Kooijman et al 2020). Despite the increased use of MSU by PTs, there is a

lack of studies on reliability (and diagnostic accuracy) of PTs using diagnostic MSU in order to detect rotator cuff disorders and for determining the choice of physical therapy treatment.

One study is available showing an excellent intra-rater-reliability of MSU when performed by an experienced PT and a high intra- and interrater-reliability when performed by inexperienced PT in healthy subjects, when measuring the acromion-greater tuberosity distance (Kumar, Bradley, & Swinkels, 2010; Kumar et al., 2011). Another study assessed the interobserver reliability of MSU between PTs and radiologists and found an overall fair agreement and a substantial agreement for full thickness tears (Thoomes-de Graaf et al., 2014)

No studies assessed the agreement and reliability of diagnostic MSU among physical therapists in patients with shoulder pain in order to detect rotator cuff disorders in routine primary care. Therefore, the aim of this study is to assess the inter-examiner agreement and reliability in physical therapists using MSU as a diagnostic tool (detecting rotator cuff disorders) for patients with shoulder complaints in primary care.

Methods

Study design

A cross-sectional inter-examiner agreement and reliability study. Agreement explores how outcomes of different examiners agree and is expressed in terms of observed agreement and proportion of specific agreement. Reliability is escribed as how patients can be distinguished from each other, despite measurement errors (de Vet et al 2013). The Medical Ethical Committee of the Erasmus University approved this study (number mec-2011-414). In the absence of standards in reporting agreement studies in the medical field, we used a formal guideline to report, named: 'Guidelines for reporting reliability and agreement studies (GRRAS) (Kottner et al., 2011).

Participants:

Over a period of 12 months, consecutive patients with shoulder pain were recruited from different physical therapy practices in the Netherlands. Inclusion criteria were: age over 18 years and adequate understanding of the Dutch language. All patients met the test cluster for SAPS described by Michener (Michener et al 2009). Patients with pathologies such as cancer, infections or fractures were excluded. Postoperative patients and patients who had

121 received a diagnostic imaging of the shoulder in the past 3 months were also excluded from
122 this study as this could affect blinding. The included patients had not previously visited a PT
123 who participated in this study. All patients signed an informed consent prior to the
124 ultrasound examinations of their shoulder.
125
126 *Examiners*
127 All MSU examinations were performed by four PTs (3 male and 1 female) more than 5 years
128 of experience in primary care and at least 2 year MSU experience (mean 3.75 years/SD 0.47)
129 evaluating more than 150 diagnostic ultrasound scans of the shoulder in primary care
130 (Mullaney 2019). trained in ultrasound of the shoulder (Table 1) performed the MSU
131 examinations. All four PTs had a certificate for their 'basic MSU skills' and 'MSU of the
132 shoulder masterclass'. In addition, the participating PTs were all holding a MSc-degree in
133 manual therapy.
134 In addition, all 4 PTs attended a 6 h-training meeting by a musculoskeletal ultrasound expert
135 about the scanning protocol of the shoulder from the European Society of Musculoskeletal
136 Radiology (Ian Beggs & Franz Kainberger, 2016) and discussed the relevant anatomy,
137 pathology, scanning technique and pitfalls.
138 The scanning protocol consisted of 9 structures to examine in a standardized sequence (1)
139 Long Head Biceps (LHB) (2) the Subscapularis (SSC) (3) anterior structures and Coraco-
140 Acromial Ligament (CAL) (4) the SupraSpinatus (SSP in crass position) (5) the SupraSpinatus
141 and Rotator Interval (SSP modified crass position and RI) (6) Subacromial Impingement Test
142 (SIT) and (7) Infraspinatus tendon and Teres minor Tendon (TmT) (8) posterior structures
143 glenohumeral joint and (9) the Acromion Clavicular joint (AC). Each PT used their own high-
144 end MSU equipment which they used in daily practice, either a PHILIPS ClearVue 550 (probe:
145 L12-4), a PHILIPS CX30 (probe: L12-4), or a Philips CX50 using a L12-3 broadband linear probe
146 with active array technology. Each transducer had a minimum frequency of 7.5 MHz and
147 appropriate software (beamforming technology) was available.
148
149 *Procedure*
150 Patients with the clinical diagnosis of SAPS were recruited for the study in the three
151 physiotherapy practices in primary care of the four participating PTs. The colleague PT
152 provided a written 'physical therapy diagnosis' in terms of the International Classification of

Functioning (ICF) Prior to MSU examination. The PTs performing the MSU examinations and the patients were not blinded for the written physical therapy diagnosis. When patients agreed to participate, they were invited for an MSU assessment in the recruiting practice by one of the PTs (PT1), followed 300 min later by the second MSU examination by one of the other PTs (PT2, 3 or 4). Each MSU examination followed the complete scanning protocol and took about 10-15 minutes. Both examinaitons were done on the same day to avoid progression bias. Each PT completed his/her ows "scan finding form" directly after ultrasound examination and the examining PTs were blinded to each other's diagnostic MSU results. Patients were not informed about the results between examinaitons, so that they were not able to influence the second MSU assessor.

Outcomes

MSU diagnoses were standardized in terms of different diagnostic outcome categories (Singh 2012): (1) Tendinopathy of the rotator cuff (RC) and/or biceps, (2) Calcification of the rotator cuff, (3) Full Thickness Tear (FTR) of the RC and/or biceps, (4) Partial Thickness Tear (PTR) of the RC and/or biceps, (5) Arthritis of the acromio-clavicular joint and (6) "No pathology found". Option (6) was only chosen when all steps of the scanning protocol were technically normal scanned and pathology was absent. The PT assessors were allowed to choose more than one outcome option.

Sample size

A sample size calculation was performed by using an online calculator (http://wnarifin.github.io). Based on a minimal acceptable kappa of 0.3, an expected kappa of 0.7, a proportion of outcome of 0.5, alpha of 0.05, beta of 0.8 and an expected drop-out rate of 10%, we needed to include at least 50 patients in this inter-examiner study.

Statistical analysis

Descriptive analyses were conducted to describe the prevalence of positive findings and the frequencies of particular diagnostic outcome categories for each of the 9 structures. For statistical analysis, the outcome categories FTR and PTR were also grouped together as any thickness rupture. Agreement is calculated by percentage agreement (AO), Specific positive-

(SPA) and Specific negative agreement (SNA). The specific positive and negative agreement is calculated according to de Vet et al (2013) and de Vet et al (2018).

Reliability is presented by A Cohen's κappa-value (k) with 95% Confidence Interval (CI) for the outcomes. The Cohen's κappa value (95% CI) is an agreement measure that corrects for chance and was interpreted in accordance with Landis and Koch (1977): <0.00: poor agreement, 0.00-0.20 slight agreement, 0.21-0.40, fair agreement; 0.41-0.60, moderate agreement; 0.61-0.80, substantial agreement; 0.81-1.00, almost perfect agreement. A kappa of ≥ 0.7 was considered acceptable (Landis and Koch 1977).

The prevalence index (PI), bias index (BI) was calculated, in order to evaluate whether kappa was influenced by high prevalence of positive or negative decisions, or by systematic bias between examiners (de Vet et al., 2013). PI reflects the absolute difference between the proportion of agreement on positive indications as compared to that of negative indications. PI ranges between 0 and 1, and is high when the prevalence of concordant positive (or negative) indications is high, chance agreement is consequently also high, and kappa is reduced accordingly (Feinstein et al., 1990). BI provides a quantification of the extent to which raters disagree on the proportions of positive (or negative) indications. BI also ranges between 0 and 1, and is high when the absolute difference between the discordant indications is high, chance agreement is consequently low, and kappa is inflated accordingly (Feinstein and Cicchetti 1990). BI provides qualification of the extent to which raters disagree on the proportions of positive (or negative) indications. BI also ranges between 0 and 1, and is high when the absolute difference between the discordant indications is high, change agreement is consequently low, and kappa is inflated accordingly (Feinstein 1990). The hsls.pitt.edu website was used to calculate BI and PI (Sim and Wright 2005). The statistical analysis were performed using SPSS version 25.0 for Windows and praph-pad software (http://wwwgraphpad.com).

Results

We finally included in total 62 patients, with a mean age of 54.4 years (SD 15.4), of which 36 was female and 26 was male. All patients had unilateral shoulder pain for more than 6 weeks. In 40 of the 62 cases the right shoulder was affected. Table 2 describes the patient characteristics.

Agreement for detecting rotator cuff tears

In all cases were ruptures seen by the MSU-assessor the supraspinatus tendon was involved. The overall agreement for detecting cuff ruptures was high and ranged from 61.7% for partial thickness cuff ruptures to 85.5% for any thickness ruptures (partial and full thickness cuff tears). The specific positive agreement was also high and ranged from 87.5% for partial thickness cuff ruptures to 91.4% for full thickness cuff ruptures. The specific negative agreement ranged from 44.4% for full thickness cuff ruptures to 79.1% for any thickness cuff ruptures.

Reliability for detecting rotator cuff tears

The kappa value was substantial (0.68) for any thickness ruptures, slight for partial thickness cuff tears fair (0.15) and for full thickness tears slight (0.35) (Table 1a).

Table 1a: Diagnostic category (n=62) for detecting rotator cuff ruptures

		PTR	FTR	ATR
Frequency	PT1	14	9	19
	PT2, 3 or 4	9	9	24
	Both	9	6	16
OA		61.7	83.9	85.5
SPA		43.9	91.4	88.8
SNA		0.78	44.4	79.1
PI		0.35	0.71	0.31
BI		0.08	0.00	0.08
Cohen's kappa		0.15 (0.02-0.3)	0.35 (0.03-0.69)	0.68 (0.49-0.87)

Prevalence, Cohen's kappa, overall kappa, percentage (%) of observed agreement (OA), % Specific Positive Agreement (SPA), % Specific Negative Agreement (SNA), Prevalence Index (PI), Bias Index (BI) of Full Thickness Rupture (FTR), Partial Thickness Rupture (PTR), Any Thickness Rupture (ATR), Physical Therapist (PT1), Physical Therapist 2–3 or 4 (PT 2–3 or 4)

Prevalence index and bias index for detecting rotator cuff ruptures

The prevalence index for detecting cuff ruptures was high and ranged from 0.31 for any thickness ruptures to 0.71 for full thickness ruptures. The bias index for detecting cuff ruptures was low and ranged from 0.00 for full thickness ruptures to 0.13 for partial thickness ruptures

Agreement for detecting other shoulder pathology

The overall agreement for detecting pathology other than rotator cuff ruptures was high and ranged from 72.6% for a cuff tendinopathy to 93.6% for calcifying tendinopathy. The specific positive agreement was also high and ranged from 77.3% for a cuff tendinopathy to 96% for calcifying tendinopathy. The specific negative agreement ranged from 52.6% for 'no details (pathology) found' to 89,8% for ACJ arthritis.

Reliability for detecting other shoulder pathology

The kappa value was moderate for cuff tendinopathy (0.43), ACJ arthritis (0.54) and no pathology found (0.44). There was substantial agreement for calcifying tendinopathy (0.80) (Table 1b).

Prevalence index and bias index for detecting other shoulder pathology

The prevalence index ranged from 0.21 for cuff tendinopathy to 0.69 for no pathology found. The bias index was low and ranged from 0.03 for any ACJ arthritis to 0.06 for calcifying tendinopathy

Tabel 1b: Diagnostic category (n=62) for detecting other pathology causing SAPS

		RC Tendinopathy (T)	RC Calcification (C)	ACJ Arthritis (ACJa)	No pathology found	Overall
Frequency	PT1	23	10	12	11	
	PT2, 3 or 4	26	14	14	8	
	Both	16	10	8	5	
OA		72.6%	93.6%	83.9%	85.5%	85.5%
SPA		77.3%	96%	89.8%	91.4%	91.4%
SNA		65.3%	83.3%	61.5%	52.6%	52.6%
PI		0.21	0.61	0.58	0.69	0.69
BI		0.05	0.06	0.03	0.05	0.05
Cohen's kappa		0.43 (0.2-0.66)	0.8 (0.6-0.99)	0.54 (0.25-0.87)	0.44 (0.14-0.68)	0.44 (0.14-0.68)

Prevalence, Cohen's kappa, overall kappa, percentage (%) of observed agreement (OA), % Specific Positive Agreement (SPA), % Specific Negative Agreement (SNA), Prevalence Index (PI), Bias Index (BI) of tendinopathy, calcification, AC arthritis and 'no details found' Physical Therapist (PT1), Physical Therapist 2–3 or 4 (PT 2–3 or 4)

Discussion

We found high overall agreement as well as high specific positive agreement for detecting rotator cuff ruptures and other pathology causing SAPS. For both, the overall agreement and the positive agreement was higher than the specific negative agreement. Physical therapists specialized in MSU agree more on the presence of rotator cuff tears and other pathology causing SAPS by using ultrasound than on the absence of pathology.

To the best of our knowledge this is the first inter-examiner agreement and reliability study between physical therapists performing diagnostic ultrasound in symptomatic shoulders in primary care. There is one inter-examiner study of US between PTs and radiologists which showed substantial agreement for full thickness tears, moderate agreement for bursitis and fair agreement for calcifying cuff tendinopathy in patients with shoulder pain (Thoomes-de Graaf et al., 2014). Differences between these results and ours can be explained by differences in profession who performed the MSU, differences in equipment, and MSU experience, as well as a difference in study population with Thoomes-de Graaf's group of patients being older compared to our study. In older patients, pathology might be found more frequently than in younger patients (Schmidt et al., 2015).

The kappa value for cuff ruptures varied from fair to moderate agreement and from moderate to substantial for other pathology causing SAPS. Remarkably, our study found a higher kappa value for PTR than for FTR. This is contrary to other studies (Rutten et al., 2006, Nazarian et al., 2013). Two of these studies were done among musculoskeletal radiologists in hospital care and showed excellent agreement on full thickness rotator cuff tears and good agreement for partial thickness rotator cuff tears (Le Correler et al., 2008, Rutten et al., 2010). A possible explanation for this difference may be that patients referred by an orthopaedic surgeon to hospital-based musculoskeletal radiologists already have a higher incidence of rotator cuff tears and have compared with patients in primary care. Furthermore, the MSU skill and experience of physical therapists and dedicated musculoskeletal radiologists is bound to be different. Another explanation for our reduced kappa values may be that we observed higher levels of positive agreement and lower levels of negative agreement resulting in a high prevalence index combined with a low bias index. In these situations,

percentages of agreement are deemed more relevant than kappa values (de Vet, Mokkink, Terwee, Hoekstra, & Knol, 2013).

The scanning protocol focused on anatomy, scan techniques and pitfalls. Although adherence to the protocol may have increased reliability, this was not determined. Some patients were not able to maintain the required position throughout both assessments because of increasing shoulder pain. Although the scanning sequence and reporting was standardized, examination presets (depth, gain, focus, frequency) were not standardized because these are operator and equipment dependent. Although all examining PTs followed a 6-h training on the study MSU protocol, not all specific diagnostic criteria for the various pathologies were discussed as mentioned in the "scan finding sheet". This may have resulted in differences in interpretation and may have negatively influenced the level of inter-examiner agreement but has increased the representativeness in clinical practice. Most discussion during the training session was about differences in a full thickness tear and a partial thickness tear of the RC and between a partial thickness tear and a tendinopathy of the RC. However, the differences in equipment and possibly imaging quality as well as the lack of standardized diagnostic ultrasound criteria are both reflective of the current practice of MSU by PTs in primary care. Another limitation of this study is that both MSU examinations by a selected group of well-trained PTs were not compared with any other imaging modality or MSU by musculoskeletal radiologists. Results from this study may therefore not be readily generalized to all MSU of the shoulder in primary care, either by PTs nor by other professional groups performing MSU, let alone radiologists. The results of this small study need to be confirmed by further research. Validity of diagnostic MSU by PTs in primary care should be examined in future studies in comparison with golden standards, with MSU by dedicated radiologists and imaging modalities such as Magnetic Resonance Imaging or Computed Tomography with Arthrography. This information is needed to confirm clinical value. Furthermore, the role of the (quality of) MSU equipment as used in primary care as well as the influence of the level of training and experience of a much larger group of PTs should be assessed.

In conclusion among a limited group of physical therapists in primary care, the inter-examiner overall agreement for detecting cuff ruptures and other pathology causing SAPS is high, although reliability values are fair for partial thickness tears and slight for full thickness tears. Physical therapists specialized in MSU agree more on the presence of pathology causing SAPS than on the absence of pathology.

Acknowledgement

This study was funded by SIA-RAAK. The Ministry of Education has made this funding available for the innovation and promotion of research. This study was also partly funded by a program grant of the Dutch Arthritis Foundation.

References

Chen, H. S., Lin, S. H., Hsu, Y. H., Chen, S. C., & Kang, J. H. (2011). A comparison of physical examinations with musculoskeletal ultrasound in the diagnosis of biceps long head tendinitis. *Ultrasound Med Biol, 37*(9), 1392-1398. doi:10.1016/j.ultrasmedbio.2011.05.842

de Vet, H. C., Mokkink, L. B., Terwee, C. B., Hoekstra, O. S., & Knol, D. L. (2013). Clinicians are right not to like Cohen's kappa. *BMJ, 346*, f2125. doi:10.1136/bmj.f2125

Diercks, R., Bron, C., Dorrestijn, O., Meskers, C., Naber, R., de Ruiter, T., . . . Dutch Orthopaedic, A. (2014). Guideline for diagnosis and treatment of subacromial pain syndrome: a multidisciplinary review by the Dutch Orthopaedic Association. *Acta Orthop, 85*(3), 314-322. doi:10.3109/17453674.2014.920991

Greving, K., Dorrestijn, O., Winters, J. C., Groenhof, F., van der Meer, K., Stevens, M., & Diercks, R. L. (2012). Incidence, prevalence, and consultation rates of shoulder complaints in general practice. *Scand J Rheumatol, 41*(2), 150-155. doi:10.3109/03009742.2011.605390

Hegedus, E. J., Goode, A., Campbell, S., Morin, A., Tamaddoni, M., Moorman, C. T., 3rd, & Cook, C. (2008). Physical examination tests of the shoulder: a systematic review with meta-analysis of individual tests. *Br J Sports Med, 42*(2), 80-92; discussion 92. doi:10.1136/bjsm.2007.038406

Ian Beggs, S. B., Angel Bueno, Michel Cohen, Michel Court-Payen, Andrew Grainger,, & Franz Kainberger, A. K., Carlo Martinoli, Eugene McNally, Philip J. O'Connor, Philippe Peetrons, Monique Reijnierse, Philipp Remplik, Enzo Silvestri, . (2016). *Musculoskeletal Ultrasound Technical Guidelines I. Shoulder*. Retrieved from https://essr.org/content-essr/uploads/2016/10/shoulder.pdf

Kottner, J., Audige, L., Brorson, S., Donner, A., Gajewski, B. J., Hrobjartsson, A., . . . Streiner, D. L. (2011). Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol, 64*(1), 96-106. doi:10.1016/j.jclinepi.2010.03.002

Kumar, P., Bradley, M., & Swinkels, A. (2010). Within-day and day-to-day intrarater reliability of ultrasonographic measurements of acromion-greater tuberosity distance in healthy people. *Physiother Theory Pract, 26*(5), 347-351. doi:10.3109/09593980903059522

Kumar, P., Chetwynd, J., Evans, A., Wardle, G., Crick, C., & Richardson, B. (2011). Interrater and intrarater reliability of ultrasonographic measurements of acromion-greater tuberosity distance in healthy people. *Physiother Theory Pract, 27*(2), 172-175. doi:10.3109/09593985.2010.481012

Lumsden, G., Lucas-Garner, K., Sutherland, S., & Dodenhoff, R. (2018). Physiotherapists utilizing diagnostic ultrasound in shoulder clinics. How useful do patients find immediate feedback from the scan as part of the management of their problem? *Musculoskeletal Care, 16*(1), 209-213. doi:10.1002/msc.1213

Nazarian, L. N., Jacobson, J. A., Benson, C. B., Bancroft, L. W., Bedi, A., McShane, J. M., . . . Yamaguchi, K. (2013). Imaging algorithms for evaluating suspected rotator cuff disease: Society of Radiologists in Ultrasound consensus conference statement. *Radiology, 267*(2), 589-595. doi:10.1148/radiol.13121947

Okoroha, K. R., Fidai, M. S., Tramer, J. S., Davis, K. D., & Kolowich, P. A. (2018). Diagnostic accuracy of ultrasound for rotator cuff tears. *Ultrasonography*. doi:10.14366/usg.18058

Ottenheijm, R. P., Jansen, M. J., Staal, J. B., van den Bruel, A., Weijers, R. E., de Bie, R. A., & Dinant, G. J. (2010). Accuracy of diagnostic ultrasound in patients with suspected subacromial disorders: a systematic review and meta-analysis. *Arch Phys Med Rehabil, 91*(10), 1616-1625. doi:10.1016/j.apmr.2010.07.017

Rutten, M. J., Jager, G. J., & Blickman, J. G. (2006). From the RSNA refresher courses: US of the rotator cuff: pitfalls, limitations, and artifacts. *Radiographics, 26*(2), 589-604. doi:10.1148/rg.262045719

Scholten-Peeters, G. G., Franken, N., Beumer, A., & Verhagen, A. P. (2014). The opinion and experiences of Dutch orthopedic surgeons and radiologists about diagnostic

musculoskeletal ultrasound imaging in primary care: a survey. *Man Ther, 19*(2), 109-113. doi:10.1016/j.math.2013.08.003

Silva, L., Andreu, J. L., Munoz, P., Pastrana, M., Millan, I., Sanz, J., . . . Fernandez-Castro, M. (2008). Accuracy of physical examination in subacromial impingement syndrome. *Rheumatology (Oxford), 47*(5), 679-683. doi:10.1093/rheumatology/ken101

Smith, T. O., Back, T., Toms, A. P., & Hing, C. B. (2011). Diagnostic accuracy of ultrasound for rotator cuff tears in adults: a systematic review and meta-analysis. *Clin Radiol, 66*(11), 1036-1048. doi:10.1016/j.crad.2011.05.007

Thoomes-de Graaf, M., Scholten-Peeters, G. G., Duijn, E., Karel, Y. H., van den Borne, M. P., Beumer, A., . . . Verhagen, A. P. (2014). Inter-professional agreement of ultrasound-based diagnoses in patients with shoulder pain between physical therapists and radiologists in the Netherlands. *Man Ther, 19*(5), 478-483. doi:10.1016/j.math.2014.04.018

Vet, H. C. W. d. (2011). *Measurement in medicine : a practical guide*. Cambridge ; New York: Cambridge University Press.