

# Reliability of Physical Examination for Diagnosis of Myofascial Trigger Points

## *A Systematic Review of the Literature*

Nicholas Lucas, BSc, MHSc, MPainMed,\*† Petra Macaskill, BA, MAppStat, PhD,†  
Les Irwig, MBBCH, PhD,† Robert Moran, BSc, MHSc,‡  
and Nikolai Bogduk, MBBS, PhD, MD, DSc§

**Background:** Trigger points are promoted as an important cause of musculoskeletal pain. There is no accepted reference standard for the diagnosis of trigger points, and data on the reliability of physical examination for trigger points are conflicting.

**Objectives:** To systematically review the literature on the reliability of physical examination for the diagnosis of trigger points.

**Methods:** MEDLINE, EMBASE, and other sources were searched for articles reporting the reliability of physical examination for trigger points. Included studies were evaluated for their quality and applicability, and reliability estimates were extracted and reported.

**Results:** Nine studies were eligible for inclusion. None satisfied all quality and applicability criteria. No study specifically reported reliability for the identification of the location of active trigger points in the muscles of symptomatic participants. Reliability estimates varied widely for each diagnostic sign, for each muscle, and across each study. Reliability estimates were generally higher for subjective signs such as tenderness ( $\kappa$  range, 0.22-1.0) and pain reproduction ( $\kappa$  range, 0.57-1.00), and lower for objective signs such as the taut band ( $\kappa$  range, -0.08-0.75) and local twitch response ( $\kappa$  range, -0.05-0.57).

**Conclusions:** No study to date has reported the reliability of trigger point diagnosis according to the currently proposed criteria. On the basis of the limited number of studies available, and significant problems with their design, reporting, statistical integrity, and clinical applicability, physical examination cannot currently be recommended as a reliable test for the diagnosis of trigger points. The reliability of trigger point diagnosis needs to be further investigated with studies of high quality that use current diagnostic criteria in clinically relevant patients.

**Key Words:** physical examination, trigger point, myofascial pain, reproducibility, reliability, systematic review

(*Clin J Pain* 2009;25:80-89)

Pain believed to be of musculoskeletal origin is a common complaint in primary care, and is a major public health concern.<sup>1-8</sup> In the majority of patients, it is difficult to come to a definitive diagnosis. One explanation for these pain symptoms is that patients have myofascial pain syndrome, a condition characterized by painful trigger points (TPs) in muscles.<sup>9-13</sup> Numerous studies report the prevalence of trigger points in various clinical conditions, with one claim that they accounted for up to 85% of patients presenting to a tertiary pain clinic.<sup>14</sup> The list of conditions reported to be associated with TPs is extensive, and includes migraine, tension type headache, temporomandibular disorder, neck pain, shoulder pain, epicondylitis, carpal tunnel syndrome, low back pain, pelvic pain, and atypical angina pectoris.<sup>10,15-18</sup> Treatments for TPs include injections,<sup>11</sup> needling,<sup>19</sup> spray and stretch,<sup>11,13</sup> and massage<sup>11,13</sup>; yet the current evidence does not demonstrate that these treatments are more efficacious than placebo.<sup>20-23</sup>

Pivotal to the appropriate and accurate prescription of any treatment is accurate diagnosis. If the diagnosis is incorrect, an otherwise effective treatment may fail because it is being applied to patients who do not have the condition. Fundamental to accurate diagnosis is the reliability of the test used to make the diagnosis. Reliability is the extent to which examiners, using the same test on the same patients, agree on the results of the test. For the investigation of TPs, various procedures have been employed, such as microdialysis,<sup>24</sup> biopsy,<sup>25</sup> imaging techniques,<sup>26</sup> and electromyography<sup>27</sup>; yet none of these are definitive or accepted as a reference standard.<sup>10,28,29</sup> Physical examination is the only means by which to establish the diagnosis and consists of firm digital pressure applied to the muscle to identify the diagnostic criteria and elicit feedback from the patient.<sup>13</sup>

The original diagnostic criteria reported to be essential for the diagnosis of TPs were tenderness within a taut band, a predictable pattern of referred pain with palpation of the taut band, and a painful limited range of movement.<sup>13,30</sup> Over time, the criteria for TPs were modified to include reproduction of the patients familiar pain with palpation of the tender point within the taut band, a local twitch response with "snapping" palpation of the taut band, and a "jump sign" when the patient quickly retracts away in response to palpation of the taut band.<sup>15,30</sup> Currently,

Received for publication March 5, 2008; revised April 22, 2008; accepted April 24, 2008.

From the \*School of Biomedical and Health Sciences, University of Western Sydney; †Screening and Test Evaluation Program, School of Public Health, University of Sydney, Sydney; §Department of Clinical Research, Newcastle Bone and Joint Institute, Royal Newcastle Hospital, University of Newcastle, Newcastle, Australia; and ‡School of Health Science, Unitec, Auckland, New Zealand.

Sources of Support: Screening and Test Evaluation Program, School of Public Health, University of Sydney. Supported in part by the Australian National Health and Medical Research Council Program grant 402764. Conflict of interest: None.

Reprints: Nicholas Lucas, BSc, MHSc, MPain Med, School of Biomedical and Health Sciences, University of Western Sydney, Campbelltown Campus, Building 24, Rm 24.4.31, Locked Bag 1797, Penrith Sth DC NSW 1797, Australia (e-mail: n.p.lucas@uws.edu.au).

Copyright © 2008 by Lippincott Williams & Wilkins

authorities describe a TP as a hyperirritable nodule located within a taut band of skeletal muscle that when palpated is tender and produces referred pain.<sup>13,15,25,30</sup> A predictable pattern of pain referral and the local twitch response are each no longer considered to be sufficient or necessary for the diagnosis.<sup>15,30</sup>

After the identification of a TP, examiners determine if it is “active” (symptomatic) or “latent” (asymptomatic). Essential for the diagnosis of an active TP is the criterion that the patient’s pain can be reproduced by palpation of the tender point within the taut band.<sup>28,30</sup> Because only patients with symptoms can have active TPs, it is necessary to demonstrate the reliability of this diagnosis in symptomatic participants. If physical examination were reliable for the identification of TPs, examiners would be able to agree upon the exact location of the taut band, and the presence of a tender nodule within it. If this finding were clinically relevant, then palpation of the nodule within the taut band should also reproduce the patients’ familiar pain.<sup>15</sup> If a diagnosis requires several diagnostic criteria to be satisfied, such as with TPs, then studies of the reliability of that diagnosis must address agreement on the simultaneous presence of all signs in each patient.

Several studies have been conducted to assess the reliability of palpation for TPs. Given the importance of reliability for the assessment of TPs, the evidence on reliability is worthy of systematic review. The primary objectives of this systematic review were to collect and summarize the available evidence regarding the reliability of physical examination for the diagnosis of TPs, and to analyze studies for their quality and for factors that might affect their applicability in clinical practice. Secondary objectives were to identify deficiencies in the available evidence, and to recommend how these might be redressed. Finally, implications for clinical practice and further research in this area are explored.

## METHODS OF THE REVIEW

### Identification of Studies

A search was conducted of electronic databases, specialist journals and textbooks on myofascial pain syndrome. Forward citation tracking, reference scanning of articles included in the final sample, and contact with experts in the field, were used to identify additional articles.

The search string for electronic databases included the following medical subject headings (MeSH) and text words (tw): (1) [Trigger Point (MeSH) OR trigger point (tw) OR myofascial pain (tw) OR muscle pain (tw)] AND [Physical Examination (MeSH) OR physical examination (tw) OR palpation (tw) OR manual examination (tw) or Diagnosis, Musculoskeletal (MeSH) OR diagnosis (tw)]. To increase the sensitivity of the search, terms relating to the study design (eg, reliability) were excluded from the search string.

The OVID platform was used to search MEDLINE, EMBASE, the cumulative index of nursing and allied health literature (*CINAHL*), Alternative Medicine (*AMED*), and SPORTSDiscus. A direct search of the Australasian Medical Index (*AMI*), and the Manual, Alternative and Natural Therapy Index (*MANTIS*) was also conducted. Each database was searched from inception until March 2008. The following journals and textbooks were manually searched for relevant articles:

The Journal of Musculoskeletal Pain (1998 to March 2008);  
Journal of Musculoskeletal Medicine (1999 to March 2008);  
International Musculoskeletal Medicine (1990 to March 2008);  
Journal of Manual and Manipulative Therapy (1993 to March 2008);  
Simons DG, Travell J, Simons LS. Travell and Simons Myofascial Pain and Dysfunction: The Trigger Point Manual Volumes 1 and 2. 2nd Edition. 1999  
Mense S, Simons DG. Muscle Pain. Understanding Its Nature, Diagnosis, and Treatment. 2001

### Study Selection

Two reviewers (N.L., R.M.) independently screened all citations retrieved from the literature search, and identified those that satisfied the inclusion and exclusion criteria.

### Inclusion Criteria

Studies that fulfilled all of the following criteria

1. investigated the reliability of physical examination (palpation) procedures designed to identify TPs.
2. used a repeated measures design and reported intraexaminer or interexaminer reliability.
3. investigated symptomatic or asymptomatic participants.
4. investigated “active” or “latent” TPs.

### Exclusion Criteria

Studies were excluded for either of the following reasons

1. only reported the reliability of identifying tenderness (because the diagnostic criteria of TPs includes more than just tenderness).
2. used technology or instrumentation to identify TPs (eg, algometry or fine needle electromyography).

The 2 reviewers met and resolved discrepancies concerning the potential eligibility of studies. Full text articles of potentially eligible studies were obtained and independently evaluated according to the eligibility criteria. Those studies that met the eligibility criteria formed the final sample, and were independently assessed by the 2 reviewers according to predefined quality assessment criteria (Appendix).

### Data Extraction and Quality Assessment

Two reviewers (N.L., R.M.) independently extracted data and assessed the quality of the studies. Discrepancies regarding study quality were discussed and resolved if necessary by consultation with a third reviewer (N.B.). Both reviewers had experience in the diagnosis and treatment of myofascial pain syndrome and the design and implementation of reliability studies. Each study in the final sample of eligible papers was assessed using a 12-item quality appraisal checklist designed to assess the quality and applicability of studies of diagnostic reliability (Appendix). The face validity of this checklist was established by consultation with methodology experts and by comparison with quality appraisal checklists used in other systematic reviews of diagnostic reliability.<sup>31–36</sup> This checklist was also developed with reference to the statement for standards for reporting studies of diagnostic accuracy (STARD),<sup>37</sup> and the quality assessment of diagnostic accuracy studies (QUADAS)<sup>38</sup> appraisal tool. Studies were not given an overall numeric quality score; instead each item was

considered separately and answered “yes,” “no,” “unclear,” or “not applicable.”

### Estimates of Reliability

Statistical estimates of reliability were extracted from the primary studies and reported for the purpose of summarizing the available evidence. Estimates of reliability were evaluated in the context of the quality of the study and its applicability to clinical practice. For categorical data,  $\kappa$  is routinely reported as a chance-corrected index of agreement, and is interpreted qualitatively in this report as: almost perfect ( $\kappa = 0.81-1.00$ ), substantial ( $\kappa = 0.61-0.80$ ), moderate ( $\kappa = 0.41-0.60$ ), fair ( $\kappa = 0.21-0.40$ ), slight ( $\kappa = 0.00-0.20$ ), or poor ( $\kappa < 0.00$ ).<sup>39</sup>

### Grouping of Studies

The eligible studies were divided into 2 main groups: those that included both symptomatic and asymptomatic participants, and those that included asymptomatic participants only. This distinction was made on the following grounds. Only active TPs produce symptoms and therefore can only be found in symptomatic participants. Only those studies that assessed the reliability of palpation to identify active TPs in symptomatic participants could be applicable to real-life clinical practice. Latent TPs may theoretically occur in both symptomatic and asymptomatic participants and it is therefore essential that examiners are able to reliably distinguish between active and latent TPs. Studies that include only asymptomatic participants can only be used to investigate the reliability of latent TP identification. As the literature describes latent trigger points as an entity that can occur in asymptomatic participants, these studies were retained and summarized separately for what they might show, in principle, about the diagnosis of active TPs in symptomatic patients.

## RESULTS

### Literature Search

A total of 2591 citations were identified from the literature search. Of these, 2563 were excluded after review of the title and abstract because they did not satisfy the eligibility criteria or were duplicates of the same study. Of the remaining 28 articles, 20 were excluded after full text review because they did not report reliability data,<sup>16,17,40-44</sup> only reported muscle tenderness,<sup>18,45-54</sup> did not assess TPs,<sup>55</sup> or did not differentiate between TPs and tenderness.<sup>56</sup> Nine studies, reported in 8 articles,<sup>57-64</sup> were included in the final sample (Fig. 1). The study characteristics of the final sample are summarized in Table 1. No additional studies were identified for inclusion from the reference lists of the final sample, forward citation tracking, or from direct contact with experts in the field. A single discrepancy between reviewers about the inclusion of one paper<sup>64</sup> was resolved by discussion. In addition, consensus between reviewers regarding the quality of included studies and their applicability to clinical practice was achieved after discussion with the third reviewer.

Five studies<sup>59,61,62,64</sup> were conducted before the proposed inclusion of a palpable nodule in the diagnostic criteria,<sup>25</sup> and so this feature was not evaluated in those studies. In the remaining 4 studies, however, only 2<sup>58,63</sup> included this feature as a diagnostic criterion.

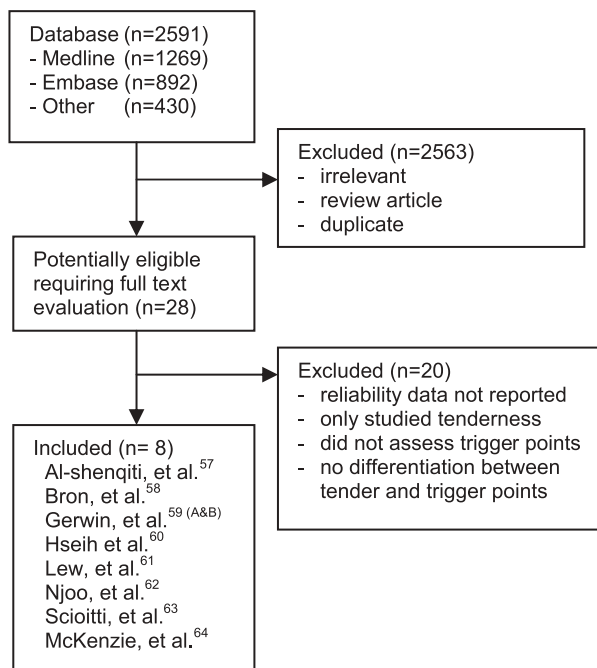


FIGURE 1. Flow chart indicating the search yields with excluded and included studies. A and B refer to 2 separate studies reported in a single paper.

### Quality and Applicability Results

None of the studies satisfied all of the relevant criteria for study quality and applicability (Table 2). None of the included studies provided data on the interexaminer reliability for the diagnosis of active TPs in symptomatic patients. One study came close to providing this information, but did not give separate reliability estimates for active and latent TPs.<sup>59</sup>

Some studies were deficient in what might be considered technicalities. Six studies<sup>57,59,61,62,64</sup> did not explicitly mention if observers were blinded to clinical information or cues that did not form part of the test under evaluation. Three of these studies<sup>59,64</sup> did not mention if examiners were blinded to the findings of other examiners. Failure to adhere to blinding would bias the results in the direction of higher reported reliability. However, these studies were nevertheless given the benefit of doubt, and accepted on the grounds that explicit statements of these critical criteria of quality were simply not included in the published manuscript.

### Studies With Symptomatic and Asymptomatic Participants

Five studies examined a mixture of symptomatic and asymptomatic participants and each reported the  $\kappa$  statistic as an estimate of reliability (Table 3).<sup>58-60,62</sup> In each of these studies, agreement was based on the ability of examiners to agree upon the presence or absence of a physical sign in a particular muscle, but not on their ability to agree about the location of the sign within the muscle. In 2 of these studies,<sup>59</sup> the examiners were expert in the field of myofascial pain and TP diagnosis, with extensive prestudy training for the specific purpose of increasing reliability. These examiners are not representative of those who would normally use the test in practice, and the results are unlikely

**TABLE 1.** Study Characteristics of Included Articles

Author (y)	Participants	Raters	Study Characteristics	Diagnostic Criteria
Al-Shenqiti and Oldham (2005) <sup>57</sup>	n = 58 diagnosed with rotator cuff tendonitis	Physiotherapist with 11 y experience	Intrarater reliability for rotator cuff muscles	LT, TB, JS, LTR, PR, RP
Bron et al (2007) <sup>58</sup>	n = 40; 32 shoulder pain, 8 controls	Physiotherapist (n = 3) with 29, 28, 16 y experience	Interrater reliability for location of 3 trigger points in rotator cuff muscles	TB-N, LT, RP, LTR, JS
Gerwin et al <sup>A</sup> (1997) <sup>59</sup>	n = 25; symptomatic/asymptomatic	Physiatrists (n = 2), neurologists (n = 2) all experts Phase A, had discussion session before data collection	Interrater reliability For phase A, for 10 paired muscles	LT, TB, LTR, PR, RP, TrP
Gerwin et al <sup>B</sup> (1997) <sup>59</sup>	n = 10; 7 symptomatic, 3 asymptomatic	Phase B, met for 3 h before practice TrP diagnosis	For phase B, for 5 paired muscles	
Hsieh et al (2000) <sup>60</sup>	n = 52; 26 low back pain patients, 26 asymptomatic	Chiropractors (n = 2) with 3 and 5 y experience, and physiatrists (n = 2) with 3 and 6 y experience each received 3 × 2 h classroom lectures on MPS	Interrater reliability  Presence of TP in 10 paired muscles	TB, LTR, RP
Lew et al (1997) <sup>61</sup>	n = 58; asymptomatic volunteers	Chiropractors (n = 2) with 15 y experience, and physiatrists (n = 2) with 2 y experience received no training Dual qualified physiotherapist and osteopath (n = 2) expert in trigger point diagnosis	Interrater reliability for the trapezius muscle Location of latent TP drawn on a 1/4 size body chart	LT, TB, RP
McKenzie et al (1997) <sup>64</sup>	n = 6; children with cerebral palsy	Clinicians (n = 3) experience and qualifications unclear	Interrater reliability study for trigger points in 16 muscles	TB-N, LT, RP
Njoo et al (1994) <sup>62</sup>	n = 124; 61 low back pain patients, 63 controls	General practitioner (n = 1) with training at a rheumatology university hospital; final year medical students (n = 4) trained with the GP	Interrater reliability for the presence of TP in the QL and GM muscles in patients with and without LBP	LT, TB, JS, LTR, PR, RP
Sciotti et al (2001) <sup>63</sup>	n = 20; asymptomatic volunteers	Clinicians (n = 4) (discipline or qualifications not specified), with 9, 13, 14, 18 y of experience Undertook 4 × 3 h training sessions on the testing procedure	Interrater reliability to identify the location of TP within the trapezius muscle	TB, LT, N, RP, LTR, JS

GM indicates gluteus maximus; GP, general practitioner; JS, jump sign; LT, local tenderness; LTR, local twitch response; MPS, myofascial pain syndrome; N, nodule; PR, pain reproduction; QL, quadratus lumborum; RP, referred pain; TB, taut band; TP, trigger point.

TABLE 2. Summary of the Quality of Studies of Diagnostic Reliability for Myofascial Trigger Points

	Al-Shengiti and Oldham <sup>57</sup>	Bron et al <sup>58</sup>	Gerwin et al <sup>59</sup>	Gerwin et al <sup>60</sup>	Lew et al <sup>61</sup>	McKenzie et al <sup>64</sup>	Njoo et al <sup>62</sup>	Sciotti et al <sup>63</sup>
Sample representative	Y	Y	Y	Y	U	N	Y	N
Raters representative	N	Y	N	Y	U	U	Y	U
Blinding (reference standard)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Blinding (other raters)	N/A	Y	U	Y	Y	U	Y	Y
Blinding (own findings)	U	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Blinding (clinical information)	U	Y	U	Y	U	U	U	Y
Blinding (covert cues)	U	Y	U	U	Y	U	U	Y
Varied order of examination	U	Y	Y	Y	Y	Y	U	Y
Appropriate statistics	Y	Y	Y	Y	N	U	Y	Y
Test appropriate	N	N	N	N	U	U	N	Y
Time interval	Y	Y	Y	Y	Y	U	Y	Y
Dropouts	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Yes	3/9	8/9	4/9	4/9	4/9	1/9	5/9	7/9
No	2/9	1/9	2/9	2/9	1/9	1/9	1/9	1/9
Unclear	4/9	0/9	3/9	3/9	4/9	7/9	3/9	1/9

Refer to the appendix for information about each item on this checklist.

to be applicable to everyday clinical practice. Nevertheless, these studies provide estimates of reliability that might constitute a benchmark of what could, or should, be obtained in conventional practice. Only 3 studies<sup>58,60,62</sup> used what might be construed as representative examiners.

Estimates of reliability varied widely between each study, for the individual muscle tested, and also for each of the individual signs that make up the criteria for the identification of TPs. None of the studies stratified their results according to asymptomatic and symptomatic groups, nor active or latent TPs. Therefore, the extent to which estimates of reliability pertain to active or latent TPs is not evident. In the majority of studies, the outcome was dichotomous with the relevant feature being recorded as either absent or present, without specifying how indeterminate results were handled. In one study,<sup>58</sup> all indeterminate results were pooled into the “absent” category, potentially inflating observed agreement and estimates of reliability between examiners.

Only one study reported agreement between examiners for the presence or absence of a TP within a group of muscles ( $\kappa = 0.66-0.95$ ), but not for the location of the TP.<sup>59</sup> All other studies<sup>58,60,62</sup> only reported estimates for the individual signs of TPs, without providing data for the overall agreement about the presence or absence of TPs. The small number of available studies differ markedly with respect to design, the prevalence of positive findings within the study sample, and the purported skill of the examiners, making comparisons between studies inappropriate.<sup>65</sup> Hence, only a qualitative summary of the results is provided.

Across all studies, reliability estimates for the taut band ranged from  $\kappa = -0.08$  to 0.75. For tenderness, the reliability estimates ranged from  $\kappa = 0.22$  to 1.0, and for reproduction of familiar pain the reliability estimates ranged from  $\kappa = 0.57$  to 1.0. For referred pain, reliability estimates ranged from  $\kappa = -0.13$  to 0.84, and for the local twitch response estimates ranged from  $\kappa = -0.05$  to 0.57. Reported estimates for the “jump sign” ranged between  $\kappa = 0.07$  and 0.71.

Because  $\kappa$  is affected by prevalence, we had aimed to report the frequency of abnormal findings in conjunction with  $\kappa$  to aid interpretation. However, the studies did not provide sufficient data to compute the relevant frequencies.

### Studies With Asymptomatic Participants Only

Two studies used asymptomatic volunteers.<sup>61,63</sup> In one of these studies,<sup>61</sup> the examiners identified latent TPs in the trapezius muscle and instructed an assistant to mark the location of the TPs on a printed body map that was reduced to approximately 25% of an average human adult. The accuracy and resolution of this recording method was not stated. In the other study,<sup>63</sup> the examination was limited to the assessment of latent TPs in the right upper trapezius muscle. Once an examiner had identified a latent TP, an assistant applied a digitizing probe to the site indicated by the examiner and the 3-dimensional coordinates of the location were recorded electronically. In both of these studies,<sup>61,63</sup> it was unclear if examiners were representative of those who would normally use the test in practice, and as all participants were asymptomatic, the findings of these studies lack applicability to everyday clinical practice. In addition, intramuscular fascial bands and neurovascular structures within the trapezius may account for the clinical

**TABLE 3.** Kappa Scores From Studies of Interrater Reliability

Study (n = participants)	Quality Criteria Met	$\kappa$ (Range) by Diagnostic Criterion					
		Tenderness	Pain Recognition	Referred Pain	Jump Sign	Taut Band	Local Twitch
Bron et al (n = 40)	8/9	•	•	-0.13, 0.64	0.07, 0.68	0.11, 0.75 <sup>N</sup>	-0.05, 0.45
Hsieh et al (n = 52)	7/9	•	•	0.13, 0.50 <sup>(1)</sup>	•	0.00, 0.19 <sup>(1)</sup>	0.02, 0.18 <sup>(1)</sup>
	7/9	•	•	0.08, 0.49 <sup>(2)</sup>	•	-0.08, 0.06 <sup>(2)</sup>	-0.02, 0.07 <sup>(2)</sup>
Njoo et al (n = 124)	5/9	0.58, 0.73	0.57, 0.58	0.36, 0.46	0.68, 0.71	0.47, 0.51	-0.02, 0.19
Gerwin et al <sup>A</sup> (n = 25)	4/9	0.22*	†	†	•	†	†
Gerwin et al <sup>B</sup> (n = 10)	4/9	0.48, 1.0	0.79, 1.0	0.57, 0.84	•	0.40, 0.46	0.11, 0.57
Range across all studies		0.22, 1.0	0.57, 1.0	-0.13, 0.84	0.07, 0.71	-0.08, 0.75	-0.05, 0.57

Studies are in descending order for total number of possible positive quality and applicability items achieved. For each sign, the range of  $\kappa$  scores that were obtained from the examination of each muscle included in the study is reported.

Criteria for which no  $\kappa$  scores were provided are indicated by •.

\*Indicates that this is the average  $\kappa$  obtained across 20 muscles. For Bron et al,<sup>58</sup> N indicates that examiners reported a palpable nodule within a taut band.

†Indicates that  $\kappa$  was not reported, but the authors reported that  $\kappa$  was poor. For Hsieh et al,<sup>60</sup>  $\kappa$  scores were provided for: (1) examiners agreement with an expert; and (2) agreement between examiners, excluding the expert. For Gerwin et al,<sup>59</sup> 2 studies, A and B, were reported in 1 paper.

phenomena of latent TPs palpated in these asymptomatic participants.<sup>66</sup>

Neither of these two studies reported conventional measures of reliability. Lew et al<sup>61</sup> argue that  $\kappa$  could not be calculated from their data due to the presence of a “zero” cell in their contingency table, however this opinion is arguable. Our calculation from the data provided by Lew et al for the identification of latent TPs is  $\kappa = -0.26$ . Irrespective of this, Lew et al reported observed agreement ( $P_o$ ) of only 21% for latent TPs that produced referred pain, and 10% for latent TPs that did not produce referred pain. Importantly, in only 2 of 42 participants identified by both examiners as having latent TPs was there agreement about the actual location of the TPs. In the study of Sciotti et al,<sup>63</sup> agreements were assessed at 2 levels. First, examiner agreement was assessed for the presence or absence of latent TPs in the right upper trapezius muscle; however the data provided did not allow calculation of  $\kappa$ . Second, examiners were assessed on their ability to identify the location of a TP using a 3-dimensional computerized tracking system. The average of the 3-dimensional coordinates obtained from each examiner was used as a reference point to represent the estimated location of the TP in that participant. The distance by which each examiner deviated from the reference point was then used to calculate examiner agreement. The examiner error associated with identifying the estimated location of TPs ranged from 3.3 to 6.6 cm.

**Other Studies**

Al-Shenqiti and Oldham<sup>57</sup> reported the only study on intrarater reliability. Patients with a clinical diagnosis of rotator cuff tendonitis were evaluated for the signs of TPs within the rotator cuff muscles. Agreement was based on the identification of physical signs within a muscle, but not on the location of those signs within the muscle. The examiner was not blinded to the participants’ clinical

history and it was not evident if the examiner was blinded to the identity of participants previously examined, or additional cues that may have influenced the test outcome. Information regarding the qualifications, training, skills and experience of the examiner were not adequately reported. The  $\kappa$  statistics reported for each sign ranged from  $\kappa = 0.75$  to 1.0, and were higher than those from interrater reliability studies. However, it is difficult to interpret these results due to issues with the quality of this study, outlined above, and its applicability to clinical practice.

McKenzie et al<sup>64</sup> provided an estimate of examiner agreement for the presence or absence of TPs in children with cerebral palsy. The authors report that out of 16 muscles examined in 6 children by 3 examiners, only 5 disagreements for the presence or absence of TPs occurred, and reported an intraclass correlation coefficient of 0.93. However, examiner reliability was not the main focus of this paper and no information was included about blinding, or how the location of the physical signs were recorded and compared between examiners. In addition, there was inadequate detail to assess if appropriate statistical analysis was performed. Lastly, the participants included in this study were diagnosed with a neurologic disorder affecting muscle tissue and are not representative of patients in typical clinical practice. The results of this study are therefore difficult to interpret, and are reported here for the purpose of completeness.

**DISCUSSION**

None of the 9 studies in this review specifically reported interrater reliability estimates for the identification of the location of active TPs in symptomatic participants. In those studies that did include symptomatic participants, the data were not stratified according to the symptom status of the participant, or the status of the TP as active or

latent.<sup>58-60,62</sup> Each of the included studies falls short of the optimal design in at least one respect thereby compromising the quality of the evidence.

Each study that examined both symptomatic and asymptomatic participants provided reliability estimates for individual signs of TPs. Only one of these studies<sup>59</sup> provided estimates of examiner agreement about the presence or absence of a TP, which varied depending upon the muscle examined ( $\kappa = 0.66-0.95$ ). However, the highest reported reliability estimate for the taut band in this study was  $\kappa = 0.46$ , which is lower than the range reported for the identification of a TP. It is therefore unclear which of the individual signs of TPs influenced the examiners in their overall decision about the presence or absence of a trigger point. The subjective signs derived from patient responses in this study, such as reports of tenderness ( $\kappa = 0.48-1.0$ ), may have influenced examiners more than the objective signs derived from palpation, such as the taut band ( $\kappa = 0.46$ ). Further, these data do not pertain to the currently proposed criteria for trigger points, which require the palpatory identification of a nodule within the taut band in addition to the reproduction of the patients' familiar pain; and nor do these data pertain to examiners agreement about the exact location of the TP. Lastly, the examiners in this study were expert in trigger point diagnosis and participated in prestudy training sessions for the purpose of increasing their reliability. For these reasons, the results from this study may not be applicable in everyday clinical practice and are likely to represent the upper limits of reliability estimates for TP diagnosis.

The remaining data pertain only to the reliability of individual signs of TPs, and not to the identification of TPs in totality. Overall, reliability was better for the subjective signs of tenderness, pain recognition, referred pain, and the jump sign, whereas reliability estimates for the taut band and twitch response were lower. Better reliability was recorded in studies that used experts who underwent special training than in studies that used conventional practitioners; and this underscores the concern that practitioners in typical clinical practice may fail to achieve the benchmarks set by experts.

In addition, and particularly prejudicial to TP theory, is the lack of any data on the reliability of pinpointing the exact location of active TPs. The relevant data on reliability pertain only to agreeing if a muscle has the signs of a TP, and not to the exact location of these signs. Yet independent identification of the location of the TP is essential to demonstrate that examiners are identifying the same phenomenon. If it is not evident that examiners can agree on the precise location of an active TP, they cannot be relied upon to place their treatment accurately. Two studies<sup>61,63</sup> evaluated agreement between examiners as to the location of latent TPs. In one study,<sup>61</sup> examiner agreement was found not to exceed 21%, and the other<sup>63</sup> reported the overall magnitude of errors between examiners to be 3.3 cm to 6.6 cm, which is not sufficiently accurate for specific treatment such as injection.

## Implications

Evidence for the diagnostic reliability of TPs is available from only a limited number of studies, and reliability data are lacking for clinically relevant, active TPs in symptomatic patients. In consideration of this, the diagnosis and treatment of TPs does not have a firm clinical basis. Until a reliable diagnostic test for TPs has been

demonstrated it is recommended that this diagnosis should not be considered as a primary, or exclusive diagnosis for patients presenting a report of pain. If a treatment or management plan is to be implemented on the basis of a diagnosis of TPs, then patients should be informed of the ambiguity of this diagnosis so that they may make an informed choice about their treatment options.

In the absence of a reliable diagnostic test for TPs, one or more of 3 main problems will confound research investigations. First, they will be confounded by the inclusion of inappropriate cases. Second, it will not be possible to determine the precision of the diagnosis, or the treatment. Third, there will be no certainty in the measurement of TP resolution after treatment. This has implications for the interpretation of prior studies investigating the prevalence of TPs in specific patient populations, and those evaluating the effectiveness of treatments for TPs in clinical trials, for these studies are based on a test of unknown reliability and validity. Reliable methods of identifying TPs should be demonstrated before the implementation of further studies investigating the prevalence or treatment of trigger points.

For practical and taxonomic purposes, proponents of TP theory might choose to redefine the diagnostic criteria to include only tenderness and pain reproduction and eliminate the need for the taut band and twitch response. If the criteria are relaxed to define TPs as a tender area in muscle that when palpated reproduces the patients familiar pain, then the present evidence indicates that worthwhile agreement might be achieved. In that event, however, the TP theory would be compromised. Even though examiners agree upon tenderness and pain reproduction, these physical signs alone do not point to a pathologic entity, and would not be distinguishable from nonspecific muscle hyperalgesia.

## Future Research

For a meaningful estimate of reliability for the diagnosis of TPs, an optimal study would have the following design features. Examiners would be blinded as to the condition of the patient, and to any clinical information other than that which they would derive from their examination. Without blinding, the study would not be a test of the clinical signs in question, for the other information might influence the examiners. Data would be collected and reported on the agreement, or otherwise, of all examiners for all participants. The raw data should be provided, so that consumers of the information are able to check and verify the estimates of reliability. For an entity whose diagnostic criteria are multiple and compound, data would be provided not only on the reliability of each component but also on the reliability of the composite findings simultaneously. For an entity such as TPs, agreement would be assessed for the identification of the exact location of the feature, for treatment is applied to the entity, and not generally to the muscle in which it is found. For an entity purported to produce symptoms, reliability would be tested in symptomatic participants, for it is only in such participants that the results apply and are of relevance. The examiners would be representative of the practitioners interested in the information, for only then might the results be applicable to everyday clinical practice. If experts are used, the results represent what the benchmark of performance might be, but they might not be

generalizable to less trained or less experienced practitioners.

The findings of this study prompt consideration of alternative explanations for the clinical phenomena attributed to trigger points. Other competing explanations include referred pain of peripheral nerve origin,<sup>67</sup> secondary hyperalgesia,<sup>68</sup> underlying joint tenderness,<sup>69</sup> and normal intramuscular anatomy.<sup>66</sup> A discussion about future research should therefore include reference to the validity of physical examination for TPs. Because there is no accepted reference standard for TPs, physical examination findings cannot yet be validated, even though they may be shown to be reliable. Researchers who investigate reliability should also consider comparing physical examination findings with the laboratory and imaging findings publicized<sup>29</sup> as evidence of the existence of TPs.

**CONCLUSIONS**

No study to date has reported the reliability of TP diagnosis according to the currently proposed criteria in symptomatic patients. Although some of the individual signs, such as tenderness and pain reproduction, may be reliable to some extent in some muscles, their detection in isolation does not qualify as a diagnosis for TPs. The interrater reliability estimates for each sign varied widely for each muscle and between different muscles, ranging from less than chance to substantial reliability. However, insufficient data were provided upon which to interpret these results with confidence. On the basis of the limited number of studies available, and significant problems with their design, reporting, statistical integrity, and clinical applicability, physical examination cannot currently be recommended as a reliable test for the diagnosis of TPs. High-quality studies are needed to investigate the reliability of identifying the exact location of TP using current diagnostic criteria in clinically relevant patients.

**Key Messages**

- No study has reported the reliability of identifying the exact location of active TPs
- No study has reported the reliability of all diagnostic criteria simultaneously
- The majority of studies did not report important aspects such as whether examiners were blinded to clinical information or additional cues that may have influenced their findings
- For each individual criterion, estimates of reliability were inconsistent and ranged from excellent to less than chance for different muscles and across different studies
- Tenderness and pain reproduction achieved acceptable reliability; however these are not specific for TPs

**ACKNOWLEDGMENTS**

*The authors thank Ruth Mitchell, Cochrane Renal Group, The Children’s Hospital at Westmead, Sydney, Australia, and Phillipa Powis, Department of Clinical Research, Royal Newcastle Center and University of Newcastle, Newcastle, Australia, for their advice regarding the development of the search strategy.*

**APPENDIX**

**Quality Appraisal of Diagnostic Reliability Checklist**

Item	Yes	No	Unclear	N/A
1. Was the test evaluated in a spectrum of subjects representative of patients who would normally receive the test in clinical practice?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
2. Was the test performed by examiners representative of those who would normally perform the test in practice?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
3. Were raters blinded to the reference standard for the target disorder being evaluated?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Were raters blinded to the findings of other raters during the study?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Were raters blinded to their own prior outcomes of the test under evaluation?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Were raters blinded to clinical information that may have influenced the test outcome?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Were raters blinded to additional cues, not intended to form part of the diagnostic test procedure?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
8. Was the order in which raters examined subjects varied?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Were appropriate statistical measures of agreement used?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
10. Was the application and interpretation of the test appropriate?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
11. Was the time interval between measurements suitable in relation to the stability of the variable being measured?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
12. If there were dropouts from the study, was this less than 20% of the sample.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>TOTAL</b>				

**REFERENCES**

1. Blyth FM, March LM, Brnabic AJ, et al. Chronic pain in Australia: a prevalence study. *Pain*. 2001;89:127–134.
2. Brattberg G, Thorslund M, Wikman A. The prevalence of pain in a general population. The results of a postal survey in a county of Sweden. *Pain*. 1989;37:215–222.
3. Cote P, Cassidy JD, Carroll LJ, et al. The annual incidence and course of neck pain in the general population: a population-based cohort study. *Pain*. 2004;112:267–273.

4. Elliott AM, Smith BH, Hannaford PC, et al. The course of chronic pain in the community: results of a 4-year follow-up study. *Pain*. 2002;99:299–307.
5. Eriksen J, Jensen MK, Sjogren P, et al. Epidemiology of chronic non-malignant pain in Denmark. *Pain*. 2003;106:221–228.
6. Gureje O, Simon GE, Von Korff M. A cross-national study of the course of persistent pain in primary care. *Pain*. 2001;92:195–200.
7. Magni G, Caldieron C, Rigatti-Luchini S, et al. Chronic musculoskeletal pain and depressive symptoms in the general population. An analysis of the 1st National Health and Nutrition Examination Survey data. *Pain*. 1990;43:299–307.
8. Blyth FM, March LM, Nicholas MK, et al. Chronic pain, work performance and litigation. *Pain*. 2003;103:41–47.
9. Cummings M, Baldry P. Regional myofascial pain: diagnosis and management. *Best Prac Res Clin Rheumatol*. 2007;21:367–387.
10. Dommerholt J, Bron C, Franssen J. Myofascial trigger points: an evidence-informed review. *J Man Manip Ther*. 2006;14:203–221.
11. Lavelle ED, Lavelle W, Smith HS. Myofascial trigger points. *Med Clinics North Am*. 2007;91:229–239.
12. Mense S, Simons DG. *Muscle Pain: Understanding Its Nature, Diagnosis, and Treatment*. Philadelphia: Lippincott Williams & Wilkins; 2001.
13. Simons DG, Travell J, Simons LS. *Travell & Simons Myofascial Pain and Dysfunction: The Trigger Point Manual*. Vol. 1. Upper Half of Body: 2nd ed. Baltimore: Lippincott Williams & Wilkins; 1999.
14. Fishbain DA, Goldberg M, Meagher BR, et al. Male and female chronic pain patients categorized by DSM-III psychiatric diagnostic criteria. *Pain*. 1986;26:181–197.
15. Simons DG. Understanding effective treatments of myofascial trigger points. *J Bodywork Mov Ther*. 2002;6:81–88.
16. Calandre EP, Hidalgo J, Garcia-Leiva JM, et al. Trigger point evaluation in migraine patients: an indication of peripheral sensitization linked to migraine predisposition? *Eur J Neurol*. 2006;13:244–249.
17. Fernández-de-las-Peñas C, Alonso-Blanco C, Luz Cuadrado M, et al. Myofascial trigger points and their relationship to headache clinical parameters in chronic tension-type headache. *Headache*. 2006;46:1264–1272.
18. Weiner DK, Sakamoto S, Perera S, et al. Chronic low back pain in older adults: prevalence, reliability, and validity of physical examination findings. *J Am Geriatr Soc*. 2006;54:11–20.
19. Baldry PE. *Acupuncture, Trigger Points and Musculoskeletal Pain*. 3rd ed. Edinburgh: Elsevier Churchill Livingstone; 2004.
20. Cummings TM, White AR. Needling therapies in the management of myofascial trigger point pain: a systematic review. *Arch Phys Med Rehabil*. 2001;82:986–992.
21. Fernandez De Las Penas C, Sorhbeck Campo M, Fernandez Carnero J, et al. Manual therapies in myofascial trigger point treatment: a systematic review. *J Bodywork Mov Ther*. 2005;9:27–34.
22. Ho K-Y, Tan K-H. Botulinum toxin A for myofascial trigger point injection: a qualitative systematic review. *Eur J Pain*. 2007;11:519–527.
23. Rickards LD. The effectiveness of non-invasive treatments for active myofascial trigger point pain: a systematic review of the literature. *Int J Osteopath Med*. 2006;9:120–136.
24. Shah JP, Danoff JV, Desai MJ, et al. Biochemicals associated with pain and inflammation are elevated in sites near to and remote from active myofascial trigger points. *Arch Phys Med Rehabil*. 2008;89:16–23.
25. Simons DG. Diagnostic criteria of myofascial pain caused by trigger points. *J Musculoskelet Pain*. 1999;7:111–120.
26. Chen Q, Bensamoun S, Basford JR, et al. Identification and quantification of myofascial taut bands with magnetic resonance elastography. *Arch Phys Med Rehabil*. 2007;88:1658–1661.
27. Simons DG, Hong CZ, Simons LS. Endplate potentials are common to midfiber myofascial trigger points. *Am J Phys Med Rehabil*. 2002;81:212–222.
28. Simons DG. Review of enigmatic MTRPs as a common cause of enigmatic musculoskeletal pain and dysfunction. *J Electro-myogr Kinesiol*. 2004;14:95–107.
29. Simons DG. New views of myofascial trigger points: etiology and diagnosis. *Arch Phys Med Rehabil*. 2008;89:157–159.
30. Tough EA, White AR, Richards S, et al. Variability of criteria used to diagnose myofascial trigger point pain syndrome—evidence from a review of the literature. *Clin J Pain*. 2007;23:278–286.
31. Gemmel H, Miller P. Interexaminer reliability of multi-dimensional examination regimens used for detecting spinal manipulable lesions: a systematic review. *Clin Chiropractic*. 2005;8:199–204.
32. Hestboek L, Leboeuf-Yde C. Are chiropractic tests for the lumbo-pelvic spine reliable and valid? A systematic critical literature review. *J Manipulative Physiol Ther*. 2000;23:258–275.
33. Hollerwoger D. Methodological quality and outcomes of studies addressing manual cervical spine examinations: a review. *Man Ther*. 2006;11:93–98.
34. May S, Littlewood C, Bishop A. Reliability of procedures used in the physical examination of non-specific low back pain: a systematic review. *Aust J Physiother*. 2006;52:91–102.
35. Stochkendahl MJ, Christensen HW, Hartvigsen J, et al. Manual examination of the spine: a systematic critical literature review of reproducibility. *J Manipulative Physiol Ther*. 2006;29:475–485, 485 e1–10.
36. van Trijffel E, Anderegg Q, Bossuyt PM, et al. Inter-examiner reliability of passive assessment of intervertebral motion in the cervical and lumbar spine: a systematic review. *Man Ther*. 2005;10:256–269.
37. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD Initiative. *Ann Intern Med*. 2003;138:40–44.
38. Whiting P, Rutjes AW, Reitsma JB, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*. 2003;3:25.
39. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.
40. Fernández-de-las-Peñas C, Alonso-Blanco C, Luz Cuadrado M, et al. Trigger Points in the Suboccipital Muscles and Forward Head Posture in Tension-Type Headache. *Headache*. 2006;46:454–460.
41. Fernández-de-las-Peñas C, Ge HY, Arendt-Nielsen L, et al. Referred pain from trapezius muscle trigger points shares similar characteristics with chronic tension type headache. *Eur J Pain*. 2007;11:475–482.
42. Couppe C, Torelli P, Fuglsang-Frederiksen A, et al. Myofascial trigger points are very prevalent in patients with chronic tension-type headache: a double-blinded controlled study. *Clin J Pain*. 2007;23:23–27.
43. Skootsky SA, Jaeger B, Oye RK. Prevalence of myofascial pain in general internal medicine practice. *West J Med*. 1989;151:157–160.

44. Wolfe F, Simons DG, Friction J, et al. The fibromyalgia and myofascial pain syndromes: a preliminary study of tender points and trigger points in persons with fibromyalgia, myofascial pain syndrome and no disease. *J Rheumatol*. 1992;19:944–951.
45. Conti PCR, dos Santos CN, Lauris JRP. Interexaminer agreement for muscle palpation procedures: the efficacy of a calibration program. *Cranio*. 2002;20:289–294.
46. Nice DA, Riddle DL, Lamb RL, et al. Intertester reliability of judgments of the presence of trigger points in patients with low back pain. *Arch Phys Med Rehabil*. 1992;73:893–898.
47. Duinkerke AS, Luteijn F, Bouman TK, et al. Reproducibility of a palpation test for the stomatognathic system. *Community Dent Oral Epidemiol*. 1986;14:80–85.
48. John MT, Dworkin SF, Mancl LA. Reliability of clinical temporomandibular disorder diagnoses. *Pain*. 2005;118:61–69.
49. List T, John MT, Dworkin SF, et al. Recalibration improves inter-examiner reliability of TMD examination. *Acta Odontol Scand*. 2006;64:146–152.
50. Stockstill JW, Gross AJ, McCall WD Jr. Interrater reliability in masticatory muscle palpation. *J Craniomandib Disord*. 1989;3:143–146.
51. 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;22:944–952.
52. Wahlund K, List T, Dworkin SF. Temporomandibular disorders in children and adolescents: reliability of a questionnaire, clinical examination, and diagnosis. *J Orofac Pain*. 1998;12:42–51.
53. Ohrbach R, Gale EN. Pressure pain thresholds, clinical assessment, and differential diagnosis: reliability and validity in patients with myogenic pain. *Pain*. 1989;39:157–169.
54. Schmitter M, Ohlmann B, John MT, et al. Research diagnostic criteria for temporomandibular disorders: a calibration and reliability study. *Cranio*. 2005;23:212–218.
55. John MT, Zwijnenburg AJ. Interobserver variability in assessment of signs of TMD. *Int J Prosthodont*. 2001;14:265–270.
56. Rosomoff HL, Fishbain DA, Goldberg M, et al. Physical findings in patients with chronic intractable benign pain of the neck and/or back. *Pain*. 1989;37:279–287.
57. Al-Shenqiti AM, Oldham JA. Test-retest reliability of myofascial trigger point detection in patients with rotator cuff tendonitis. *Clin Rehabil*. 2005;19:482–487.
58. Bron C, Franssen J, Wensing M, et al. Interrater reliability of palpation of myofascial trigger points in three shoulder muscles. *J Man Manip Ther*. 2007;15:203–215.
59. Gerwin RD, Shannon S, Hong CZ, et al. Interrater reliability in myofascial trigger point examinations. *Pain*. 1997;69:65–73.
60. Hsieh CYJ, 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;81:258–264.
61. Lew PC, Lewis J, Story I. Inter-therapist reliability in locating latent myofascial trigger points using palpation. *Man Ther*. 1997;2:87–90.
62. Njoo KH, Van Der Does E. The occurrence and inter-rater reliability of myofascial trigger points in the quadratus lumborum and gluteus medius: a prospective study in non-specific low back pain patients and controls in general practice. *Pain*. 1994;58:317–323.
63. Sciotti VM, Mittak VL, DiMarco L, et al. Clinical precision of myofascial trigger point location in the trapezius muscle. *Pain*. 2001;93:259–266.
64. McKenzie M, Taylor L, Cummings GS, et al. Prevalence of muscle trigger points in children with cerebral palsy. *Phys Ther Occup Ther Pediatr*. 1997;17:47–59.
65. Lantz CA, Nebenzahl E. Behavior and interpretation of the k statistic: a resolution of the two paradoxes. *J Clin Epidemiol*. 1996;49:431–434.
66. Wade B. Trigger points in the upper trapezius or normal subtrapezial anatomy? *Physiother Can*. 2001;53:219–222.
67. Quintner JL, Cohen ML. Referred pain of peripheral nerve origin: an alternative to the “myofascial pain” construct. *Clin J Pain*. 1994;10:243–251.
68. Gunn C. Radiculopathic pain: diagnosis and treatment of segmental irritation or sensitisation. *J Musculoskelet Pain*. 1997;5:119–134.
69. Bogduk N, Simons DG. Neck pain: joint pain or trigger points. In: Værøy H, Merskey H, eds. *Progress in Fibromyalgia and Myofascial Pain*. Amsterdam, The Netherlands: Elsevier; 1993:267–273.