Accepted Manuscript

Title: Inter-rater reliability of shoulder measurements in middle-aged women

Author: A. De Groef M. Van Kampen N. Vervloesem E. Clabau M.-R. Christiaens P. Neven I. Geraerts F. Struyf N. Devoogdt

PII: S0031-9406(16)30038-4
DOI: <http://dx.doi.org/doi:10.1016/j.physio.2016.07.002>
Reference: PHYST 923

To appear in: *Physiotherapy*

Received date: 16-9-2015

Please cite this article as: De Groef A, Van Kampen M, Vervloesem N, Clabau E, Christiaens M-R, Neven P, Geraerts I, Struyf F, Devoogdt N, Inter-rater reliability of shoulder measurements in middle-aged women, *Physiotherapy* (2016), <http://dx.doi.org/10.1016/j.physio.2016.07.002>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Inter-rater reliability of shoulder measurements in middle-aged women

A. De Groef^{a,*}, M. Van Kampen^a, N. Vervloesem^a, E. Clabau^a, M-R. Christiaens^{b,c}, P. Neven^{b,d}, I. Geraerts^a, F. Struyf^e, N. Devoogdt^a

^a*Department of Rehabilitation Sciences, KU Leuven – University of Leuven, and Department of Physical Medicine and Rehabilitation, University Hospitals Leuven, Leuven, Belgium*

^b*Multidisciplinary Breast Centre, University Hospitals Leuven, Leuven, Belgium*

^c*Department of Surgical Oncology, KU Leuven – University of Leuven, Leuven, Belgium*

^d*Department of Obstetrics and Gynaecology, University Hospitals Leuven, Leuven, Belgium*

^e*Department of Rehabilitation Sciences and Physiotherapy, University of Antwerp, Antwerp, Belgium*

*Corresponding author. Address: Department of Physical Medicine and Rehabilitation, University Hospitals Leuven, Herestraat 49, 3000 Leuven, Belgium. Tel.: +32 16 342171.

E-mail address: an.degroef@faber.kuleuven.be (A. De Groef).

1 **Abstract**

2 **Objectives** To investigate inter-rater reliability of a set of shoulder measurements including inclinometry
3 [shoulder range of motion (ROM)], acromion–table distance and pectoralis minor muscle length (static
4 scapular positioning), upward rotation with two inclinometers (scapular kinematics) and pain pressure
5 thresholds (muscle tenderness) in middle-aged women.

6 **Design** Observational study.

7 **Participants** Thirty symptom-free middle-aged women (first cohort) were measured by two raters. All
8 measurements with an intraclass correlation coefficient (ICC) below 0.75 were retested after an additional
9 training period in a second cohort of 30 symptom-free middle-aged women.

10 **Main outcome measures** Inter-rater reliability of all variables was measured with the ICC (95% confidence
11 interval) and standard error of measurement (SEM).

12 **Results** Acromion–table distance (ICC=0.91, SEM 0.22 to 0.28% of body length), pectoralis minor muscle
13 length (ICC=0.91, SEM 0.16% of body length), pain pressure thresholds (ICC=0.78 to 0.85, SEM 0.39 to
14 0.70 kg) and abduction ROM (ICC=0.77, SEM 5°) showed good to excellent inter-rater reliability in the first
15 cohort. After an additional training period, forward flexion ROM showed good inter-rater reliability
16 (ICC=0.83, SEM 5°), scapular upward rotation in resting position showed moderate reliability (ICC=0.52,
17 SEM 2°), and other scaption angles showed weak reliability (ICC=0.26 to 0.43, SEM 3 to 8°).

18 **Conclusions** In a battery of clinical tools to evaluate factors contributing to shoulder pain, static scapular
19 positioning and pressure pain thresholds were found to have good to excellent inter-rater reliability in
20 middle-aged women. Additional training is recommended for measurements with a gravity inclinometer.

21
22 **Keywords:** Inter-rater reliability; Inclinometry; Algometry; Static scapular positioning

<A>Introduction

Depending on a person's age, sex and activity level, prevalence rates of shoulder problems vary between 7% and 67% [1]. Shoulder pain is a common musculoskeletal problem influenced by several factors. The most common factors contributing to the emergence or maintenance of shoulder pain are impaired shoulder mobility, altered static posture, and altered scapular kinematics and muscle recruitment patterns [2–5]. In addition, muscle tenderness and myofascial trigger points have been suggested as sources of shoulder pain [6,7]. However, at present, there is no consensus on the definitive relationship between these factors and painful shoulder pathologies, such as subacromial impingement syndrome [3].

During clinical examination, possible contributing factors to pain should be evaluated. In addition to visual evaluation and clinical tests, clinical measurement tools can be used, and these should be valid, reliable and easy to use. For the evaluation of static posture, several tools have been found to have good to excellent reliability [8,9]. In clinical practice, shoulder protraction can be evaluated by measuring the distance of the posterior border of the acromion to the table with a sliding caliper. A sliding caliper can also be used to measure pectoralis minor muscle length, which correlates to the acromion–table distance [8,10–12]. An inclinometer is recommended for the evaluation of shoulder mobility, as this can be applied with one hand [13]. For dynamic evaluation of the scapula, the lateral scapular slide test is less suitable and reliable than the measurement of scapular upward rotation using two inclinometers [14,15]. For the evaluation of muscle tenderness and myofascial trigger points, the pressure pain threshold (PPT) can be measured by an algometer [16]. An overview of the reliability of these clinical tools [intraclass correlation coefficient (ICC) and standard error of measurement (SEM)], characteristics of the study population and average values is given in Table 1.

<insert Table 1 near here>

Reliability has mainly been tested in symptomatic cohorts of middle-aged men and women [10,17–22]. The asymptomatic cohorts were all of a younger age, and differentiation was only made between men and women in one study [16]. However, middle-aged women are most likely to consult a general

practitioner because of shoulder complaints [23]. Additionally, it is well known that total and regional fat increases with age [24], and is greater in menopausal women [25,26]. This feature may compromise palpation for bony landmarks, and therefore the accuracy of static and dynamic evaluation of the scapula. It may also complicate muscle palpation when evaluating pressure pain thresholds. To the authors' knowledge, no previous studies have investigated the reliability of a set of clinical tools to evaluate factors contributing to shoulder pain in a cohort of middle-aged women. Also, most previous studies have only investigated intrarater reliability [15,16,19,22,27]. No previous studies have investigated the reliability of a set of clinic tools in a single population. A set of tests is needed to capture the factors contributing to shoulder pain. Finally, previous studies have not determined if certain tools need a learning period before they can be applied reliably in clinical practice.

As such, the aim of this study was to test inter-rater reliability of a set of clinical assessment tools for several factors contributing to shoulder pain that are easy to use in clinical practice. This set of clinical tools consists of static scapular positioning (acromion–table distance and pectoralis minor muscle length), shoulder mobility (inclinometry), dynamic scapular evaluation (scapular upward rotation with two inclinometers) and PPTs (algometry).

<A>Methods

The Guidelines for Reporting Reliability and Agreement Studies (GRASS) were used as a basis to report this reliability study [28]. This study was approved by the Ethical Committee of the University Hospitals Leuven (Ref No. s54579).

Subjects

Two convenience samples of asymptomatic middle-aged women (40 to 60 years) with no history of shoulder pathologies were recruited from the Department of Physical Medicine, University Hospital of Leuven and through the families and friends of the investigators. Women had to score 0 on the visual analogue scale for shoulder pain over the past 3 months. The first cohort of 30 women was measured between December 2012 and July 2013. In order to evaluate the need for a learning period, all measurements with ICC below 0.75

were retested after an additional training period of 3 months. This second cohort consisted of 30 different women, and measurements were taken between September and December 2013.

Procedure

Measurements were taken independently by two raters, one of whom had a Masters in Physical Therapy and Rehabilitation, and the other had 1 year of clinical experience in musculoskeletal rehabilitation. Prior to the study, they underwent two types of training. First, a 4-hour training session was held for accuracy of measurements, and second, training was performed on 20 healthy subjects. Between the first and second cohort, both raters were active in clinical practice; an additional 2-hour training session was implemented, and retraining was performed on 20 healthy subjects.

Both shoulders of all women were measured within a single testing session. The order of testing (right or left shoulder and the choice of rater) was chosen at random. Both raters were blinded to the results of each other's measurements as measurements were conducted in separate rooms. The rater was alone in the room to avoid a Hawthorne effect (i.e. where individuals modify or improve an aspect of their behaviour in response to their awareness of being observed) [29].

<C>Static scapular positioning (Fig. 1)

Static scapular positioning was evaluated using acromion–table distance and pectoralis minor muscle length. For acromion–table distance, the subject was in a supine position and was asked to adopt a natural relaxed position. The distance between the most posterior border of the acromion and the table was measured with a sliding caliper (Hogetex, 0 to 300 mm). The measurement was repeated with the patient actively retracting both shoulders. The score obtained was divided by body length, resulting in the acromion–table index [12].

The resting muscle length of the pectoralis minor was measured between the caudal edge of the 4th rib and the inferomedial aspect of the coracoid process with a sliding caliper. This measurement was also normalised to body length, resulting in the pectoralis minor index.

<insert Fig 1 near here>

<C>*Shoulder range of motion (Fig. 2)*

Shoulder range of motion (ROM) was measured with a Dr. Rippstein Plurimeter-V gravity reference analogue inclinometer (www.dr-rippstein.ch). The inclinometer was velcro taped perpendicular to the humeral shaft, just below the deltoid tuberosity. The subject was instructed to perform one arm abduction movement in the coronal plane and one forward flexion arm movement in the frontal plane with full elbow extension, neutral wrist flexion/extension and with the thumb leading to ensure vertical alignment of the inclinometer. The rater corrected the patient for compensations [13]. The degree of maximum shoulder ROM of both movements was recorded.

<insert Fig 2 near here>

<C>*Scapular upward rotation (Fig. 3)*

The measurement of scapular upward rotation was performed using two Dr. Rippstein Plurimeter-V gravity reference analogue inclinometers, as described by Watson *et al.* [15]. Movement in the scapular plane was chosen because it is widely used in other clinical trials, and it is a more functional direction of arm elevation [30–35]. One inclinometer was velcro taped perpendicular to the humeral shaft, just below the deltoid tuberosity. The second inclinometer was aligned manually along the scapular spine. The subject was asked to perform one arm abduction movement in the scapular plane, i.e. scaption (30° anterior to the coronal plane) with full elbow extension, neutral wrist flexion/extension and with the thumb leading. The movement was guided by the raters' instructions. The subject was asked to stop at 45°, 90°, 135° and at their maximum achievable arm movement. At each of the scaption positions, the degree of scapular upward rotation was measured with the second inclinometer [15].

<insert Fig 3 near here>

<C>Pressure pain thresholds (Fig. 4)

Pressure pain thresholds at different locations in the upper body were measured by a digital Wagner FPX algometer (Greenwich, CT, USA). Points of measurement were defined by palpation for tender muscle points in the region of the upper trapezius (between the C7 spinous process and the acromion), supraspinatus (above the spine of the scapula), infraspinatus (muscle belly under the spine of the scapula), pectoralis major (under the clavicle), pectoralis minor (between the caudal edge of the 4th rib and the inferomedial aspect of the coracoid process) and serratus anterior (below the axilla, on the muscle belly which branches to the ribs). The assessor placed the algometer on the examination point and pressed against the device in a vertical direction while increasing the force at a constant rate of 1 kg/second. The subject was asked to say 'stop' when the sensation of pressure first changed to pain. After examining all muscles, the subjects rested for 5 minutes. After the resting period, the procedure was repeated. The mean value of the two measurements was calculated and used for the analysis [36].

<insert Table 4 near here>

Statistical analysis

Statistical analysis was performed using Statistical Package for the Social Sciences Version 22 (IBM Corp., Armonk, NY, USA). The ICC for single measurements [ICC(2.1)] based on a two-way random effects analysis of variance model was used to determine the inter-rater reliability of acromion–table distance, pectoralis minor muscle length, shoulder ROM and scapular upward rotation [37]. For PPTs, average measurements were used for the ICC [ICC(2.2)] [38]. Data from both dominant and non-dominant sides were pooled. According to the guidelines offered by Portney and Watkins [39], an ICC below 0.50 indicates poor reliability, an ICC between 0.51 and 0.75 indicates moderate reliability, an ICC between 0.75 and 0.90 indicates good reliability, and an ICC above 0.90 indicates excellent reliability.

To quantify the amount of error inherent in a measurement, the standard error of measurement (SEM) was calculated. The following formula was used: $SEM = SD_{12} \sqrt{1 - ICC}$, where SD_{12} indicated the average standard deviation (SD) of the two ratings [38]. The SEM provides information on whether

observed change is likely to be due to measurement error or to real change. The 2SEM provides the clinician with 95% confidence that the true measurement value lies within ± 2 SEM from the clinical measurement. Finally, Bland and Altman limits of agreement analyses were performed [37]. For each measurement, a graph was drawn with the mean of the two assessments on the X-axis and the difference between the two assessments on the Y-axis. One line was drawn to indicate the mean difference between two assessments, and two lines for the 95% limits of agreement (average difference ± 1.96 standard deviation of the difference). If the differences within these limits are not clinically important, the two ratings may be used interchangeably.

<A>Results

In the first cohort, 60 shoulders of 30 asymptomatic women were available for reliability analysis. The mean age of the women was 50.3 (SD 7.3) years, with a range from 37.7 to 70.7 years. Mean body mass index (BMI) was 25.2 (SD 3.0) kg/m², with a range from 21.0 to 31.6. Table 2 shows the ICC of the inter-rater measurements, 95% confidence intervals (CI) and average values. ICC values for static scapular position and PPTs reached the critical value of 0.75, indicating good reliability. Reliability of the tests with the gravity inclinometer did not reach the critical value of 0.75, except for shoulder abduction ROM (Table 2). Therefore, 60 different asymptomatic shoulders were measured to retest the inter-rater reliability of these measurements.

<insert Table 2 near here>

The mean age of the second cohort of women was 54.7 (11.3) years, with a range from 37.9 to 88.4 years. Mean BMI was 25.8 (3.8) kg/m², with a range from 19.1 to 34.9 kg/m². The ICC of forward flexion ROM increased to 0.83 (95% CI 0.73 to 0.89). ICC values of scapular upward rotation increased but did not reach the critical value of 0.75. Table 3 shows the ICC and 95% CI and average values of the inter-rater measurements in the second cohort.

<insert Table 3 near here>

The SEM and 2SEM of the first and second cohort are given in Tables 2 and 3, respectively.

Additionally, Bland and Altman limits of agreement analyses were performed. First, for all measurements, the mean difference was situated around zero, indicating that no systematic differences between raters were present (Table 4). Second, the scatter plots were equally divided around the mean difference. Third, 95% limits of agreement were all clinically acceptable (Table 4). Except for scapular upward rotation, 95% limits of agreement became larger with increasing scaption angles.

<insert Table 4 near here>

<A>Discussion

To the authors' knowledge, this is the first study to investigate the reliability of a set of clinical tools for the evaluation of several factors contributing to shoulder pain in a cohort of asymptomatic women with moderate to high mean BMI and age, and thus more local subcutaneous fat tissue mass around the shoulder region. For static scapular position, PPTs and abduction ROM ICC values were 0.75 or more, indicating good to excellent inter-rater reliability. SEM and 2SEM values were relatively small. After additional training, forward flexion ROM showed good inter-rater reliability, and scapular upward rotation in resting position showed moderate reliability. Weak inter-rater reliability was found for the other scaption angles, even after additional training.

The inter-rater reliability of static scapular position measured by acromion–table distance was comparable with the study of Nijs *et al.* [20] (Table 1). They found good to excellent reliability in symptomatic men and women, despite the fact that they used a less solid measurement tape [20].

Struyf *et al.* found moderate inter-rater reliability for pectoralis minor muscle length, measured with the same device as the present study in younger asymptomatic men and women, compared with good to excellent reliability in the present study of middle-aged women alone [10]. A possible explanation may be

the short training period and limited experience in palpation of the bony landmarks among the students who performed the tests. Secondly, variability in the pectoralis minor muscle length was low in their study population [10]. Lewis *et al.* used a comparable measurement device but only investigated intrarater reliability [19]. The SEM values and 95% limits of agreement of the Bland and Altman analyses for static scapular positioning were very small, indicating very small error of measurement.

Abduction ROM was the only measurement with the gravity inclinometer that reached good reliability without the need for additional training. The slight increase in reliability and decrease in measurement error for forward flexion ROM after the additional training may be attributed to an improvement in palpation skills and increased accuracy of placing the inclinometer on the upper arm. ICC values in the present study (Tables 2 and 3) are comparable or better than the inter-rater ICC values in the studies by Green *et al.* [17], Hoving *et al.* [18] and Tveita *et al.* [21] (Table 1). These studies indicate moderate to good inter-rater reliability of shoulder ROM measurements in men and women with and without symptoms [17,18,21]. A study investigating the intrarater reliability of shoulder ROM reported SEM values between 2° and 3° in asymptomatic women and between 5° and 14° in symptomatic subjects [22], compared with 5° in the present study. Differences in SEM may be explained by differences in the study population. Greater variability, and therefore greater SEM values, in performing the movements may be due to pain or altered biomechanics in the symptomatic population [22], or higher average age and thus age-related differences in shoulder ROM in the present study. Even after additional training, inter-rater reliability for scapular upward rotation for different angles of scaption only increased from fairly reliable to fairly to moderately reliable (Tables 2 and 3). Compared with the measurements of the first cohort, the SEM decreased, indicating the possible advantage of additional training. To the authors' knowledge, no studies have investigated the inter-rater reliability of scapular upward rotation measured by two gravity inclinometers. One study investigated intrarater reliability and found good to excellent reliability [15]. Several possible mechanisms contributing to their high reliability values should be mentioned. First, in the study by Watson *et al.*, patients had a lower average age and 71% of the patients participated in sports activities [15]. Due to more subcutaneous fat tissue mass at the level of the scapula and upper arm in the population of the present study, palpation of the bony landmarks and positioning of the inclinometer may

have been more difficult, resulting in less accurate measurements. Additionally, the inclinometers are placed on muscles, which are not rigid [40]. Second, Watson *et al.* used shoulder abduction in the coronal plane. In the present study, shoulder abduction movement was measured in the scapular plane as this is a more functional plane for arm elevation and is widely used in other studies [30–35]. However, this movement can be more difficult to perform in a standardised manner. Moreover, in the study by Watson *et al.*, subjects were standing compared with sitting in the present study. Third, when there is enough variation in the sample, ICC values will be high. As the present study used a homogenous sample of asymptomatic women, variation may be low, resulting in lower ICC values [41,42]. Likewise, ICC values will be higher in study populations with more variation, such as the symptomatic women in the study by Watson *et al.* [15]. In addition, scapular upward rotation and scapular abduction may be inherently variable, depending on, for example, patients' warm up. Additionally, by asking the patient to stop and hold their arm statically at different degrees of scaption, an isometric muscle contraction is required, and this may also influence the measurement. Finally, ICC values for single measurements were tested in the present study. It has been shown in the literature that ICC values for average measurements are generally higher [39]. The SEM values in the study by Watson *et al.* are comparable to the SEM values in the present study. SEM values increase with increasing scaption or abduction [15]. In the present study, PPTs reached good reliability (Table 2). To the authors' knowledge, no studies have investigated the inter-rater reliability of PPTs in the upper limb region. Jones *et al.* found excellent intrarater reliability of algometry of upper limb and upper torso muscles in asymptomatic women with a lower mean age compared with the population in the present study [16]. Comparison with the study by Vanderweeën *et al.* is difficult because of varying age ranges (15 to 75 years), omission of information about BMI, inclusion of symptomatic men and women, and only one rater performed the tests [27]. A recent study reported an SEM of 0.19 kg/m² for PPT of the paravertebral muscles of the back [43].

This study had a few limitations. First, reliability was only tested in healthy asymptomatic women and not in women with shoulder problems. As variability in asymptomatic subjects will be smaller, ICC values will also be smaller. Second, in the present study, only inter-rater reliability was investigated. Intrarater reliability was not tested as it was assumed that intrarater reliability would be as good as or even

better than inter-rater reliability [44]. Third, the authors acknowledge that there may be other factors contributing to shoulder pain that should be considered. Fourth, the most easy-to-use clinical tools were used in this study, but other clinical tools may be equally important. Finally, patients' age ranged between 37 and 88 years instead of the proposed 40 to 60 years due to recruitment difficulties. This study also had several strengths. First, the measurements were performed in the 'field' with the same disadvantages as when performed for clinical purposes, such as time limitations and physical limitations of the patient. Second, a set of different clinical tools to test several factors contributing to shoulder pain was tested in the same population at the same time. Third, the need for additional training for the clinical tools was explored.

In clinical practice, the evaluation of several factors contributing to shoulder pain using a combination of the presented clinical tools is recommended. The clinical tools may be useful in all women with shoulder pain and in women at risk of developing shoulder pain; for example, women after breast cancer treatment [45,46] or patients who have had a stroke [47,48].

For measurements with the gravity inclinometer, the implementation of an extensive training period to train palpation skills and accuracy of placing the inclinometer is recommended. However, performing the measurement in a standing position may increase reliability. Further, use of a reference stick to guide the movement of the arm in the scapular plane is recommended. One study that investigated the reliability of an electronic inclinometer of scapular upward rotation found good intrarater reliability when using a reference stick [33].

Inter-rater reliability of scapular upward rotation measurements with two gravity inclinometers should be further investigated in future research. Tucker *et al.*'s protocol for measurement with a reference stick can be used [33]. In addition, ICC values for average measurements should be established, instead of single measurements. Further research is needed to define the exact duration and intensity of the training period for measurements with an inclinometer.

<A>Conclusions

In conclusion, a set of clinical tools for the evaluation of static scapular positioning and PPTs have good to excellent reliability in women with moderate to high mean BMI and age. A training period for measurement

with a gravity inclinometer is recommended. Only moderate reliability was found for evaluation of upward rotation of the scapula using two gravity inclinometers in resting position.

Ethical approval: This study was approved by the Ethical Committee of the University Hospitals Leuven (Ref No. s54579).

Funding: This study was funded by the agency for Innovation by Science and Technology (Applied Biomedical Research) (IWT 110703).

Conflict of interest: None declared.

<A>References

1. Luime JJ, Koes BW, Hendriksen IJ, 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;33:73–81.
2. Borsa PA, Laudner KG, Sauers EL. Mobility and stability adaptations in the shoulder of the overhead athlete: a theoretical and evidence-based perspective. *Sports Med* 2008;38:17–36.
3. Ratcliffe E, Pickering S, McLean S, Lewis J. Is there a relationship between subacromial impingement syndrome and scapular orientation? A systematic review. *Br J Sports Med* 2014;48:1251–6.
4. Ludewig PM, Cook TM. Alterations in shoulder kinematics and associated muscle activity in people with symptoms of shoulder impingement. *Phys Ther* 2000;80:276–91.
5. Kibler WB, Sciascia A. Current concepts: scapular dyskinesis. *Br J Sports Med* 2010;44:300–5.
6. Bron C, Dommerholt J, Stegenga B, Wensing M, Oostendorp RA. High prevalence of shoulder girdle muscles with myofascial trigger points in patients with shoulder pain. *BMC Musculoskel Disord* 2011;12:139–51.
7. Sergienko S, Kalichman L. Myofascial origin of shoulder pain: a literature review. *J Bodywork Movement Ther* 2015;19:91–101.

8. Struyf F, Nijs J, Mottram S, Roussel NA, Cools AM, Meeusen R. Clinical assessment of the scapula: a review of the literature. *Br J Sports Med* 2012;48:883–90.
9. Larsen CM, Juul-Kristensen B, Lund H, Sogaard K. Measurement properties of existing clinical assessment methods evaluating scapular positioning and function. A systematic review. *Physiother Theory Pract* 2014;30:453–82.
10. Struyf F, Meeus M, Fransen E, Roussel N, Jansen N, Truijen S, *et al.* Interrater and intrarater reliability of the pectoralis minor muscle length measurement in subjects with and without shoulder impingement symptoms. *Man Ther* 2014;19:294–8.
11. Borstad JD, Ludewig PM. The effect of long versus short pectoralis minor resting length on scapular kinematics in healthy individuals. *J Orthopaed Sports Phys Ther* 2005;35:227–38.
12. Host HH. Scapular taping in the treatment of anterior shoulder impingement. *Phys Ther* 1995;75:803–12.
13. Kolber MJ, Hanney WJ. The reliability and concurrent validity of shoulder mobility measurements using a digital inclinometer and goniometer: a technical report. *Int J Sports Phys Ther* 2012;7:306–13.
14. Struyf F, Nijs J, De Coninck K, Giunta M, Mottram S, Meeusen R. Clinical assessment of scapular positioning in musicians: an intertester reliability study. *J Athl Train* 2009;44:519–26.
15. Watson L, Balster SM, Finch C, Dalziel R. Measurement of scapula upward rotation: a reliable clinical procedure. *Br J Sports Med* 2005;39:599–603.
16. Jones DH, Kilgour RD, Comtois AS. Test–retest reliability of pressure pain threshold measurements of the upper limb and torso in young healthy women. *J Pain* 2007;8:650–6.
17. Green S, Buchbinder R, Forbes A, Bellamy N. A standardized protocol for measurement of range of movement of the shoulder using the Plurimeter-V inclinometer and assessment of its intrarater and interrater reliability. *Arthritis Care Res* 1998;11:43–52.
18. Hoving JL, Buchbinder R, Green S, Forbes A, Bellamy N, Brand C, *et al.* How reliably do rheumatologists measure shoulder movement? *Ann Rheum Dis* 2002;61:612–6.

19. Lewis JS, Valentine RE. The pectoralis minor length test: a study of the intra-rater reliability and diagnostic accuracy in subjects with and without shoulder symptoms. *BMC Musculoskel Disord* 2007;8:64–74.
20. Nijs J, Roussel N, Vermeulen K, Souvereyns G. Scapular positioning in patients with shoulder pain: a study examining the reliability and clinical importance of 3 clinical tests. *Arch Phys Med Rehabil* 2005;86:1349–55.
21. Tveita EK, Ekeberg OM, Juel NG, Bautz-Holter E. Range of shoulder motion in patients with adhesive capsulitis; intra-tester reproducibility is acceptable for group comparisons. *BMC Musculoskel Disord* 2008;9:49–58.
22. Valentine RE, Lewis JS. Intraobserver reliability of 4 physiologic movements of the shoulder in subjects with and without symptoms. *Arch Phys Med Rehabil* 2006;87:1242–9.
23. Greving K, Dorrestijn O, Winters JC, Groenhof F, van der Meer K, Stevens M, *et al.* Incidence, prevalence, and consultation rates of shoulder complaints in general practice. *Scand J Rheumatol* 2012;41:150–5.
24. Kuk JL, Saunders TJ, Davidson LE, Ross R. Age-related changes in total and regional fat distribution. *Ageing Res Rev* 2009;8:339–48.
25. Douchi T, Yamamoto S, Yoshimitsu N, Andoh T, Matsuo T, Nagata Y. Relative contribution of aging and menopause to changes in lean and fat mass in segmental regions. *Maturitas* 2002;42:301–6.
26. Toth MJ, Tchernof A, Sites CK, Poehlman ET. Menopause-related changes in body fat distribution. *Ann NY Acad Sci* 2000;904:502–6.
27. Vanderweeen L, Oostendorp RA, Vaes P, Duquet W. Pressure algometry in manual therapy. *Man Ther* 1996;1:258–65.
28. Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hrobjartsson A, *et al.* Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol* 2011;64:96–106.
29. Wickstrom G, Bendix T. The "Hawthorne effect" – what did the original Hawthorne studies actually show? *Scand J Work Environ Health* 2000;26:363–7.

30. Borsa PA, Timmons MK, Sauers EL. Scapular-positioning patterns during humeral elevation in unimpaired shoulders. *J Athl Train* 2003;38:12–7.
31. Johnson MP, McClure PW, Karduna AR. New method to assess scapular upward rotation in subjects with shoulder pathology. *J Orthopaed Sports Phys Ther* 2001;31:81–9.
32. Cools AM, Palmans T, Johansson FR. Age-related, sport-specific adaptations of the shoulder girdle in elite adolescent tennis players. *J Athl Train* 2014;49:647–53.
33. Tucker WS, Ingram RL. Reliability and validity of measuring scapular upward rotation using an electrical inclinometer. *J Electromyog Kinesiol* 2012;22:419–23.
34. Scibek JS, Carcia CR. Assessment of scapulohumeral rhythm for scapular plane shoulder elevation using a modified digital inclinometer. *World J Orthoped* 2012;3:87–94.
35. Wassinger CA, Sole G, Osborne H. Clinical measurement of scapular upward rotation in response to acute subacromial pain. *J Orthopaed Sports Phys Ther* 2013;43:199–203.
36. Fischer AA. Pressure algometry over normal muscles. Standard values, validity and reproducibility of pressure threshold. *Pain* 1987;30:115–26.
37. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
38. Lexell JE, Downham DY. How to assess the reliability of measurements in rehabilitation. *Am J Phys Med Rehabil* 2005;84:719–23.
39. Portney LG WM. *Foundations of clinical research: applications to practice*. 3rd edn. 2009.
40. Reenders SJ TC. Validity and reliability of the Blankenship electronic inclinometer. *J Othopaed Sports Phys Ther* 1998;27:59.
41. Struyf F, Nijs J, Horsten S, Mottram S, Truijen S, Meeusen R. Scapular positioning and motor control in children and adults: a laboratory study using clinical measures. *Man Ther* 2011;16:155–60.
42. Struyf F, Nijs J, De Graeve J, Mottram S, Meeusen R. Scapular positioning in overhead athletes with and without shoulder pain: a case–control study. *Scand J Med Sci Sports* 2011;21:809–18.

43. Koo TK, Guo J-Y, Brown CM. Test–retest reliability, repeatability, and sensitivity of an automated deformation-controlled indentation on pressure pain threshold measurement. *J Manipul Physiol Ther* 2013;36:84–90.
44. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;86:420–8.
45. Hayes SC, Johansson K, Stout NL, Prosnitz R, Armer JM, Gabram S, *et al.* Upper-body morbidity after breast cancer: incidence and evidence for evaluation, prevention, and management within a prospective surveillance model of care. *Cancer* 2012;118(Suppl.):2237–49.
46. Hidding JT, Beurskens CH, van der Wees PJ, van Laarhoven HW, Nijhuis-van der Sanden MW. Treatment related impairments in arm and shoulder in patients with breast cancer: a systematic review. *PloS One* 2014;9:e96748.
47. Langhorne P, Stott DJ, Robertson L, MacDonald J, Jones L, McAlpine C, *et al.* Medical complications after stroke: a multicenter study. *Stroke* 2000;31:1223–9.
48. McLean DE. Medical complications experienced by a cohort of stroke survivors during inpatient, tertiary-level stroke rehabilitation. *Arch Phys Med Rehabil* 2004;85:466–9.

Overview of the reliability of clinical assessment tools to evaluate the contributing factors of shoulder pain

Clinical test	Author	Population characteristics	Age and BMI/weight Mean (SD)	Rating 1 Mean (SD)	Rating 2 Mean (SD)	Reliability	SEM
Posterior acromion–table distance (mm)	Nijs <i>et al.</i> 2005 (20)	Men and women with a variety of shoulder disorders ($n=29$)	56.6 (14.9) years	I. Symptomatic side Relaxed: 72 (15) Retracted: 46 (21) II. Asymptomatic side Relaxed: 72 (13) Retracted: 48 (20)	I. Symptomatic side Relaxed: 72 (14) Retracted: 47 (20) II. Asymptomatic side Relaxed: 72 (14) Retracted: 47 (20)	I. Relaxed: ICC 0.88 to 0.94 II. Retracted: ICC 0.92 to 0.91 (two raters)	I. Symptomatic side Relaxed: 3 to 4 Retracted: 4 to 5 II. Asymptomatic side Relaxed: 2 Retracted: 4
	Lewis <i>et al.</i> 2007 (19)	I. Men and women with unilateral shoulder pain and/or restriction ($n=45$) II. Asymptomatic men and women ($n=45$)	I. 42.8 (16.6) years; 71.4 (11.8) kg II. 32.1 (7.3) years; 70.4 (14.2) kg	I. 61 (13) II. 60 (14)	n.a.	I. ICC 0.91 II. ICC 0.92 to 0.93 (single measurement, one rater) I. ICC 0.95 II. ICC 0.92 to 0.96 (average	I. 4 to 5 mm II. 3 to 4 mm

BMI, body mass index; ROM, range of motion; PPT, pressure pain threshold; SD, standard deviation; SEM, standard error of measurement; ICC, intraclass correlation coefficient; n.a., not applicable; $\diamond$, information not provided by authors.

^aMean movement measured in six patients by six raters, round 1.

^bMean movement measured in six patients by six raters, round 2.

Table 2

Inter-rater reliability ($n=30$) of different clinical measurement tools for the shoulder in a population of women with higher average body mass index and age: first cohort

Mean $\pm$ SD							
		Rater 1	Rater 2	ICC	95% CI	SEM	2SEM
Scapular positioning using a sliding caliper							
Acromion–table index	(neutral position) (% of body length)	2.7 (0.7)	2.7 (0.7)	0.91	0.85 to 0.95	0.22	0.44
Acromion–table index	(retraction) (% of body length)	1.6 (0.8)	1.6 (0.8)	0.87	0.79 to 0.92	0.28	0.56
Pectoralis minor index	(% of body length)	9.2 (0.6)	9.3 (0.6)	0.91	0.85 to 0.94	0.16	0.32
Acromion–table distance	(neutral position) (mm)	4.4 (1.3)	4.5 (1.3)	0.91	0.85 to 0.95	0.38	0.76
Acromion–table distance	(retraction) (mm)	2.6 (1.3)	2.7 (1.4)	0.87	0.79 to 0.92	0.49	0.98
Pectoralis minor length	(mm)	15.3 (0.9)	15.4 (0.9)	0.91	0.85 to 0.94	0.28	0.56
Shoulder mobility using a gravity inclinometer							
Forward flexion (°)		163 (10)	164 (11)	0.67	0.51 to 0.79	6	12
Abduction (°)		150 (11)	151 (13)	0.77	0.65 to 0.86	5	10
Scapular upward rotation using two gravity inclinometers							
Shoulder resting position	(°)	-2 (4)	-1 (3)	0.21	-0.04 to 0.43	3	6

45° scaption (°)	1 (5)	3 (5)	0.16	-0.92	to 4	8
				0.39		
90° scaption (°)	11 (6)	12 (6)	0.14	-0.11	to 5	10
				0.38		
135° scaption (°)	33 (9)	31 (9)	0.23	-0.03	to 8	16
				0.46		
Maximal scaption (°)	47 (10)	43 (11)	0.39	0.14 to 0.59	8	16
PPTs using a digital algometer						
Upper trapezius (kg/cm ²)	4.5 (1.7)	4.6 (1.7)	0.89	0.82 to 0.93	0.70	1.40
Supraspinatus (kg/cm ²)	4.6 (1.5)	4.9 (1.6)	0.87	0.79 to 0.96	0.56	1.12
Infraspinatus (kg/cm ²)	4.7 (1.7)	4.9 (1.7)	0.89	0.81 to 0.93	0.50	1.00
Serratus anterior (kg/cm ²)	4.5 (1.4)	4.5 (1.5)	0.90	0.83 to 0.94	0.59	1.18
Pectoralis major (kg/cm ²)	2.8 (1.1)	2.9 (1.2)	0.91	0.84 to 0.94	0.45	0.90
Pectoralis minor (kg/cm ²)	2.9 (1.1)	2.8 (1.0)	0.92	0.87 to 0.95	0.39	0.78

SEM, standard error of measurement; SD, standard deviation; ICC, intraclass correlation coefficient; CI, confidence interval; PPT, pressure pain threshold.

Table 3

Inter-rater reliability ($n=30$) of different clinical measurement tools for the shoulder in a population of women with higher average body mass index and age: second cohort

Mean $\pm$ SD	Rater 1	Rater 2	ICC	95% CI	SEM	2SEM
Shoulder mobility using a gravity inclinometer						
Forward flexion (°)	161 (12)	161 (11)	0.83	0.73 to 0.89	5	10
Scapular upward rotation using two gravity inclinometers						
Resting position (°)	0 (2)	0 (3)	0.52	0.31 to 0.68	2	8
45° scaption (°)	3 (4)	3 (3)	0.35	0.10 to 0.55	3	6
90° scaption (°)	13 (5)	13 (5)	0.43	0.20 to 0.61	4	8
135° scaption (°)	36 (6)	35 (6)	0.39	0.16 to 0.59	5	10
Maximal scaption (°)	43 (8)	44 (7)	0.26	0.01 to 0.48	8	16
SD, standard deviation; ICC, intraclass correlation coefficient; CI, confidence interval.						

Table 4

Bland and Altman limits of agreement analyses

	Mean difference	95% limits of agreement
Acromion–table index (neutral position) (% of body length)	0.04	-0.59 to 0.67
Acromion–table index (retraction) (% of body length)	0.07	-0.75 to 0.89
Pectoralis minor index (% of body length)	0.05	-0.41 to 0.52
Acromion–table distance (neutral position) (mm)	0.07	-0.99 to 1.12
Acromion–table distance (retraction) (mm)	0.11	-1.26 to 1.49
Pectoralis minor length (mm)	0.09	-0.67 to 0.85
Forward flexion (°)	0.09	-13.61 to 13.77
Abduction (°)	1.33	-14.56 to 17.23
Shoulder resting position (°)	0.25	-4.62 to 5.12
45° scaption (°)	-0.15	-8.43 to 8.13
90° scaption (°)	-0.72	-10.76 to 9.33
135° scaption (°)	-1.00	-13.96 to 11.96
Maximal scaption (°)	1.48	-16.92 to 19.89
Upper trapezius (kg/cm ²)	0.12	-1.96 to 2.20
Supraspinatus (kg/cm ²)	0.23	-1.86 to 2.32
Infraspinatus (kg/cm ²)	0.16	-2.06 to 2.37
Serratus anterior (kg/cm ²)	0.05	-1.73 to 1.83
Pectoralis major (kg/cm ²)	0.13	-1.17 to 1.44
Pectoralis minor (kg/cm ²)	-0.09	-1.24 to 1.06

Mean difference and 95% limits of agreement are given. Measurements of the first cohort were taken, except for forward flexion and scapular upward rotation.

Accepted Manuscript