Pain Medicine 2017; 0: 1–12
doi: 10.1093/pm/pnx315

OXFORD

Original Research Article

Validity and Reliability of Clinical Examination in the Diagnosis of Myofascial Pain Syndrome and Myofascial Trigger Points in Upper Quarter Muscles

Orlando Mayoral del Moral, PT, PhD,*,† María Torres Lacomba, PhD,† I. Jon Russell, MD, PhD, ACR Master,‡ Óscar Sánchez Méndez, PT, MSc,* and Beatriz Sánchez Sánchez, PT, PhD*

*Physical Therapy Department, Provincial Hospital, Toledo, Spain; †Physiotherapy in Women's Health Research Group, Department of Physical Therapy, University of Alcalá, Alcalá de Henares, Madrid, Spain; ‡Fibromyalgia Research and Consulting, San Antonio, Texas, USA

Correspondence to: María Torres-Lacomba, Physiotherapy in Women's Health Research Group. Department of Physical Therapy, University of Alcalá. Alcalá de Henares, Madrid, Spain. Tel & Fax: 34 918854828, E-mail: maria.torres@uah.es.

Conflicts of interest: None declared.

Abstract

Objectives. To determine whether two independent examiners can agree on a diagnosis of myofascial pain syndrome (MPS). To evaluate interexaminer reliability in identifying myofascial trigger points in upper quarter muscles. To evaluate the reliability of clinical diagnostic criteria for the diagnosis of MPS. To evaluate the validity of clinical diagnostic criteria for the diagnosis of MPS.

Design. Validity and reliability study.

Setting. Provincial Hospital. Toledo, Spain.

Participants. Twenty myofascial pain syndrome patients and 20 healthy, normal control subjects, enrolled by a trained and experienced examiner.

Methods. Ten bilateral muscles from the upper quarter were evaluated by two experienced examiners. The second examiner was blinded to the diagnosis group. The MPS diagnosis required at least one muscle to have an active myofascial trigger point. Three to four days separated the two examinations. The primary outcome measure was the frequency with which the two examiners agreed on the classification of the subjects as patients or as healthy controls. The kappa statistic (K) was used to determine the level of agreement between both examinations, interpreted as very good (0.81–1.00), good (0.61–0.80), moderate (0.41–0.60), fair (0.21–0.40), or poor ($\leq$0.20).

Results. Interexaminer reliability for identifying subjects with MPS was very good (K = 1.0). Interexaminer reliability for identifying muscles leading to a diagnosis of MPS was also very good (K = 0.81). Sensitivity and specificity showed high values for most examination tests in all muscles, which confirms the validity of clinical diagnostic criteria in the diagnosis of MPS.

Conclusions. Interrater reliability between two expert examiners identifying subjects with MPS involving upper quarter muscles exhibited substantial agreement. These results suggest that clinical criteria can be valid and reliable in the diagnosis of this condition.

Key Words. Myofascial Pain Syndromes; Diagnostic Criteria; Validity; Reliability; Clinical Examination; Palpation; Myofascial Trigger Point

Introduction

Myofascial pain syndrome (MPS) is a clinical condition described as a combination of sensory, motor, and autonomic signs and symptoms caused by myofascial trigger points (MTrPs). An MTrP is defined as a

hyperirritable spot in a skeletal muscle that is associated with a hypersensitive palpable nodule within a taut band [1]. The spot is painful on compression and can refer pain or other sensory symptoms, such as paresthesia, to a zone of reference. It can also cause motor dysfunction and autonomic phenomena in the region of reference [1]. According to their clinical manifestations, MTrPs are usually classified as active when able to cause spontaneous sensory symptoms, usually pain, or latent when clinically silent regarding pain [1]. While both can create motor and autonomic dysfunction, only the active MTrPs can cause pain or other sensory symptoms, and the referred symptoms elicited by its compression are usually recognized by the patient as his/her characteristic complaint [1]. Both active and latent MTrPs can be present in an MPS, but at least one active MTrP in at least one skeletal muscle is required to make a diagnosis of MPS, and this diagnosis should identify the specific MTrPs that are relevantly contributing to that MPS [1]. This leads us to the use of the concept of "relevant MTrPs" as those, either active or latent, that contribute to a specific MPS [2].

The apparent prevalence of the MPS has been studied in populations with pain [3–5] and among people who suffer from MPS concomitant with other pathologies or diseases [6,7]. It is estimated that 93% of patients attending pain clinics [3] are suffering from MPS. Other studies show that 100% of the patients suffering from shoulder pain [8] or chronic nonspecific neck pain [7] meet criteria for MPS. Unfortunately, the limited number of studies on MPS prevalence, the heterogeneity of the samples in these studies, and the lack of consensus about the diagnostic criteria [9] make it difficult to compare such epidemiological data.

One of the most controversial aspects of the MPS still relates to the recognition of MTrPs: The field still awaits clinical diagnostic criteria that can be validated against a gold standard, such us needle EMG examination of the MTrP zone [10]. For this reason, agreement is the primary tool currently used to judge the accuracy of clinical diagnosis in this area. Several studies have assessed the interexaminer reliability on different body locations [11–14], leading to the conclusion that both training [11,13] and experience [12,14] are needed to achieve an acceptable interrater reliability. In 2011, a systematic review concluded that the published studies did provide satisfactory reliability for some of the characteristics of MTrPs [15]. However, a recently published systematic review regarding the interrater agreement for the diagnosis of MPS [16] shows agreement with earlier reviews published in 2008 [17] and 2009 [18] and concludes that palpation is a tool that presents moderate reliability for the clinical diagnosis of MPS. All reviews called for an increase in both the quantity and the quality of research regarding the clinical diagnosis of MPS [15–18].

While the diagnosis of MTrPs is academically important, the diagnosis of the MPS is clinically critical to the management of the suffering patient. Unfortunately, this aspect has not been investigated enough, and high-quality studies that assess whether clinical examination is capable of discriminating a healthy individual from an MPS patient are scarce and have great limitations regarding the relevant MTrPs contributing to the symptoms of a specific MPS [19,20].

The aims of this study were 1) to determine whether two independent examiners, one of them blinded to the diagnosis group, can agree on a diagnosis of MPS, 2) to evaluate interexaminer reliability in identifying relevant MTrPs in upper quarter muscles, 3) to evaluate the reliability of specific clinical criteria for the diagnosis of MPS, and 4) to evaluate the validity of clinical diagnostic criteria for the diagnosis of MPS.

Methods

Participants

Subjects were recruited from the Physical Therapy Department of Provincial Hospital, Toledo, Spain, between March 2003 and November 2004. Twenty subjects with MPS and an equal number of healthy normal control subjects were recruited by a physical therapist (OMM) expert in the diagnosis and treatment of MPS.

The MPS expert identified each subject to be within one of the two diagnostic groups MPS or healthy normal control (HNC). Once the diagnosis was established, the subject remained in that diagnosis group for the duration of the study. The inclusion criteria for the MPS subjects were MPS with at least one active MTrP in at least one of the 10 study-designated muscles (splenius capitis, sternocleidomastoid, upper trapezius, levator scapulae, infraspinatus, supraspinatus, anterior deltoid, latissimus dorsi, teres major, and pectoralis major); a history of regional pain persisting and stable for at least two weeks; subject's pain drawing showing regional pain involving the head, neck, and/or shoulder girdle in at least one side of the body; age 18 to 80 years; and not meeting American College of Rheumatology diagnostic criteria for fibromyalgia syndrome (FMS) [21]. Study subjects were excluded if their pain drawing showed pain in the low back, abdomen, or lower extremities or if they had a rheumatic disease, a major arthritic disorder, a recognized pain syndrome such as FMS or chronic fatigue syndrome, a serious illness involving any organ other than the musculoskeletal system, a clinically apparent affective disorder, an untreated endocrine or metabolic disorder, a diagnosis of malignancy within the past five years, use of opioid drugs within two weeks prior to either of the two study examinations, or an unwillingness to cooperate fully with the study protocol procedures.

The inclusion criteria for the HNC subjects were no troublesome pain problem of any kind in the prior two months, pain diagrams with no areas of painful symptoms, and no active MTrPs in any of the 10 examined muscles.

Healthy subjects were excluded if they had used pain relievers within two weeks prior to either of the study examinations or if they were unwilling to cooperate fully with the study protocol procedures.

Design

A validity and reliability study was carried out using a case-control design. A nonprobabilistic consecutive sampling design was used to select participants. The study was approved by the Ethics and Clinical Investigation Committee of the "Complejo Hospitalario de Toledo" (protocol number 82/2002), and written informed consent was obtained from all participants.

Assessment Procedure

The team responsible for implementing the procedure was composed of a first examiner, the MPS expert (OMM), and a blinded examiner (MTL) trained to recognize MTrPs and with more than five years' experience in performing the maneuvers required for the study examination protocol. Both examiners were physical therapists. It was decided that the blinded examiner should not be a member of the staff of the hospital where the study was being carried out as this could have compromised his/her blinding. For this reason, the selected blinded examiner came from another town (Madrid) to the study setting for every recruitment session, which helped to guarantee proper blinding.

Before starting the recruitment of subjects for the study, the two examiners performed training sessions in which they evaluated seven patients with MPS and seven asymptomatic subjects using the same procedures to be used during the study, and reconciling discrepancies. Once the training period was completed, the first examiner began to enroll MPS and HNC subjects. The blinded examiner was not informed of the diagnostic group classification for any study subject before completion of the entire study.

Given the expertise of the first examiner in the diagnosis of MPS and the unblinded nature of his examinations, the results of his classification of subjects were considered the reference standard [22] for the diagnosis of MPS, with which the results of the second examiner would have to be compared in terms of validity of the results.

Each subject was asked to complete a series of self-report questionnaires prior to physical examination. These questionnaires collected demographic data, including age, race, gender, ethnicity, work status, educational level, information regarding medical history, visual analog scale (VAS) for pain [23], SF-36 Health Survey [24] for health-related quality of life, and body pain diagrams. These self-report questionnaires were administered to each subject by the first examiner on two separate occasions, shortly before each of the two examinations.

Three to four days after the examination by the first examiner, the blinded examiner assessed each subject using the same type of standard case report form as used by the first examiner for recording the results of the examination. This case report form listed each of the 10 muscles to be examined and each of the maneuvers included in the examination (Supplementary Data). The 10 muscles were examined bilaterally. The blinded examiner was not allowed to ask the subject whether there was a personal "history of pain" or whether referral reproduced a "familiar pain." This obviously meant that the blinded examiner couldn't be certain of the active or latent status of the MTrPs that she found. Instead, she was asked to estimate whether the different MTrPs found in her examination could be "relevant" for the diagnosis of MPS, in which case the muscles harboring them would be termed "MPS muscles" in the evaluation of the results.

Regarding the pressure pain threshold (PPT) test, it was recorded in kg/cm^2 using an analog algometer (Wagner Instruments, Greenwich, USA) with a hard rubber surface area at the flat tip of 1.0 cm^2 (25). The pressure of compression was increased gradually at a rate of approximately 1 $kg/cm^2/s$. To standardize the rate of application, both examiners practiced, in the training phase prior to the study, increasing the pressure linearly to 5 kg/cm^2 over five seconds according to the method recommended by Fischer [25,26]. The participants were asked to say "yes" as soon as pain or discomfort appeared, and immediately the compression was stopped. At each disclosed MTrP, three repetitive measurements were recorded, with an interval of 30 seconds between each of the measurements. The highest reading was discarded, and the mean of the two remaining readings was used in the statistical analysis [1].

Sample Size Estimation

The sample size was estimated by calculation for the primary objective. The specifications were considering two examiners and a dichotomous variable (presence and absence of MPS) in order to detect a kappa value of 0.6, a one-tailed test with an alpha level of 0.05, and power of 80%. Therefore, a total of 40 participants (20 HNC subjects and 20 MPS subjects) were recruited [27].

Statistical Analysis

The primary outcome variable for this study was the frequency with which the diagnosis of the blinded examiner agreed with the diagnosis recorded by the first examiner. The secondary outcomes measures were the levels of agreement between the first examiner and the blinded examiner regarding 1) the presence of MTrPs in each of the 10 study muscles and 2) the outcomes of each of the examination tests performed in each muscle.

Qualitative data were represented by absolute and relative frequencies, and quantitative data were calculated by computing the mean and standard deviation or median and interquartile range (IQR) of each variable, for parametric and nonparametric data, respectively. The Shapiro Wilks test was used to determine normality.

Descriptive statistics were provided to describe whether the two groups (MPS and HNC) differed with regard to demographic variables and baseline characteristics, and whether the two groups differed on repeated testing 1) before the first examination and 2) before the blinded examiner's assessment. The Fisher exact test was used for qualitative dichotomous variables, and the Student *t* test was used for quantitative variables.

Analyses of agreement on MPS diagnosis for each of the 10 muscles and on clinical diagnostic criteria were performed using Cohen's kappa coefficient (K) and its 95% confidence interval. The K statistic is a chance-corrected measure of agreement. In addition to examining the proportion of observed agreements, K also considers the proportion of agreements expected by chance. Kappa values were calculated by a specific muscle and side of the body, and by nine of the 10 examination maneuvers in a specific muscle and side of the body. The degree of agreement was determined following the criteria proposed by Altman [28] as very good (0.81–1.00), good (0.61–0.80), moderate (0.41–0.60), fair (0.21–0.40), or poor ($\leq$0.20). The intraclass correlation coefficient (ICC) was calculated for PPT at each side of the body of each specific muscle. According to Fleiss [29], ICC values above 0.75 generally mean "excellent" reliability.

Sensitivity and specificity values of both examiners were calculated in order to analyze validity. For this aim, we considered the results obtained by the first unblinded examiner as the reference standard [22] of the MPS/MTrP diagnosis.

The analyses were carried out using the Statistical Package for the Social Sciences software (17.0 version; SPSS Inc., Chicago, IL, USA). A P value of < 0.05 was considered statistically significant.

Results

Forty participants were included in the study, 20 in the MPS group and 20 in the HNC group, as designated by the first examiner. Table 1 compares the demographics and other characteristics of the subjects in both diagnosis groups. The main significant differences ($P < 0.001$) between MPS and HNC subjects were the VAS pain score and the duration of MPS. As these are indicators of MPS, such differences were expected between the two groups.

In addition, MPS subjects exhibited a poorer quality of life than HNC subjects as all of the questionnaires showed statistically significant differences ($P < 0.001$) between the two groups (Table 2).

Table 2 also shows that there were no significant changes in any of the two groups between the first and the diagnosis-blinded assessments.

Agreement on the Diagnosis of MPS

Agreement on the primary outcome variable between the first and the blinded examiners was very good (K = 1, agreement = 100%) for identifying MPS subjects.

Interexaminer Reliability in Identifying Myofascial Trigger Points in Upper Quarter Muscles

Table 3 shows the differences for each of the examination tests in each group of subjects found by the blinded examiner. There were very significant differences between MPS and HNC subjects for all but three of the tests: palpable taut band ($P = 0.09$), local twitch response ($P = 0.214$), and matchstick test ($P = 0.47$).

Regarding PPT, the mean value obtained by the blinded examiner for muscles in the healthy side of MPS subjects with unilateral involvement was 3.02 (0.42). The mean PPT value for muscles in the involved side was 1.59 (0.15). Thus, the difference in the PPT scores between the healthy and the involved side in MPS subjects with unilateral involvement was 1.43 kg/cm^2.

Agreement between the first and the blinded examiners for one of the secondary outcome variables, namely identifying MPS muscles, was very good (K = 0.81, agreement = 81%).

MPS muscles' agreement between the first and blinded examiners is shown in Table 4. The muscles with the better results were supraspinatus (K = 1), anterior deltoid (K = 0.92), sternocleidomastoid (K = 0.96), levator scapulae (K = 0.88), latissimus dorsi (K = 0.77), and infraspinatus (K = 0.77).

Reliability of Clinical Diagnostic Criteria for the Diagnosis of Myofascial Pain Syndrome

Agreement between the first and blinded examiners for each muscle and each examination maneuver is shown in Table 5. All but two of the maneuvers (matchstick test K = 0.20; skin rolling test K = 0.30) showed good ($0.61 < K < 0.80$) or very good ($K > 0.81$) agreement.

The percentage of agreement between the two evaluators for each muscle and each examination maneuver exceeds 70% in all cases, meaning an optimal percentage for a good reliability.

Table 1 Demographics and subject characteristics

Demographics	MPS Subjects (N = 20)	HNC Subjects (N =20)	Total Sample (N = 40)	*P* Value
Gender, No. (%)				
Male	6 (30)	10 (50)	16 (40)	NS*
Female	14 (70)	10 (50)	24 (60)	NS*
Age $\bar{x}$ (SD), y	30.1 (8.8)	30.1 (10)	30 (9)	NS*
Ethnicity: white, No. (%)	20 (100)	20 (100)	40 (100)	NS*
Duration of MPS, mo	1.9 (0.4)	—	1.9 (0.4)	NS*
Current VAS pain median (IQR), mm	40.1 (33.2)	0 (2)	40.1 (35.2)	<0.001
Education $\bar{x}$ (SD), y	13.5 (0.6)	12.9 (0.9)	13.2 (0.7)	NS*
Employed, No. (%)	17 (85)	18 (90)	35 (87.5)	NS*

HNC = healthy normal control; MPS = myofascial pain syndrome; NS = not significant; VAS = visual analog scale.
*$P \geq 0.05$.

Table 2 Clinical characteristics of subjects in the first and in the blinded examinations

	First Assessment Day 1		Blinded Assessment Day 3–5		MPS Day 1 vs MPS Day 3–5	HNS Day 1 vs HNS Day 3–5
	MPS Subjects (N = 20)	HNC Subjects (N = 20)	MPS Subjects (N = 20)	HNC Subjects (N = 20)	*P* Value	*P* Value
Pain VAS, median (IQR) mm	48 (16.7)	0 (2)	48 (20.2)	0 (1)	NS*	NS*
SF-36 $\bar{x}$ (SD)						
SF-36 physical function	86.6 (9.1)	90.7 (8.3)	88.1 (10.2)	90.6 (8.1)	NS*	NS*
SF-36 role-physical	74.4 (11.7)	91.1 (8.1)	76.8 (11.2)	90.8 (7.3)	NS*	NS*
SF-36 pain index	74.1 (12.5)	92.6 (5.3)	74.3 (11.6)	93.1 (5.2)	NS*	NS*
SF-36 general health perception	70.5 (15.7)	80.1 (17.2)	68.33 (16.4)	80.6 (18.1)	NS*	NS*
SF-36 vitality	57.1 (16.8)	72.5 (18.2)	57 (15.9)	73 (18.6)	NS*	NS*
SF-36 social functioning	78.6 (16.6)	92 (6.3)	80.5 (15.5)	91 (6.8)	NS*	NS*
SF-36 role-emotional	72.6 (10.6)	90.5 (7.2)	72.2 (8.8)	91 (6.9)	NS*	NS*
SF-36 mental health index	67.8 (14.9)	78.9 (17.4)	68.3 (15.2)	79.1 (16.6)	NS*	NS*
Standardized physical component scale	51.4 (4.1)	56 (8)	51.6 (4.2)	57.7 (8.6)	NS*	NS*
Standardized mental component scale	47.5 (6.3)	54.5 (7.4)	47.7 (5.9)	53.5 (8)	NS*	NS*

HNC = healthy normal control; MPS = myofascial pain syndrome; NS = not significant; SF-36 = Short Form-36; VAS = visual analog scale.
*$P \geq 0.05$.

Validity of Clinical Diagnostic Criteria for the Diagnosis of Myofascial Pain Syndrome

Considering the first examiner as the reference standard [22] for the MPS/MTrP diagnosis, Table 6 shows sensitivity and specificity values from the blinded examiner for each muscle and each examination. High values were demonstrated for both sensitivity and specificity for each examination tests in all tested muscles.

Discussion

The main purpose of this study was to determine whether subjects suffering from MPS could be diagnosed and distinguished from HNC subjects using clinical examination. The study was designed as a validity and interrater reliability trial comparing the assessments performed by an experienced-but-blinded examiner with those performed by a first examiner identified as an MPS expert.

Table 3 Differences for each of the examination maneuvers in each group of subjects in blinded assessment

Examination Maneuvers (10 × 10 Muscles × Both Sides = 200), No. (%)	MPS Subjects (N = 20) (%)	HNC Subjects (N = 20) (%)	*P* Values
Painfully restricted passive range of motion observed	81 (2)	4 (0.1)	<0.0001*
Muscle strength limited by pain	85 (2)	1 (0)	<0.0001*
Palpable taut band	360 (9)	318 (8)	0.09*
Area of spot tenderness	378 (9.5)	306 (7.6)	0.0045*
Jump sign	277 (7)	142 (3.6)	<0.0001*
Pain referral reported by the subject	225 (5.6)	72 (2)	<0.0001*
Local twitch response	270 (6.7)	242 (6)	0.214*
Pressure pain threshold $\bar{x}$ (SD)	2.12 (0.42)	3.83 (0.21)	<0.0001†
Matchstick test positive	2 (0.05)	0 (0)	0.47*
Skin rolling test positive	27 (0.7)	0 (0)	<0.0001*

$\bar{x}$ (SD) = mean (standard deviation).
**P* values obtained by $\chi 2$.
†*P* values obtained by Student *t* test.

Table 4 Agreement between first and blinded examiners for identifying MPS muscles

Muscle	First Exam. MPS Dx	Blinded Exam. Frequency, No. (%)	K Value for Right Side (95% CI)*	First Exam. MPS Dx	Blinded Exam. Frequency, No. (%)	K Value for Left Side (95% CI)*
Splenius capitis	10	6 (60)	0.69 (0.42–0.96)	7	2 (28.6)	0.55 (0.18–0.92)
Sternocleidomastoid	9	9 (100)	1.0 (1.00–1.00)	9	9 (100)	0.93 (0.80–1.00)
Upper trapezius	12	9 (75)	0.70 (0.45–0.94)	9	7 (77.7)	0.78 (0.54–1.00)
Levator scapulae	9	8 (88.8)	0.93 (0.78–1.00)	9	7 (77.7)	0.84 (0.64–1.00)
Infraspinatus	4	4 (100)	0.77 (0.47–1.00)	4	4 (100)	0.77 (0.47–1.00)
Supraspinatus	3	3 (100)	1.0 (1.00–1.00)	1	1 (100)	1.0 (1.00–1.00)
Anterior deltoid	4	3 (75)	0.84 (0.55–1.00)	4	4 (100)	1.0 (0.54–1.00)
Latissimus dorsi	4	2 (50)	0.77 (0.48–0.90)	2	1 (50)	0.78 (0.47–0.90)
Teres major	1	0 (0)	0.68 (0.47–1.00)	3	2 (66.6)	0.79 (0.38–1.00)
Pectoralis major	3	1 (33.3)	0.64 (0.47–1.00)	1	0 (0)	0.77 (0.29–1.00)

Dx = diagnosis including both sides; Frequency = frequency of agreement with first examination; MPS = myofascial pain syndrome.
**P* < 0.001.

The study also evaluated the interrater reliability of the diagnosis of the selected muscles and the reliability of each maneuver performed on each muscle.

The results show a very good reliability (K = 1, agreement = 100%) for distinguishing HNC subjects from MPS subjects. Two previous studies have tried to do something similar. Tunks et al. [19], using pressure algometry and digital palpation in a blinded fashion, could distinguish patients with MPS or FMS from HNCs, but could not discriminate between patients with these two conditions or the body region affected. Gerber et al. [20] could successfully discriminate between patients with chronic cervical pain and healthy subjects by means of physical examination (range of motion measurement, a 10-point manual muscle test, and manual and algometric palpation) and the use of a verbal analog scale (0–10) to rate pain, as well as the Brief Pain Inventory and Oswestry Disability Scale (with a sleep subscale), the Short-Form 36 Health Survey, and the Profile of Mood States. The results of the current study add to these and confirm the capability of clinical examination to distinguish between MPS patients and HNCs. In the current study, questionnaires were

Table 5 Interexaminer reliability for each of the examination tests in each muscle

		Examination Tests*																	
		ROM		Strength Limited by Pain		Taut Band		Spot Tenderness		Jump Sign		Pain Referral		LTR		PPT	Matchstick Test		Skin Rolling Test	
Muscles		Agreement, %	K Values (95% CI)	Agreement, %	K Values (95% CI)	Agreement, %	K Values (95% CI)	Agreement, %	K Values (95% CI)	Agreement, %	K Values (95% CI)	Agreement, %	K Values (95% CI)	Agreement, %	K Values (95% CI)	K Values (95% CI)	Agreement, %	K Values (95% CI)	Agreement, %	K Values (95% CI)
Splenius capitis	Right	92.50	0.53 (0.07–0.99)	97.50	0.79 (0.39–1.00)	95.00	0.87 (0.70–1.00)	97.50	0.90 (0.73–1.00)	87.50	0.74 (0.52–0.95)	82.50	0.59 (0.33–0.86)	97.50	—	0.96 (0.94–0.98)	—	0 (0.00–0.00)	97.50	0.65 (0.03–1.00)
	Left	97.50	0.53 (0.07–0.99)	97.50	0.79 (0.39–1.00)	92.50	0.81 (0.62–1.00)	92.50	0.73 (0.45–1.00)	95.00	0.88 (0.73–1.00)	82.50	0.56 (0.30–0.83)	—	—	0.95 (0.92–0.98)	—	0 (0.00–0.00)	95.00	0.02 (0.01–0.03)
Sternocleido-mastoid	Right	95.00	0.84 (0.63–1.00)	97.50	0.93 (0.79–1.00)	100.00	1.00 (1.00–1.00)	97.50	0.65 (0.03–1.00)	90.00	0.79 (0.60–0.98)	85.00	0.69 (0.48–0.91)	92.50	0.53 (0.07–0.99)	0.85 (0.76–0.93)	—	0 (0.00–0.00)	95.00	0.01 (0.00–0.03)
	Left	95.00	0.77 (0.47–1.00)	92.50	0.72 (0.44–1.00)	100.00	1.00 (1.00–1.00)	92.50	0.36 (0.00–0.92)	92.50	0.84 (0.60–0.98)	80.00	0.59 (0.35–0.84)	92. 50	0.68 (0.35–1.00)	0.91 (0.84–0.96)	—	0 (0.00–0.00)	95.00	0.01 (0.00–0.03)
Upper trapezius	Right	9000	0.73 (0.49–0.97)	71.00	0.83 (0.60–1.00)	97.50	1.00 (1.00–1.00)	—	1.00 (1.00–1.00)	82.50	0.60 (0.35–0.87)	87.50	0.74 (0.54–0.95)	87.50	0.40 (0.01–0.80)	0.91 (0.84–0.96)	90.00	1.00 (1.00–1.00)	92.50	0.53 (0.07–0.99)
	Left	95.00	0.86 (0.66–1.00)	95.00	0.77 (0.48–1.00)	97.50	1.00 (1.00–1.00)	100.00	1.00 (1.00–1.00)	85.00	0.67 (0.43–0.91)	82.50	0.65 (0.41–0.88)	92.50	0.54 (0.10–0.98)	0.91 (0.86–0.96)	92.50	1.00 (1.00–1.00)	85.00	0.35 (0.00–0.71)
Levator scapulae	Right	92.50	0.75 (0.49–1.00)	97.50	0.65 (0.03–1.00)	97.50	0.87 (0.64–1.00)	90.00	0.78 (0.58–0.98)	92.50	0.82 (0.63–1.00)	95.00	0.88 (0.74–1.00)	85.00	0.63 (0.37–0.89)	0.91 (0.86–0.96)	95.00	1.00 (1.00–1.00)	92.50	0 (0.00–0.00)
	Left	97.50	0.92 (0.76–1.00)	100.00	1.00 (1.00–1.00)	97.50	0.79 (0.39–1.00)	92.50	0.84 (0.67–1.00)	92.50	0.81 (0.62–1.00)	97.50	0.94 (0.82–1.00)	—	0.54 (0.26–0.82)	0.94 (0.91–0.97)	95.00	1.00 (1.00–1.00)	95.00	0 (0.00–0.00)
Infraspinatus	Right	80.00	0.56 (0.32–0.81)	92.50	0.62 (0.24–1.00)	97.50	0.79 (0.39–1.00)	95.00	0.64 (0.18–1.00)	92.50	0.85 (0.68–1.00)	87.50	0.75 (0.55–0.95)	75.00	0.39 (0.10–0.68)	0.92 (0.88–0.97)	90.00	0.03 (0.00–0.06)	90.00	0.46 (0.04–0.88)
	Left	92.50	0.79 (0.58–1.0)	97. 50	0.87 (0.64–1.0)	100.00	1.0 (1.0–1.0)	100.00	1.0 (1.0–1.0)	95.00	0.90 (0.76–1.00)	90.00	0.79 (0.60–0.98)	75.00	0.32 (0.03–0.62)	0.94 (0.90–0.97)	90.00	0.03 (0.00–0.06)	92.50	0.37 (0.00–0.90)
Supraspinatus	Right	97.50	0.84 (0.55–1.00)	95.00	0.77 (0.47–1.00)	92.50	0.54 (0.50–1.00)	97.50	0.94 (0.83–1.0)	97.50	0.80 (0.64–1.00)	85.00	0.64 (0.39–0.90)	—	—	0.88 (0.82–0.95)	90.00	0 (0.00–0.00)	87.50	0.39 (0.00–0.79)
	Left	—	0.88 (0.74–1.00)	90.00	0.54 (0.10–0.98)	92.50	0.75 (0.49–1.00)	97.50	0.94 (0.83–1.00)	95.00	0.81 (0.55–1.00)	90.00	0.73 (0.50–0.97)	—	—	0.84 (0.75–0.93)	95.00	0 (0.00–0.00)	85.00	0.34 (0.00–0.71)
Anterior deltoid	Right	97.50	0.84 (0.55–1.0)	—	0.75 (0.49–1.00)	—	0.85 (0.68–1.00)	97.50	0.84 (0.55–1.00)	87.50	0.75 (0.49–1.00)	85.00	0.67 (0.44–0.91)	100.00	1.00 (1.00–1.00)	0.94 (0.91–0.98)	—	0 (0.00–0.00)	95.00	0.48 (0.00–1.00)
	Left	100.00	1.00 (1.00–1.00)	100.00	1.00 (1.00–1.00)	97.50	0.75 (0.49–1.00)	97.50	0.65 (0.03–1.00)	87.50	0.75 (0.49–1.00)	92.50	0.83 (0.66–1.00)	100.00	1.00 (1.00–1.00)	0.96 (0.94–0.99)	97.50	0 (0.00–0.00)	95.00	0.02 (0.01–0.03)
Latissimus dorsi	Right	100.00	1.00 (1.00–1.00)	92.50	0.36 (0.00–0.92)	100.00	1.00 (1.00–1.00)	100.00	1.00 (1.00–1.00)	90.00	0.77 (0.56–0.98)	95.00	0.88 (0.74–1.00)	97.50	0.84 (0.55–1.00)	0.87 (0.80–0.95)	—	0 (0.00–0.00)	97.50	0 (0.00–0.00)
	Left	—	1.00 (1.00–1.00)	95.00	0.64 (0.18–1.00)	97.50	0.78 (0.58–0.98)	100.00	1.00 (1.00–1.00)	92.50	0.81 (0.62–1.00)	92.50	0.84 (0.63–1.00)	95.00	0.47 (0.00–1.00)	0.84 (0.74–0.93)	—	0 (0.00–0.00)	97.50	0 (0.00–0.00)
Teres major	Right	100.00	1.00 (1.00–1.00)	100.00	1.00 (1.00–1.00)	90.00	0.69 (0.41–0.97)	100.00	1.00 (1.00–1.00)	97.50	0.95 (0.85–1.00)	92.50	0.84 (0.66–1.00)	82.50	0.36 (0.00–0.74)	0.94 (0.91–0.98)	97.50	0 (0.00–0.00)	97.50	0 (0.00–0.00)
	Left	97.50	0.65 (0.03–1.00)	95.00	0.72 (0.37–1.00)	100.00	1.00 (1.00–1.00)	97.50	0.90 (0.73–1.00)	95.00	0.89 (0.76–1.00)	100.00	1.00 (1.00–1.00)	90.00	0.65 (0.34–0.96)	0.96 (0.94–0.98)	97.50	0 (0.00–0.00)	97.50	0 (0.00–0.00)
Pectoral major	Right	97.50	0.84 (0.55–1.00)	—	1.00 (1.00–1.00)	100.00	1.00 (1.00–1.00)	97.50	0.79 (0.38–1.00)	92.50	0.82 (0.62–1.00)	90.00	0.79 (0.60–0.98)	95.00	0.83 (0.60–1.00)	0.96 (0.94–0.98)	—	0 (0.00–0.00)	92.50	0.54 (0.10–0.98)
	Left	97.50	——	—	———	—	—	97.50	0.79 (0.38–1.00)	95.00	0.87 (0.69–1.00)	85.00	0.67 (0.44–0.91)	95.00	0.84 (0.55–1.00)	0.95 (0.93–0.98)	92.50	0 (0.00–0.00)	87.50	0.54 (0.10–0.98)
Average values			0.80		0.76		0.87		0.84		0.80		0.71		0.62	0.92		0.20		0.30

All values are expressed in K values, and their 95% confidence intervals, except the PPTs, are expressed in intraclass correlation coefficients. All values are expressed in %.

ICC = intraclass correlation coefficient; LTR = local twitch response; PPT = pressure pain threshold value; ROM = painfully restricted passive range of motion.

*$P < 0.001$.

Table 6 Sensitivity and specificity for each of the applied tests on each examined muscle

		Examination Tests								
		Sensitivity/Specificity								
Muscles		Rom	Strength Limited by Pain	Taut Band	Spot Tenderness	Jump Sign	Pain Referral	LTR	Matchstick Test	Skin Rolling Test
Splenius capitis	Right	66.67/94.59	66.67/100	93.33/100	100/85.71	92.86/84.62	81.82/82.76	NED/100	NED/100	100/97.44
	Left	NED/100	66.67/100	96.43/83.33	100/62.5	100/92.86	100/78.79	NED/100	NED/100	NED/97.44
Sternocleidomastoid	Right	87.5/96.88	90/100	100/100	100/97.44	95.65/82.35	95.24/73.68	66.67/94.59	NED/100	NED/100
	Left	80/97.10	71.43/96.97	100/100	97.30/33.33	92.31/92.86	80.95/78.95	94.29/80	NED/100	NED/97.44
Upper trapezius	Right	72.73/96.55	75/100	97.50/NED	100/NED	85.19/76.92	84/93.33	86.84/100	NED/100	66.67/94.59
	Left	100/93.75	100/94.44	NED/100	100/100	85.19/84.62	80.95/84.21	92.11/100	NED/95	84.21/100
Levator scapulae	Right	75/96.88	50/100	97.22/100	92.31/85.71	76.92/100	92.86/100	96.3/61.54	NED/95	NED/94.87
	Left	87.50/100	100/100	97.31/100	100/81.25	90.91/93.10	100/96.55	100/46.15	NED/100	NED/94.87
Infraspinatus	Right	100/74.19	75/94.44	100/66.67	97.30/66.67	95.24/89.47	94.12/82.61	43.86/92.31	NED/92.31	100/89.47
	Left	100/90.63	80/100	100/100	100/100	95/95	93.33/88	30.77/96.30	NED/92.31	100/92.31
Supraspinatus	Right	100/97.3	100/94.44	66.67/100	100/92.31	100/97.2	81.82/86.21	NED/100	NED/95	100/86.84
	Left	NED/100	100/89.19	40/100	100/92.86	100/94.29	100/87.50	NED/100	NED/97.44	100/84.21
Anterior deltoid	Right	75/100	NED/100	100/NED	97.3/100	100/77.27	91.67/82.14	100/100	NED/NED	33.33/100
	Left	100/100	100/100	100/NED	100/50	100/73.68	100/89.29	100/100	NED/100	NED/97.44
Latissimus dorsi	Right	100/100	50/94.74	100/100	100/100	96.15/78.87	85.71/100	100/75	NED/100	NED/100
	Left	NED/100	66.67/97.30	100/NED	100/100	93.10/90.91	93.33/92.00	97.37/50.00	NED/100	NED/100
Teres major	Right	100/100	100/100	90.91/85.71	100/100	100/87.5	92.31/92.86	90.91/42.86	NED/100	NED/100
	Left	100/97.44	60/100	100/100	100/100	95.83/93.75	100/100	93.94/71.43	NED/100	NED/100
Pectoralis major	Right	100/97.3	NED/100	100/100	100/66.67	93.10/90.91	81.25/95.83	94.12/100	NED/100	NED/100
	Left	NED/100	NED/100	NED/100	100/66.67	100/81.82	78.57/88.46	96.88/87.50	NED/97.44	NED/95

All values are expressed in %.
LTR = local twitch response; NED = not enough data to calculate the value; ROM = painfully restricted passive range of motion.

used, but the clinical decision to classify subjects was taken in a blinded fashion only by the use of physical examination. Despite the great technical advances in imaging techniques [30–35], the clinical diagnosis of MPS is still based mainly on physical examination [36], which confers relevance to these studies.

The current study also evaluated the interrater reliability of the diagnosis of MTrPs in different muscles. The agreement between the two examiners was globally very good ($K = 0.81$, agreement = 81%). Systematic reviews have established that the interrrater reliability in diagnosing MTrPs is, at its best, moderate ($0.41 < K < 0.60$) [16], but the 10 muscles included in our study scored higher than that, the lowest being the splenius capitis muscle, whose reliability reached a kappa value of 0.62 (Table 4). One of the possible explanations for this disparity is the very thorough systematic examinations performed in the current study, using a combination of 10 diagnostic tests. Most of the studies included in previous reviews used a smaller combination of the diagnostic criteria proposed by Simons et al. [1] or did not follow the recommended algometric measuring procedures [1,25,26].

The current results show that most of the diagnostic tests used in this study disclosed a very significant difference between MPS and HNC subjects (Table 3). This observation suggests an explanation for why the blinded examiner achieved such excellent results when deciding group allocation. Only three criteria failed to show a significant difference between MPS and HNC groups: identification of a taut band ($P = 0.09$), LTR ($P = 0.214$), and the matchstick test ($P = 0.47$) (Table 3).

The matchstick test (Supplementary Data) showed no power to discriminate between diagnosis groups in the present study. Actually, this test is not directly related to MTrPs but to changes in skin associated with nerve entrapment, as suggested by Gunn [37]. In this regard, it is no surprise that it fails to be helpful in the diagnosis of MTrPs.

On the other hand, the reported presence of latent MTrPs in scapular muscles in almost 90% of the healthy population [38] could account for the small differences found between HNC and MPS subjects regarding taut band and LTR.

Differences between MPS patients and HNCs seem to be mostly related to muscle mechano-sensitivity caused by the proven sensitization existing in MTrPs [39,40] and revealed by the mechanical deformations imposed by maneuvers such as pressure, contraction, or stretch. Previous studies have documented the importance of PPT to differentiate between patients and healthy subjects [19,20]. According to Fischer, a muscle PPT equaling 3 kg/cm^2 or less can be considered abnormally low [25]. In our study (Table 3), the mean PPT in MPS muscles was 2.12, with no muscle having a PPT higher than 3, while the mean PPT among HNC muscles was 3.83, with no PPTs lower than 3.

The mean difference in PPTs between MPS subjects and HNCs in the current study was 1.71 kg/cm^2. In the study by Gerber et al. [20], this difference was 1.41 kg/cm^2. A recent report by Calvo Lobo et al. [41] showed that the minimal clinically important difference in PPT for patients with nonspecific shoulder pain for anterior deltoid muscle is 1.17 kg/cm^2, which seems to correlate with the current results (differences in PPTs between MPS patients and HNCs for anterior deltoid were 1.3 kg/cm^2 for the first examiner and 1.4 kg/cm^2 for the blinded examiner) and reinforces the idea that algometric differences can be valid to discriminate between MPS patients and HNCs.

In the current study, the mean PPT difference between healthy and involved sides in MPS patients was 1.43 kg/cm^2. Other studies have found side-to-side differences of just 0.16 kg/cm^2 between the primary pain site and its contralateral control [42], although with a questionable way to select the contralateral control. This means that Fischer's statement that a side-to-side difference exceeding 2 kg/cm^2 can be considered abnormal, according to clinical experience [25], could be an overestimation, and as the current results show, smaller differences can be sufficiently relevant to differentiate between healthy and involved sides of an MPS patient. Future research is needed to clarify this issue.

The difference between PPT and referred pain threshold has been shown to be significantly different in active and latent MTrPs [43]. Referred pain threshold was not included in the study. Future reliability studies should include it in order to determine whether the difference with PPT could help to create an easier distinction between MPS and HNC subjects.

To our knowledge, the criterion "painfully restricted passive range of motion," considered an essential criterion by Simons et al. [1], had never been tested in any previous reliability study. Although the interrater reliability of this measure seems to be highly dependent on the muscle, ranging from $K = 0.53$ of the splenius capitis muscle to $K = 1.0$ for the left anterior deltoid, the right latissimus dorsi, and the right teres major muscles) (Table 5), it shows very significant global differences between its results in MPS and those in HNC subjects ($P < 0.0001$), which could indicate that it could be a useful criterion in the clinical diagnosis of MPS, at least in some muscles and, most likely, in nonhypermobile joints or patients.

In order to assess the validity of the clinical criteria in the diagnosis of MPS, the results obtained by the unblinded examination of the first examiner were considered the reference standard [22] to which the results of the blinded examiner were compared. In this regard, the sensitivity and the specificity of the results of the different diagnostics tests obtained by the blinded examiner

in the different muscles were analyzed. Briefly explained, in dichotomous variables, sensitivity refers to the probability of correctly classifying a real patient, while specificity refers to the probability of correctly classifying a healthy subject [22]. For the purposes of the current study, in which classifying both MPS patients and HNC subjects is equally important, both sensitivity and specificity should score high for the variable to be considered valid. In the current study, most of the tests scored very high in both characteristics, which accounts for a high validity of most tests, except for the matchstick test and the skin rolling test, whose sensitivity values could not be calculated in most of the muscles. The sensitivity/specificity results obtained in the current study confirm high-validity values for all but these two tests, which could account for the very good results in the discrimination between MPS and HNC subjects.

An interesting protocol detail is the time interval between first and second assessments. In most studies, this variable is usually either not specified, or is specified as not being longer than 10–15 minutes [16]. In the present study, the time interval between examinations was three to four days, which made it less likely that palpation by the first examiner would produce any changes, such as visible cutaneous signs or local alterations in soft tissue sensitivity, that might have risked blinding of the second evaluator. On the other hand, this contravenes the opinion of authors such as Rathbone et al. [16], who recommended not to exceed 24 hours between assessments as this could produce a variation on the interrater's agreement due to the risk of changes being produced in study subject's underlying characteristics. It is indeed interesting to observe that the clinical characteristics of the subjects did not vary in the time interval between assessments. Future studies should evaluate whether shorter intervals could lead to alterations on the interrater's agreement when diagnosing MPS.

Even though the combination of the diagnostic criteria used in this study allows us a better understanding of the possibility of properly diagnosing a subject suffering from MPS, it is worth mentioning that, when questioned, the blinded examiner admitted that, although in a very minor degree, her decision on some subjects was also influenced or reinforced by other factors, different from the used criteria, that is, body language or facial gestures (cautious movements, grimace, moan), usually referred to as "pain behavior" in the literature. To our knowledge, these criteria have never been used in studies about the diagnosis of MPS, despite being known and used in other fields, such as temporomandibular disorders [44–46] or low back pain [47–49], where they are integrated successfully with traditional diagnostic tests by expert evaluators [50]. It would be interesting in future studies to determine the extent to which the observation of pain behavior could be utilized by expert examiners in the diagnosis of MPS.

Limitations

Although very unlikely, and limited to the few final presenting subjects, the fact that the blinded examiner was aware of the number of subjects to be included in both groups could have unintentionally biased her judgments as the enrollment progressed, based on the number of subjects that she had previously classified.

The fact that most muscles included in the study were superficial could have biased the results in favor of reliability. Future studies in this region should also include deep muscles such as cervical multifidii, semiespinalis capitis, or longus colli, as well as the consideration of the body mass index of the subjects, so that validity and reliability could also be evaluated in deep muscles, or even in superficial muscles of obese subjects.

Conclusions

Interrater reliability between two examiners identifying MPS subjects with MTrPs in upper quarter muscles exhibited substantial agreement. These results suggest that clinical criteria can be valid and reliable in the diagnosis of MPS. Some muscles, mainly supraspinatus, sternocleidomastoid, anterior deltoid, levator scapulae, latissimus dorsi, and infraspinatus, were more reliably identified by the physical therapists in their assessments. Most of the examination maneuvers showed a high degree of agreement ($K \geq 0.71$) and very high sensitivity and specificity, which account for the validity of clinical criteria in discriminating between MPS and HNC subjects. Future studies will be required to confirm these findings, evaluating the validity of the clinical diagnostic criteria of MTrPs as compared with a gold standard such as needle electromyography, and in different body regions.

Supplementary Data

Supplementary Data may be found online at http://painmedicine.oxfordjournals.org.

References

1 Simons DG, Travell JG, Simons LS. Myofascial Pain and Dysfunction. The Trigger Point Manual. Upper Half of Body, 2nd edition. Baltimore: Williams & Wilkins; 1999.

2 Mayoral del Moral O. Diagnóstico y fisioterapia invasiva del síndrome de dolor miofascial y de los puntos gatillo miofasciales. Doctoral Dissertation. Alcalá University; 2017.

3 Gerwin RD. A study of 96 subjects examined both for fibromyalgia and myofascial pain. J Musculoskelet Pain 1995;3(suppl 1):121.

4 Fishbain DA, Goldberg M, Meagher BR, Steele R, Rosomoff H. Male and female chronic pain patients categorized by DSM-III psychiatric diagnostic criteria. Pain 1986;26(2):181–97.

5 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 Musculoskelet Disord 2011;12(1):139.

6 Mayoral O, Salvat I, Martín MT, et al. Efficacy of myofascial trigger point dry needling in the prevention of pain after total knee arthroplasty: A randomized, double-blinded, placebo-controlled trial. Evid Based Complement Alternat Med 2013;2013:8.

7 Cerezo-Tellez E, Torres-Lacomba M, Mayoral-Del Moral O, et al. Prevalence of myofascial pain syndrome in chronic non-specific neck pain: A population-based cross-sectional descriptive study. Pain Med 2016;17(12):2369–77.

8 Bron C, Wensing M, Franssen JL, Oostendorp RA. Treatment of myofascial trigger points in common shoulder disorders by physical therapy. A randomized controlled trial [ISRCTN75722066]. BMC Musculoskelet Disord 2007;8(1):107.

9 Tough EA, White AR, Richards S, Campbell J. Variability of criteria used to diagnose myofascial trigger point pain syndrome—evidence from a review of the literature. Clin J Pain 2007;23(3):278–86.

10 Simons DG. Neuromusculoskeletal medicine—dawning of a new day. J Man Manip Ther 2006;14(4):199–201.

11 Gerwin RD, Shannon S, Hong CZ, Hubbard D, Gevirtz R. Interrater reliability in myofascial trigger point examination. Pain 1997;69(1–2):65–73.

12 Hsieh CY, Hong CZ, Adams AH, et al. Interexaminer reliability of the palpation of trigger points in the trunk and lower limb muscles. Arch Phys Med Rehabil 2000;81(3):258–64.

13 Sciotti VM, Mittak VL, DiMarco L, et al. Clinical precision of myofascial trigger point location in the trapezius muscle. Pain 2001;93(3):259–66.

14 Mora-Relucio R, Nunez-Nagy S, Gallego-Izquierdo T, et al. Experienced versus inexperienced interexaminer reliability on location and classification of myofascial trigger point palpation to diagnose lateral epicondylalgia: An observational cross-sectional Study. Evid Based Complement Alternat Med 2016;2016:6059719.

15 McEvoy J, Huijbregts P. Reliability of myofascial trigger point palpation: A systematic review. In: Dommerholt J, Huijbregts P, eds. Myofascial Trigger Points. Pathophysiology and Evidence-Informed Diagnosis and Management. Sudbury, Canada: Jones and Bartlett Publishers; 2011:65–88.

16 Rathbone AT, Grosman-Rimon L, Kumbhare DA. Interrater agreement of manual palpation for identification of myofascial trigger points: A systematic review and meta-analysis. Clin J Pain 2017;33:715–29.

17 Myburgh C, Larsen AH, Hartvigsen J. A systematic, critical review of manual palpation for identifying myofascial trigger points: Evidence and clinical significance. Arch Phys Med Rehabil 2008;89(6):1169–76.

18 Lucas N, Macaskill P, Irwig L, Moran R, Bogduk N. Reliability of physical examination for diagnosis of myofascial trigger points: A systematic review of the literature. Clin J Pain 2009;25(1):80–9.

19 Tunks E, McCain GA, Hart LE, et al. The reliability of examination for tenderness in patients with myofascial pain, chronic fibromyalgia and controls. J Rheumatol 1995;22(5):944–52.

20 Gerber LH, Sikdar S, Armstrong K, et al. A systematic comparison between subjects with no pain and pain associated with active myofascial trigger points. PM R 2013;5(11):931–8.

21 Wolfe F, Smythe HA, Yunus MB, et al. The American College of Rheumatology 1990 criteria for the classification of fibromyalgia. Report of the Multicenter Criteria Committee. Arthritis Rheum 1990;33(2):160–72.

22 Portney LG, Watkins MP. Foundations of Clinical Research: Applications to Practice, 3rd edition. New Jersey: Pearson; 2008.

23 Price DD, McGrath PA, Rafii A, Buckingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. Pain 1983;17(1):45–56.

24 Alonso J, Prieto L, Anto JM. [The Spanish version of the SF-36 Health Survey (the SF-36 Health Questionnaire): An instrument for measuring clinical results]. Med Clin (Barc) 1995;104(20):771–6.

25 Fischer AA. Reliability of the pressure algometer as a measure of myofascial trigger point sensitivity. Pain 1987;28(3):411–4.

26 Fischer AA. Algometry in diagnosis of musculoskeletal pain and evaluation of treatment outcome: An update. J Musculoskelet Pain 1998;6(1):5–32.

27 Donner A, Rotondi MA. Sample size requirements for interval estimation of the kappa statistic for interobserver agreement studies with a binary outcome and multiple raters. Int J Biostat 2010;6(1):Article 31.

28 Altman DG. Practical Statistics for Medical Research. London: Chapman & Halls/CRC Press; 1990.

29 Fleiss JL. The Design and Analysis of Clinical Experiments. New York: John Wiley & Sons, Inc.; 1999.

30 Chen Q, Bensamoun S, Basford JR, Thompson JM, An KN. Identification and quantification of myofascial taut bands with magnetic resonance elastography. Arch Phys Med Rehabil 2007;88(12):1658–61.

31 Sikdar S, Shah JP, Gebreab T, et al. Novel applications of ultrasound technology to visualize and characterize myofascial trigger points and surrounding soft tissue. Arch Phys Med Rehabil 2009;90(11):1829–38.

32 Ballyns JJ, Shah JP, Hammond J, et al. Objective sonographic measures for characterizing myofascial trigger points associated with cervical pain. J Ultrasound Med 2011;30(10):1331–40.

33 Ballyns JJ, Turo D, Otto P, et al. Office-based elastographic technique for quantifying mechanical properties of skeletal muscle. J Ultrasound Med 2012;31(8):1209–19.

34 Turo D, Otto P, Shah JP, et al. Ultrasonic characterization of the upper trapezius muscle in patients with chronic neck pain. Ultrason Imaging 2013;35(2): 173–87.

35 Turo D, Otto P, Hossain M, et al. Novel use of ultrasound elastography to quantify muscle tissue changes after dry needling of myofascial trigger points in patients with chronic myofascial pain. J Ultrasound Med 2015;34(12):2149–61.

36 Russell IJ. Reliability of clinical assessment measures for the classification of myofascial pain syndrome. J Musculoskelet Pain 1999;7(1/2):309–24.

37 Gunn CC. The Gunn Approach to the Treatment of Chronic Pain, Intramuscular Stimulation for Myofascial Pain of Radiculopathic Origin, 2nd edition. New York: Churchill Livingstone; 1996.

38 Lucas KR, Rich PA, Polus BI. How common are latent myofascial trigger points in the scapular positioning muscles? J Musculoskelet Pain 2008;16(4): 279–86.

39 Dommerholt J. Dry needling—peripheral and central considerations. J Manual Manipul Ther 2011;19 (4):223–37.

40 Shah JP, Phillips TM, Danoff JV, Gerber LH. An in vivo microanalytical technique for measuring the local biochemical milieu of human skeletal muscle. J Appl Physiol 2005;99(5):1977–84.

41 Calvo Lobo C, Romero Morales C, Rodriguez Sanz D, et al. Comparison of hand grip strength and upper limb pressure pain threshold between older adults with or without non-specific shoulder pain. PeerJ 2017;5:e2995.

42 Ohrbach R, Gale EN. Pressure pain thresholds, clinical assessment, and differential diagnosis: Reliability and validity in patients with myogenic pain. Pain 1989;39(2):157–69.

43 Hong C-Z, Chen YN, Twehous DA, Hong DH. Pressure threshold for referred pain by compression on the trigger point and adjacent areas. J Musculoskelet Pain 1996;4(3):61–79.

44 Keefe FJ, Dolan EA. Correlation of pain behavior and muscle activity in patients with myofascial pain-dysfunction syndrome. J Craniomandib Disord 1988;2(4):181–4.

45 LeResche L, Dworkin SF. Facial expressions of pain and emotions in chronic TMD patients. Pain 1988; 35(1):71–8.

46 LeResche L, Dworkin SF, Wilson L, Ehrlich KJ. Effect of temporomandibular disorder pain duration on facial expressions and verbal report of pain. Pain 1992;51(3):289–95.

47 Craig KD, Versloot J, Goubert L, Vervoort T, Crombez G. Perceiving pain in others: Automatic and controlled mechanisms. J Pain 2010;11(2): 101–8.

48 Prkachin KM, Hughes E, Schultz I, Joy P, Hunt D. Real-time assessment of pain behavior during clinical assessment of low back pain patients. Pain 2002;95(1–2):23–30.

49 Prkachin KM, Schultz I, Berkowitz J, Hughes E, Hunt D. Assessing pain behaviour of low-back pain patients in real time: Concurrent validity and examiner sensitivity. Behav Res Ther 2002;40(5): 595–607.

50 Courbalay A, Deroche T, Descarreaux M, et al. Facial expression overrides lumbopelvic kinematics for clinical judgements about low back pain intensity. Pain Res Manag 2016;2016:7134825.