Background: Chronic mid back and upper back pain caused by thoracic facet joints has been reported in 34% to 48% of the patients based on their responses to controlled diagnostic blocks. Systematic reviews have established moderate evidence for controlled comparative local anesthetic blocks of thoracic facet joints in the diagnosis of mid back and upper back pain.

Objective: To determine the diagnostic accuracy of thoracic facet joint nerve blocks in the assessment of chronic upper back and mid back pain.

Study Design: Systematic review of the diagnostic accuracy of thoracic facet joint nerve blocks.

Methods: A methodological quality assessment of included studies was performed using Quality Appraisal of Reliability Studies (QAREL). Only diagnostic accuracy studies meeting at least 50% of the designated inclusion criteria were utilized for analysis. Studies scoring less than 50% are presented descriptively and critically analyzed.

The level of evidence was classified as good, fair, and limited (or poor) based on the quality of evidence developed by the United States Preventive Services Task Force (USPSTF). Data sources included relevant literature identified through searches of PubMed and EMBASE from 1966 to March 2012, and manual searches of the bibliographies of known primary and review articles.

Outcome Measures: Controlled placebo or local anesthetic blocks were utilized using at least 50% pain relief as the reference standard.

Results: Three studies were identified utilizing controlled comparative local anesthetic blocks, with ≥50% pain relief as the criterion standard. The evidence is good for the diagnosis of thoracic pain of facet joint origin with controlled diagnostic blocks.

Limitations: The limitations of this systematic review include a paucity of literature for the diagnosis of thoracic facet joint pain, with all included manuscripts originating from one group of authors.

Conclusions: Based on this systematic review, the evidence for the diagnostic accuracy of thoracic facet joint nerve injections is good.

Key words: Chronic thoracic pain, mid back or upper back pain, thoracic facet or zygapophysial joint pain, facet joint nerve blocks, medial branch blocks, controlled comparative local anesthetic blocks
Among chronic pain disorders, pain arising from various structures of the spine constitutes the majority of problems, with a reported lifetime prevalence of spinal pain of 54% to 80% (1-3). However, the proportion of patients suffering from chronic upper or mid back pain secondary to thoracic disorders is relatively small, specifically in interventional pain management settings, ranging from 3% to 22% (1,2,4-13). Linton et al (7,11) estimated the prevalence of thoracic pain in 15% of the general population in contrast to 56% in the low back and 44% in the neck. Leboeuf-Yde et al (1) reported that low back pain in the past year was most frequent in 43%, followed by neck pain in 32%, followed by mid back pain in 13%. Leboeuf-Yde et al (13) also showed that regardless of the area of the complaint, care seeking and reduced physical activities were the most commonly reported consequences. The role of thoracic facet joints as a cause of chronic upper or mid back pain has received very little attention with only a few publications discussing these joints as the source of pain (6,14-20). Even though the description of the involvement of thoracic facet joints as a cause of chronic mid back and upper back pain dates back to 1987 (12), thoracic facet joint pain patterns were not described until 1994 and 1997 by Dreyfuss et al (19) and Fukui et al (20). Subsequent multiple studies have described thoracic facet joints as the source of chronic pain in 34% to 48% of patients with chronic mid back and upper back pain (6,14-18).

Based on postulates of Bogduk (21), thoracic facet joints have been shown to have an abundant nerve supply (19,20,22-30); shown to be capable of causing pain similar to that seen clinically, in normal volunteers with persistent mid back and upper back pain and referred pain into the chest wall (19,20); been shown to be affected by osteoarthritis, rheumatoid arthritis, spondylitis, degeneration, inflammation, and injury leading to pain upon joint motion and restriction of motion (31-33); and to be a source of pain in patients, using diagnostic techniques of known reliability and validity (6,14,15).

Conventional clinical and radiologic techniques are unreliable in diagnosing facet or zygapophysial joint pain (3,16-19,34-46). Consequently, controlled local anesthetic blocks of thoracic facet joints or medial branch blocks are employed to diagnose facet joint pain (30). The rationale is that anesthetic blockade of a painful joint will abolish pain arising from the joint for the duration of the anesthetic effect, while anesthetic blockade of a non-painful joint will not alter the pain report. The probability that the blocked joint is the actual source of pain is increased if repeating the block with an anesthetic agent that has a different duration of action reproduces the analgesic response (34-36,39). To ensure accuracy and validity, these blocks must be controlled and verified for delivery of a local anesthetic agent and placebo response (34-36,39). Either placebo controlled or comparative local anesthetic blocks are employed to eliminate placebo responses (42,44). Single facet joint injections are not recommended, as they do not control for a false-positive response, even though some have advocated therapeutic interventions without any diagnostic blocks (35-48). Moreover, accuracy of facet joint nerve blocks has been demonstrated with long-term follow-up (49-58). Despite all of the issues, including fraud and abuse, facet joint interventions, including thoracic facet joint interventions, have become more prevalent along with other interventional techniques (59-69).

This systematic review is undertaken to determine the accuracy of thoracic facet joint blocks in the diagnosis of chronic mid back and upper back pain. This is an update of a 2008 systematic review (16).

1.0 METHODS

The methodology utilized in this systematic review followed the review process derived from evidence-based systematic reviews and meta-analysis of diagnostic accuracy studies (41-44,46,70-78).

1.1 Criteria for Considering Studies for This Review

1.1.1 Types of Studies

Diagnostic accuracy studies evaluating thoracic facet joint pain.

1.1.2 Types of Participants

Participants of interest were adults aged at least 18 years with chronic upper and mid back pain of at least 3 months duration.

Participants must have failed previous pharmacotherapy, exercise therapy, etc., prior to starting diagnostic interventional pain management techniques.

1.1.3 Types of Interventions

The interventions were diagnostic thoracic facet joint blocks appropriately performed with proper technique under fluoroscopic or CT guidance.
1.1.4 Types of Outcome Measures

- The primary outcome parameter was pain relief concordant with the type of controlled diagnostic blocks performed.
- The secondary outcome measures were the ability to perform previously painful movements without significant pain or complications.
- At least 2 of the review authors independently, in an unblinded standardized manner, assessed the outcomes measures. Any disagreements between reviewers were resolved by a third author and consensus.

1.2 Literature Search

Searches were performed from the following sources without language restrictions:

1. PubMed from 1966
2. EMBASE from 1980
   www.embase.com/
3. Cochrane Library
   www.thecochranelibrary.com/view/0/index.html
   www.guideline.gov/
5. Previous systematic reviews and cross references
6. Clinical Trials
   clinicaltrials.gov/

The search period was from 1966 through March 2012.

1.3 Search Strategy

The search strategy emphasized chronic thoracic pain of facet joint origin with a focus on all types of diagnostic interventions. Search terminology included thoracic facet joint, thoracic facet joint pain, thoracic diagnostic facet joint blocks, thoracic facet joint intraarticular injections, and medial branch blocks.

This systematic review focused only on diagnostic studies, including invasive techniques and reports of complications. Only thoracic facet joint injections performed under fluoroscopy or CT imaging techniques were evaluated. Interventional techniques performed blindly or using other identification modalities were excluded. All studies describing appropriate outcome evaluations with proper statistical evaluations were reviewed. Reports without appropriate diagnosis, non-systematic reviews, book chapters, and case reports were excluded.

At least 2 of the review authors independently, in an unblinded standardized manner, performed each search. Accuracy was confirmed by a statistician. All searches were combined to obtain a unified search strategy. Any disagreements between reviewers were resolved by a third author and consensus.

1.4 Data Collection and Analysis

The quality of each individual article used in this assessment was based on the Quality Appraisal of Reliability Studies (QAREL) checklist (Table 1) (71). This checklist has been validated and utilized in multiple studies.

Table 1. Quality Appraisal of Diagnostic Reliability (QAREL) checklist.

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

TOTAL

systematic reviews (72). Each study in the final sample of eligible manuscripts was assessed using a 12-item appraisal checklist designed to assess the quality and applicability of studies. The face validity of these checklists was established by consultation with methodology experts (71) and comparison with quality appraisal checklists used in other systematic reviews examining diagnostic reliability (79-84). This checklist was also developed in accordance with the Standards for Reporting Studies of Diagnostic Accuracy (STARD) (75), and the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) (75,76) appraisal tool. Studies were not given an overall numeric quality score; instead, each item was considered separately and graded as “yes,” “no,” “unclear,” or “not applicable.”

1.4.1 Selection of Studies
- In an unblinded standardized manner, 2 review authors screened the abstracts of all identified studies against the inclusion criteria.
- All articles with possible relevance were then retrieved in full text for comprehensive assessment of internal validity, quality, and adherence to inclusion criteria.

1.4.2 Inclusion and Exclusion Criteria

Inclusion Criteria
Prospective and retrospective studies published on the diagnosis of thoracic facet joint pain in patients with chronic pain of greater than 3 months duration were included for review. Only the studies utilizing controlled diagnostic blocks under fluoroscopy were included. The criterion standard for diagnosis of thoracic facet joint pain was at least greater than 50% pain relief for the duration of local anesthetic and ability to perform previously painful movements.

Exclusion Criteria
All non-clinical studies were excluded. Ultrasound guided injections, case reports, book chapters, non-evidence-based guidelines, letters, and expert opinions were also excluded.

1.4.3 Clinical Relevance
The clinical relevance of the included studies was evaluated according to 5 questions recommended by the Cochrane Back Review Group (Table 2) (85,86). Each question was scored as positive (+) if the clinical relevance item was met, negative (−) if the item was not met, and unclear (?) if data were not available to answer the question.

1.4.4 Methodological Quality or Validity Assessment
Each study was evaluated by at least 2 authors for stated criteria and any disagreements discussed with a third reviewer. Authors with a perceived conflict of interest for any manuscript were recused from reviewing the manuscript.

Only diagnostic accuracy studies meeting at least 50% of applicable inclusion criteria were included for analysis. Studies scoring less than 50% were reported descriptively with critical analysis.

1.4.5 Data Extraction and Management
Two review authors independently, in an unblinded standardized manner, extracted the data from the included studies. Disagreements were resolved by discussion between the 2 reviewers; if no consensus could be reached, a third author was called in to break the impasse.

1.4.6 Assessment of Heterogeneity
Analysis of the evidence was based on diagnostic criteria as follows: 1) blocks in which the reference stan-

Table 2. Clinical relevance questions.

<table>
<thead>
<tr>
<th>Question</th>
<th>P (+)</th>
<th>N (-)</th>
<th>U (unclear)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A) Are the patients described in detail so that one can decide whether they are comparable to those who are treated practice?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B) Are the interventions and treatment settings described in sufficient detail to apply its use in clinical practice?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C) Were clinically relevant outcomes measured and reported?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D) Is the size of the effect clinically meaningful?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E) Do the likely treatment benefits outweigh the potential harms?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

standard for diagnosis was between 50% to 79% pain relief with a single block; 2) blocks in which the reference standard for diagnosis was between 50% to 79% pain relief with dual blocks; 3) blocks in which the reference standard for diagnosis was between 80% to 100% pain relief with a single block; and 4) blocks in which the reference standard for diagnosis was between 80% to 100% pain relief with dual blocks, to reduce clinical heterogeneity.

1.4.7 Measurement of Treatment Effect in Data Synthesis (Meta-Analysis)

Data was summarized separately using meta-analysis when at least 5 studies per type of diagnostic criteria were available that met the inclusion criteria (e.g., single block, double blocks, and 50% to 80% relief).

The minimum acceptable relief was considered to be 50%; however, data were sub-analyzed for ≥ 80% and 50% to 79% relief as the cutoff threshold for a positive block during the performance of previously painful movements. Four separate diagnostic categories were evaluated (i.e., 50% to 79% relief as the cutoff threshold with single and dual blocks; and 80% to 100% relief as the cutoff threshold with single or dual blocks). For dual blocks, there had to have been concordant response with short-acting and long-acting local anesthetics, or placebo.

1.4.8 Integration of Heterogeneity

A meta-analysis was performed only if there were at least 5 studies meeting inclusion criteria for each variable.

Statistical heterogeneity was explored using univariate meta-regression (87).

1.5 Summary Measures

Summary measures included 50% to 79% or 80% to 100% pain relief with the capability of performing previously painful movements concordant with the duration of local anesthetic.

1.6 Analysis of Evidence

The analysis of the evidence was performed based on United States Preventive Services Task Force (USPSTF) criteria (88) as illustrated in Table 3, which has been utilized by multiple authors (43,45,70,88-90).

The analysis was conducted using 3 levels of evidence ranging from good, fair, and limited (or poor) (43,45,70,88-90).

At least 2 of the review authors independently, in an unblinded standardized manner, analyzed the evidence. Any disagreements between reviewers were resolved by a third author and consensus. If there were any conflicts of interest (e.g., authorship), those reviewers were recused from assessment and analysis.

1.7 Outcome of the Studies

Outcomes included the prevalence of thoracic facet joint pain and false-positive rate. Based on the above parameters, the reliability of the data derived from each study was assessed.

2.0 Results

Figure 1 shows the flow diagram and study selection of diagnostic accuracy studies of thoracic facet joint pain. Only 3 studies met the inclusion criteria evaluating the prevalence and false-positive rate of facet joint nerve blocks in the diagnosis of mid back and upper back pain (6,14,15). Table 4 shows excluded studies.

Table 3. Method for grading the overall strength of the evidence for an intervention.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
<td>Evidence includes consistent results from well-designed, well-conducted studies in representative populations that directly assess effects on health outcomes (at least 2 consistent, higher-quality RCTs or studies of diagnostic test accuracy).</td>
</tr>
<tr>
<td>Fair</td>
<td>Evidence is sufficient to determine effects on health outcomes, but the strength of the evidence is limited by the number, quality, size, or consistency of included studies; generalizability to routine practice; or indirect nature of the evidence on health outcomes (at least one higher-quality trial or study of diagnostic test accuracy of sufficient sample size; 2 or more higher-quality trials or studies of diagnostic test accuracy with some inconsistency; at least 2 consistent, lower-quality trials or studies of diagnostic test accuracy, or multiple consistent observational studies with no significant methodological flaws).</td>
</tr>
<tr>
<td>Limited or Poor</td>
<td>Evidence is insufficient to assess effects on health outcomes because of limited number or power of studies, large and unexplained inconsistency between higher-quality trials, important flaws in trial design or conduct, gaps in the chain of evidence, or lack of information on important health outcomes.</td>
</tr>
</tbody>
</table>

Adapted and modified from methods developed by U.S. Preventive Services Task Force (70,88,89).
2.1 Descriptive Characteristics

Descriptive characteristics of these studies is included in Table 5. All 3 studies (6,14,15) were performed by the same group, with utilization of the same methodology, with controlled comparative local anesthetic blocks with 80% pain relief based on the duration of local anesthetics with lidocaine administered first, followed by bupivacaine, and with ability to perform maneuvers which were painful prior to injection therapy, and also the duration of the relief with the second block exceeding the first block irrespective of the duration in hours, days, or months. These studies evaluated not only the prevalence but also false-positive rate with confidence intervals. There was no significant difference among the 3 studies with prevalence or false-positive rate. The selection criteria, inclusion, and exclusion criteria of the patients was the same in all 3 studies.

2.2 Clinical Relevance

Among the 3 studies assessed for clinical relevance (6,14,15), all studies met criteria with a score of 5. Table 6 illustrates the assessment of clinical relevance.

2.3 Methodological Quality Assessment

A methodological quality assessment of diagnostic accuracy studies meeting inclusion criteria was carried
Table 5. Descriptive characteristics of diagnostic thoracic facet joint interventions.

<table>
<thead>
<tr>
<th>Study/Methods</th>
<th>Participants</th>
<th>Intervention(s)</th>
<th>Outcome(s)</th>
<th>Result(s)</th>
<th>Conclusion(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manchikanti et al 2004 (6)</td>
<td>500 consecutive patients with chronic, non-specific spine pain</td>
<td>Controlled comparative local anesthetic blocks (1% lidocaine or 1% lidocaine followed by 0.25% bupivacaine).</td>
<td>80% pain relief with the ability to perform previously painful movements. The relief with bupivacaine to last longer than lidocaine.</td>
<td>The prevalence of facet joint pain in patients with chronic cervical spine pain was 55% (95% CI, 49% - 61%), with thoracic spine pain was 42% (95% CI, 30% - 53%), and in patients with lumbar spine pain was 31% (95% CI, 27% - 36%). The false-positive rate with single blocks with lidocaine was 63% (95% CI, 54% - 72%) in the cervical spine, 55% (95% CI, 39% - 78%) in the thoracic spine, and 27% (95% CI, 22% - 32%) in the lumbar spine.</td>
<td>Facet joints are clinically important spinal pain generators in a significant (42%) proportion of patients with chronic spinal pain, with a false-positive rate of 55%.</td>
</tr>
<tr>
<td>Manchikanti et al 2002 (14)</td>
<td>46 consecutive patients with chronic midback and upper back pain</td>
<td>Diagnostic facet joint nerve blocks using lidocaine 1%, initially followed by bupivacaine 0.5% on separate occasions, usually 3 to 4 weeks apart.</td>
<td>80% pain relief with the ability to perform previously painful movements. The relief with bupivacaine to last longer than lidocaine.</td>
<td>46 patients underwent single blocks with lidocaine and 36 of these patients, or 78%, were positive for facet joint pain, reporting a definite response. Confirmatory blocks with bupivacaine were performed in all patients who were lidocaine-positive, with 61%, or 48% (95% CI; 34% to 62%), of the total sample of the lidocaine-positive group, reporting a definite response with improvement in their pain.</td>
<td>Comparative local anesthetic blocks showed the prevalence of facet joint pain to be 48%, with single blocks carrying a false-positive rate of 58%.</td>
</tr>
<tr>
<td>Manchukonda et al 2007 (15)</td>
<td>500 consecutive patients with chronic facet or zygapophysial joint pain</td>
<td>Diagnostic blocks using 0.5 mL of 1% lidocaine per nerve. Patients with lidocaine positive results were further studied using 0.5 mL of 0.25% bupivacaine per nerve on a separate occasion.</td>
<td>80% pain relief with the ability to perform previously painful movements. The relief with bupivacaine to last longer than lidocaine.</td>
<td>Prevalence of facet joint pain was 39% in the cervical spine (95% CI, 32%-45%), 34% (95% CI, 22%-47%) in the thoracic pain; and 27% (95% CI, 22%-33%) in the lumbar spine. The false-positive rate with a single block in the cervical region was 45%, in the thoracic region was 42%, and in the lumbar region 45%.</td>
<td>Significant prevalence of facet joint pain in chronic spinal pain, with 34% prevalence and 42% false-positive rate.</td>
</tr>
</tbody>
</table>

Table 6. Clinical relevance of included studies.

<table>
<thead>
<tr>
<th>Manuscript Author(s)</th>
<th>A) Patient description</th>
<th>B) Description of interventions and treatment settings</th>
<th>C) Clinically relevant outcomes</th>
<th>D) Clinical importance</th>
<th>E) Benefits versus potential harms</th>
<th>Total Criteria Met</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manchikanti et al (6)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>5/5</td>
</tr>
<tr>
<td>Manchikanti et al (14)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>5/5</td>
</tr>
<tr>
<td>Manchukonda et al (15)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>5/5</td>
</tr>
</tbody>
</table>

+= positive; - = negative

out utilizing QAREL criteria as shown in Table 7. Studies achieving 50% or higher scores were included. Scores of 67% or higher were considered to be high quality, 50% to 66% were considered to be moderate quality, and studies scoring less than 50% were considered to be of poor quality and excluded.

All 3 eligible studies utilized ≥ 80% relief (6,14,15).

Table 7. Quality Appraisal of Diagnostic Reliability checklist.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Was the test evaluated in a spectrum of subjects representative of patients who would normally receive the test in clinical practice?</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2. Was the test performed by examiners representative of those who would normally perform the test in practice?</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>3. Were raters blinded to the reference standard for the target disorder being evaluated?</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>4. Were raters blinded to the findings of other raters during the study?</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>5. Were raters blinded to their own prior outcomes of the test under evaluation?</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6. Were raters blinded to clinical information that may have influenced the test outcome?</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7. Were raters blinded to additional cues, not intended to form part of the diagnostic test procedure?</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>8. Was the order in which raters examined subjects varied?</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>9. Were appropriate statistical measures of agreement used?</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>10. Was the application and interpretation of the test appropriate?</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>11. Was the time interval between measurements suitable in relation to the stability of the variable being measured?</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>12. If there were dropouts from the study, was this less than 20% of the sample.</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>TOTAL</td>
<td>10/12</td>
<td>10/12</td>
<td>10/12</td>
</tr>
</tbody>
</table>

Y=yes; N=no; U=unclear; N/A=not applicable


2.4 Meta-Analysis

Only 3 studies were available utilizing 80% or greater relief as criterion standard with dual blocks (6,14,15). Consequently, there was no meta-analysis performed.

2.5 Diagnostic Accuracy

Accuracy was established in 3 studies based on a false-positive rate of 42% to 58%. Confidence intervals (95% CI) ranged from 26% to 78% (Table 8). Results of a combination of 3 studies showed prevalence of 40% (95% CI of 33% to 48%) with dual blocks and a false-positive rate of 42% (95% CI of 33% to 51%) with a single block.

The prevalence was illustrated to be 34% to 48%. Confidence intervals (95% CI) ranged from 22% to 62% (Table 8). The combination of results of all 3 studies yielded a prevalence rate of 40% (with a 95% CI of 33% to 48%) and a false-positive rate of 42% (with a 95% CI of 33% to 51%).

2.6 Confounding Factors

Influence of psychological factors was evaluated in the diagnosis of thoracic facet joint pain in only one study (91). Based on this evaluation, the prevalence of facet joint pain in patients suffering with chronic upper or mid back pain involving thoracic facet joints was shown to be present in 40% (95% CI 18% to 62%) of
patients without psychopathology, whereas it was 31% (95% CI 16% to 47%) in patients with vs 37% (95% CI 19% to 54%) without major depression, 33% (95% CI 19% to 48%) versus 35% (95% CI 15% to 55%) in patients with or without generalized anxiety disorder, and 36% (95% CI 7% to 65%) versus 33% (95% CI 21% to 46%) in patients with or without somatization disorder without any significant differences between the patients with psychological disorders and without psychopathology. However, due to small numbers in the study, there was a wide variation in 95% confidence intervals. This report is not considered conclusive with regards to the influence of psychological factors. Sedation as a confounding factor was evaluated in the cervical and lumbar spine (92-95). However, no such studies were available in the thoracic spine.

### 2.7 Analysis of Evidence

The evidence was synthesized based on the relief criteria when thoracic facet joint nerve blocks were performed.

There were a total of 3 studies meeting the inclusion criteria (6,14,15).

Using between 80% and 100% pain relief with dual blocks as the criterion standard has been advocated by some as the most rigorous means for diagnosing thoracic facet joint pain (16). All the studies evaluating prevalence and false-positive rates of facet joint pain showed a prevalence of 34% (15), 42% (6), and 48% (14); with false-positive rates of 44% (15), 55% (6), and 58% (14).

### 2.8 Level of Evidence

Based on the USPSTF criteria, the evidence was classified to be either good, fair, and limited (or poor).

Based on the 3 studies (6,14,15) with 80% or greater relief, with all 3 studies being of high quality, the evidence is good.

### 3.0 Discussion

This systematic review implicated thoracic facet joints as the source of chronic pain in 34% to 48% of patients with chronic mid back and upper back pain based on response to controlled diagnostic blocks of these joints (6,14,15). Based on this systematic review, false-positive rates of single local anesthetic blocks have been shown to range from 42% to 58%. The combined results of all 3 studies yielded a prevalence rate of 40% (95% CI, 33%-48%) and a false-positive rate of 42% (95% CI, 36%-53%) which may be defined as narrow confidence intervals both for prevalence as well as for false-positive rate.

This systematic review found good evidence for diagnostic accuracy of thoracic facet joint blocks.

The diagnostic thoracic facet joint blocks have been shown to be valid. The rationale for diagnostic blocks of the facet or zygapophysial joint(s) by blocking the nerve supply with an intraarticular injection of local anesthetic or by the blockade of the medial branches of the dorsal rami that innervate the target joint is based on the belief that one must test to determine whether a particular joint is the source of the pain. The rationale for using thoracic facet joint blocks for diagnosis is based on the fact that facet joints are capable of causing pain and that they have a nerve supply (6,12,14,15,19-27,31-35). Neuroanatomic studies have demonstrated free and encapsulated nerve endings in facet joints, as well as nerves containing substance P and calcitonin gene-related peptide (39,96,97). Further, thoracic facet joints have been shown to be a source of pain in the upper back, mid back, and referred pain in the chest wall (12,19,20,39).
The diagnosis of facet joint pain by controlled local anesthetic blocks is considered as valid. Controlled diagnostic blocks with 2 local anesthetics with placebo control are the important means of confirming the diagnosis of facet joint pain. The face validity of thoracic medial branch blocks has been established by injecting small volumes of local anesthetic and contrast material onto the target points. Benefit of diagnostic blocks on pain processing have been described (98).

Construct validity of thoracic facet joint blocks is important to eliminate placebo effect as a source of confounding results and to secure true-positive results as with all other medial branch blocks in the spine (6,14,19,39-44). In addition, the hypothesis that testing a patient first with lidocaine and subsequently with bupivacaine provides a means of identifying that the placebo responses have been tested and proven (99,100). Furthermore, the value and validity of facet joint nerve blocks based on long-term follow-up has been evaluated extensively in the lumbar spine (49,50). Even though Cohen et al (47) published that in the lumbar spine there it was cost-effective to provide radiofrequency neurotomy without diagnostic blockade, the study has been illustrated to be associated with multiple flaws. In addition, therapeutic medial branch blocks have shown to provide significant relief over a period of 2 years (51,52,101), even though literature on radiofrequency neurotomy or other modalities of treatments is scarce in thoracic spine (10,23,102,103). In addition, there has not been any further diagnostic literature since publication of our 2008 systematic review (16).

No tissue diagnosis (biopsy or autopsy) techniques are available to diagnose facet joint pain and confirm specificity and sensitivity of diagnostic blocks. However, pain relief and stability of the diagnosis with long-term follow-up are employed as the criterion standards and are accepted across different medical disciplines (49,50,104-106). Long-term relief of facet joint interventions has been demonstrated (16,37,38,45,49,57,101,104,107,108).

Thoracic medial branch blocks may be the only means available to diagnose thoracic facet joint pain, as there are no specific markers to diagnose facet joint pain in any region, specifically the thoracic region (6,14-16,39). Conventional clinical and radiologic techniques are unreliable in diagnosing facet or zygapophysial joint pain and various patterns of referred pain described for facet joints in the spine are similar to other structures, such as discs. Moreover, most maneuvers of physical examination are difficult to perform in the thoracic spine and such maneuvers are likely to stress several structures simultaneously, thus failing to provide any reasonable diagnostic criteria. The evidence thus far on physical examination and diagnosis has been controversial.

However, the major disadvantage of assessment of diagnostic utility of thoracic facet joint blocks appears to be that all the evidence is derived from one group of authors, even though methodological quality assessment is high and 95% confidence intervals are low.

Complications of thoracic facet joint nerve blocks are minor; however, serious complications can occur during this procedure related to the technique, as well as injection of other agents, specifically with relationship to the radicular artery which may be punctured with a poor technique (6,8,14,109-132).

This systematic review once again illustrates good evidence for the accuracy of thoracic facet joint nerve blocks in the diagnosis of chronic mid back and upper back pain based on a response to controlled diagnostic blocks. The review is based on 3 high quality studies, with the limitation that all the studies were performed by the same group of authors.

4.0 Conclusion

Diagnostic thoracic facet joint nerve blocks are safe, valid, and reliable. Based on the review of available studies that met inclusion criteria, the strength of evidence for diagnostic facet joint nerve blocks is good.

Acknowledgments

The authors wish to thank Vidyasagar Pampati, MSc, for statistical assistance, Sekar Edem for assistance in the search of the literature, Alvaro F. Gómez, MA, for manuscript review, and Tonie M. Hatton and Diane E. Neihoff, transcriptionists, for their assistance in preparation of this manuscript. The authors also thank Pain Physician for permission to reproduce Atluri et al’s manuscript from 2008. We would also like to thank the editorial board of Pain Physician for review and criticism in improving the manuscript.
Diagnostic Accuracy of Thoracic Facet Joint Nerve Blocks

References


www.painphysicianjournal.com


66. Manchikanti L, Singh V, Boswell MV. Interventional pain management at cross-
Diagnostic Accuracy of Thoracic Facet Joint Nerve Blocks


91. Bosczyk BM, Bosczyk AA, Putz R, E495


