Systematic Review

An Update of the Appraisal of the Accuracy of Thoracic Discography as a Diagnostic Test for Chronic Spinal Pain

Vijay Singh, MD1, Laxmaiah Manchikanti, MD2, Obi Onyewu, MD3, Ramsin M. Benyamin, MD4, Sukdeb Datta, MD5, Stephanie Geffert, MLIS6, Allan T. Parr, MD7 and Frank J.E. Falco, MD8

Background: Even though the prevalence of thoracic pain has been reported to be 13% of the general population and up to 22% of the population in interventional pain management settings, the role of thoracic discs as a cause of chronic thoracic and extrathoracic pain has not been well studied. The intervertebral discs, zygapophysial or facet joints, and other structures including the costovertebral and costotransverse joints have been identified as a source of thoracic pain.

Study Design: A systematic review of provocation thoracic discography.

Objective: To systematically assess and update the quality of clinical studies evaluating the diagnostic accuracy of provocation thoracic discography.

Methods: A systematic review of the literature was performed to assess the diagnostic accuracy of thoracic discography with respect to chronic, function limiting, thoracic or extrathoracic pain.

The available literature on thoracic discography was reviewed. A methodological quality assessment of included studies was performed using Quality Appraisal of Reliability Studies (QAREL).

The level of evidence was classified as good, fair, and limited (or poor) based on the quality of evidence developed by the U.S. Preventive Services Task Force (USPSTF).

Data sources included relevant literature identified through searches of PubMed and EMBASE from 1966 to June 2012, and manual searches of the bibliographies of known primary and review articles.

Results: The evidence and clinical value of thoracic provocation discography is limited (poor) with a paucity of evidence, with only 2 studies meeting inclusion criteria.

Limitations: The limitation of this study continues to be the paucity of literature.

Conclusion: Based on the available evidence for this systematic review, due to limited evidence, thoracic provocation discography is rarely recommended for the diagnosis of discogenic pain in the thoracic spine, if conservative management has failed and facet joint pain has been excluded.

Key words: Thoracic pain, chest wall pain, intervertebral disc, thoracic intervertebral disc, thoracic disc herniation, discogenic pain, thoracic provocation discography, false-positive response, diagnostic accuracy

Pain Physician 2012; 15:E757-E776
The lifetime prevalence of spinal pain has been reported as 54% to 80%; however, the proportion of patients suffering from chronic upper or mid back pain secondary to thoracic disorders is relatively small. The most common spinal regions studied are the lumbar and cervical spine, due to their strong and well-defined associations with pain conditions, work-related injuries, intervertebral disc degenerations, headaches, psychosocial disturbances, and expenses associated with managing these problems (1). It is not only that the thoracic spine has received less attention in terms of clinical, genetic, and epidemiologic research compared to the lumbar and cervical spine, it also has been assumed that the low prevalence of thoracic spinal pain is secondary to relative immobility in support of the thoracic region in contrast to other regions of the spine (1-7). However, pain experienced in the thoracic spine can be equally disabling, imposing similar burdens on the individual, community, and workforce (8-11). The spinal pain experienced in the region of the thoracic spine may arise from a number of sources including thoracic and cervical spinal structures, the thorax, and the gastrointestinal, cardiopulmonary, and renal systems (12-16). In addition, the thoracic spine is a common site for inflammatory, degenerative, metabolic, infective, and neoplastic conditions which may also contribute to pain and disability (15). Even though the limited research on prevalence and risk factors for thoracic spinal pain likely reflects the belief that the clinical and public health significance of thoracic spinal pain is less compared to other spinal levels, it has been argued that thoracic spinal pain should be considered as a discrete and important clinical entity, independent of pain experienced in other areas of the spine (11). In addition, in young adults, thoracic spinal pain is common and disabling with increasing incidence with age during adolescence (11,17). Even though the majority of the thoracic spinal pain and dysfunction may be associated with vertebral fractures (18-21), and hyperkyphosis arising from vertebral bone loss (22), ankylosing spondylitis (23), osteoarthritis (24), and Scheuermann’s disease (25), a significant proportion of patients with thoracic pain may suffer with degenerative disorders of the thoracic spine (26). Similar to the lumbar spine, degenerative signs identified in the thoracic spine with imaging modalities are not necessarily associated with pain, suggesting that non-specific thoracic spinal pain is highly prevalent (12,27-30). The emerging evidence suggests that thoracic spinal pain significantly impacts function (1-11,16,29,30). This is also evidenced by the proportion of patients presenting to interventional pain management settings with thoracic pain (31-41), which shows a variable presence of 3% to 22%. Linton et al (5) estimated that pain in the general population has a prevalence of 15% in the thoracic area, in contrast to 56% in the low back, and 44% in the neck. Briggs et al (3) in a systematic review of 33 studies meeting the inclusion criteria evaluated prevalence, incidence, and associated factors of thoracic spinal pain in the general population. They showed that thoracic spinal pain was significantly associated with concurrent musculoskeletal pain, along with the following factors: growth, physical, lifestyle, social, backpack, postural, psychological, and environmental. The biopsychosocial association was limited in the available literature. This systematic review showed that thoracic spinal prevalence data ranged from 3.5% to 34.8% of one-year prevalence and 15.6% to 19.5% of lifetime prevalence with a point prevalence of 4% to 72%. In another literature review, Briggs et al (4) published the results of prevalence and associated factors for thoracic spine pain in the adult working population. They identified 52 studies. Prevalence varied with the occupational group and time period. The one-year prevalence of thoracic spinal pain ranged from 3% to 55%, with most occupational groups having medians around 30%. Leboeuf-Yde et al (2,16) reported the prevalence of mid back pain at 13% with significant impact on quality of life.

Even though the thoracic spine has not been studied specifically, it appears that pain generators follow the same pattern as the lumbar spine. Kuslich et al (42) identified intervertebral discs, facet joints, ligaments, fascia, muscles, and nerve root dura as tissues capable of transmitting pain in the low back. Similarly, chronic thoracic or chest wall pain may also be transmitted by intervertebral discs, facet joints, ligaments, fascia, muscles, and nerve root dura, the tissues capable of transmitting pain in the mid back and upper back (33). Chronic, persistent thoracic and chest wall pain, and rare radicular pain may be secondary to disc herniation, discogenic pain, spinal stenosis, or post thoracic surgery syndrome. Furthermore, thoracic facet joints have been shown to be responsible for a significant proportion of pain in the thoracic spine (32,34,39,40). A diagnosis of the structures causing pain is crucial in providing appropriate treatment. Recent health care policy decisions have focused on increasing interventions in managing spinal pain (43-61). However, while facet joint pain has been shown to be significant among a proportion
Thoracic Discography in Chronic Mid-Back Pain

of patients responding to epidural steroid injections after failure of facet joint pain, symptomatic thoracic disc herniation is an uncommon condition, accounting for approximately 5 of every 1,000 disc herniations encountered in the clinical setting (60). The majority of thoracic disc herniations are asymptomatic (27), with radicular chest pain being the most common presenting complaint. Overall, very few patients require invasive treatment and most are conservatively treated, returning to their prior level of activity (62). However, a small proportion of cases require interventional techniques or surgical interventions, similar to the cervical and lumbar spine (54-59). Controlled diagnostic interventions have been described in the cervical and lumbar regions with good evidence in the diagnosis of facet joint pain and sacroiliac joint pain, and fair evidence in the diagnosis of lumbar discogenic pain (32-34,40,41,63-71). However, with studies in the thoracic spine studies being very few and far between, the diagnosis of thoracic facet joint pain appears to be fair (40), whereas that of discogenic pain with provoked discography, diagnosis was poor with limited evidence (41).

Provocation discography as performed today was first described in 1948 by Lindblom (72) when he used the term “diagnostic disc puncture.” This procedure provisionally replaced oil-contrast myelography described by Dandy (73) in 1929 for the diagnosis of a herniated disc as a cause of radicular pain. During the “herniated disc” era, both axial and referred radicular pain was thought to be due to a herniated disc compressing neural elements (71). It is well known that Mixter and Barr (74) were the first to create widespread interest in the disc as a source of pain in American literature with their 1934 hallmark description of the herniated nucleus pulposus. However, soon after, Mixter and Ayers (75) in 1935 demonstrated that radicular pain could occur without disc herniation. Since then, numerous investigators (27,71,72,76-120) have described provocation syndromes emanating from intervertebral discs that are not associated with evidence of the mechanical compression of neural structures. Consequently, internal disc derangement without a specific disc herniation has assumed a major role as a cause of non-specific spinal pain.

Discography has mainly been used as an imaging tool over the years and has been considered to be superior to radiographs, myelography, magnetic resonance imaging (MRI), and computed axial tomography (CT) scanning in imaging intervertebral disc morphology (41,43,68-71,82,109,110). Advances in CT and MRI scanning have added to the knowledge of disc pathology, structural abnormalities such as degenerative disc changes, herniations, associated end plate changes, and annular tears. There are no definitive tests for the diagnosis of discogenic pain, even in the lumbar and cervical spine. Structural abnormalities are present in patients asymptomatic of spinal pain, thus increasing the importance of discography as the most specific and sensitive test to assess if a disc is painful (41,68,69,100,104-107,112-114). Discography continues to be the criterion standard (96,98,110,111) to determine whether or not a particular disc is painful, irrespective of the evidence or lack thereof for degenerative changes utilizing other imaging modalities. The appropriate performance and diagnostic value of lumbar discography, and, to a somewhat lesser extent, cervical discography, has been extensively documented, practiced, and refined over the past 6 decades since its first descriptions in the 1940s. However, thoracic discography continues to be in its nascent stages of clinical application, specifically in the arena of evidence-based medicine, with the first descriptions of thoracic discography appearing in 1975 (87), approximately 30 years after the description of lumbar discography (72).

In 1975, Simmons and Segil (87) described thoracic discography and nucleography in the evaluation of a man with mid-thoracic radicular pain with a diagnosis of a posterior annular tear that reproduced his thoracic symptoms. In 1994, Schellhas et al (88) published a retrospective review of 100 outpatient thoracic discographies performed on patients whose MRI findings revealed thoracic disc degeneration. In 1999, Wood et al (90) published a prospective study of MRI and thoracic discography in the evaluation of thoracic disc herniations and annular tears. There are no definitive tests for the diagnosis of discogenic pain, even in the lumbar and cervical spine. Structural abnormalities such as degenerative disc changes, herniations, associated end plate changes, and annular tears. There are no definitive tests for the diagnosis of discogenic pain, even in the lumbar and cervical spine. Structural abnormalities are present in patients asymptomatic of spinal pain, thus increasing the importance of discography as the most specific and sensitive test to assess if a disc is painful (41,68,69,100,104-107,112-114). Discography continues to be the criterion standard (96,98,110,111) to determine whether or not a particular disc is painful, irrespective of the evidence or lack thereof for degenerative changes utilizing other imaging modalities. The appropriate performance and diagnostic value of lumbar discography, and, to a somewhat lesser extent, cervical discography, has been extensively documented, practiced, and refined over the past 6 decades since its first descriptions in the 1940s. However, thoracic discography continues to be in its nascent stages of clinical application, specifically in the arena of evidence-based medicine, with the first descriptions of thoracic discography appearing in 1975 (87), approximately 30 years after the description of lumbar discography (72).

The Task Force on Taxonomy of Classification of Chronic Pain in 1994 described criteria for the diagnosis of discogenic pain (112-114). The Task Force (112) defined thoracic discogenic pain as thoracic spinal pain, with or without referred pain. The key diagnostic criteria of thoracic discogenic pain is that the patient's pain must be shown conclusively to stem from an intervertebral disc by provocation discography of the putatively symptomatic disc that reproduces the patient's accustomed pain, and with provocation of at least 2 adjacent intervertebral discs that clearly do not reproduce the patient's pain, and provided that the pain...
cannot be ascribed to some other source innervated by the same segments that innervate the putatively symptomatic disc. The Task Force (112) cautioned that thoracic discography alone is insufficient to conclusively establish a diagnosis of discogenic pain because of the propensity for false-positive responses, either because of apprehension on the part of the patient or because of the coexistence of a separate source of pain within the segment under investigation.

Degeneration of the thoracic disc, along with end-plate irregularities and changes due to osteophyte formation, are common findings (99-101,116). Three systematic reviews evaluating the role of provocation discography in the diagnosis of spinal pain have presented limited evidence supporting the role of discography in identifying the subset of patients with thoracic discogenic pain (41,110,111). Furthermore, multiple concerns have been raised in regard to the reported high false-positive rate, the lack of concordance, potential confounding factors, and safety of controlled diagnostic blocks (68-71,76-78,110,111,116-120). In a systematic review of lumbar provocation discography in asymptomatic subjects with a meta-analysis of false-positive rates, Wolfer et al (71), after extensive evaluation, concluded that the strength of evidence is Level II-2 based on the USPSTF criteria (121) and the criteria for assessment of accuracy of diagnostic studies (122). Singh et al (41) in determining the accuracy of thoracic provocation discography concluded that the clinical value of thoracic provocation discography was limited Level II-3 with 2C/weak recommendation derived from low quality or very low quality evidence, indicating that other alternatives may be equally reasonable.

In this systematic review, we sought to update the current evidence of provocation discography (41) in the diagnosis of thoracic discogenic pain.

1.0 Methods

1.1 Definition and Criteria
The International Association for the Study of Pain (IASP) criteria (112) for thoracic discogenic pain includes reproduction of a patient’s typical pain with disc stimulation, with a failure to provoke pain in 2 adjacent intervertebral discs through injection. In addition, the pain cannot be ascribed to some other source innervated by the same segments that innervate the putatively symptomatic disc.

The methodology utilized in this systematic review followed the review process derived from evidence-based systematic reviews of diagnostic accuracy studies (123-136).

1.2 Criteria for Considering Studies for the Review

1.2.1 Types of Studies
Diagnostic accuracy studies of thoracic provocation discography.

1.2.2 Types of Participants
Participants of interest were adults aged at least 18 years with chronic mid and/or upper back pain and of at least 3 months duration. Participants must have failed previous pharmacotherapy, exercise therapy, etc., prior to discography.

1.2.3 Types of Interventions
The interventions were thoracic provocation discography.

1.2.4 Types of Outcome Measures
♦ The primary outcome parameter was pain provocation with or without control discs.
♦ At least 2 of the review authors independently, in an unblinded standardized manner, assessed the outcomes measures. A third author and consensus resolved any disagreements between reviewers.

1.3 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 June 2012.
1.4 Search Strategy
The search strategy emphasized chronic thoracic pain and diagnostic interventional techniques with special emphasis on provocation or discography.

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. A third author and consensus resolved any disagreements between reviewers.

1.5 Data Collection and Analysis
This systematic review focused only on provocation discography. The population of interest was patients suffering with chronic thoracic pain with or without chest wall pain for at least 3 months. Only the diagnostic accuracy of thoracic discography with respect to chronic thoracic pain was evaluated. Reports without appropriate diagnosis, non-systematic reviews, book chapters, and case reports were excluded.

The quality of each individual article used in this assessment was based on QAREL checklist (Table 1) (126). This checklist has been validated and also utilized in multiple systematic reviews (127). 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 (126) and comparison with quality appraisal checklists used in other systematic reviews examining diagnostic reliability (128-132). This checklist was also developed in accordance to the Standards for the Reporting Studies of Diagnostic Accuracy Studies (STARD) (124), and the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) (124,125) 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.”

1.5.1 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 a comprehensive assessment of internal validity, quality, and adherence to inclusion criteria.

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>

1.5.2 Inclusion and Exclusion Criteria

The following are the inclusion and exclusion criteria:

1. Are the patients described in sufficient detail to allow one to decide whether they are comparable to those who are treated in interventional pain management clinical practices?
   A. Setting – office, hospital, outpatient, inpatient
   B. Physician – interventional pain physician, general physician, anesthesiologist, physiatrist, neurologist, rheumatologist, orthopedic surgeon, neurosurgeon, etc.
   C. Patient characteristics - duration of pain
   D. Non-interventional techniques or surgical intervention in the past

2. Is the intervention described in sufficient detail to enable one to apply its use to patients in interventional pain management settings?
   A. Nature of intervention
   B. Frequency of intervention
   C. Duration of intervention

3. Were clinically relevant outcomes measured?
   A. Proportion of pain relief
   B. Disorder/specific disability
   C. Functional improvement
   D. Allocation of eligible and non-eligible patients to return to work
   E. Ability to work

1.5.3 Clinical Relevance

The clinical relevance of the included studies was evaluated according to 5 questions recommended by the Cochrane Back Review Group (Table 2) (133,134). Each question was scored 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.

1.5.4 Methodological Quality or Validity Assessment

Each study was evaluated by at least 2 authors for stated criteria. Any disagreements were 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% are reported descriptively with critical analysis.

1.5.5 Data Extraction & 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 planned third author was called in to break the impasse.

1.6 Analysis of Evidence

The analysis of the evidence was performed based on USPSTF criteria as illustrated in Table 3, which has been utilized by multiple authors (136,137).

The analysis was conducted using 3 levels of evidence ranging from good, fair, and limited or poor.

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

1.7 Outcome of the Studies

Outcome evaluations included the prevalence of thoracic discogenic pain and false-positive results.

Table 2. Clinical relevance questions.

<table>
<thead>
<tr>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are the patients described in detail so that one can decide whether they are comparable to those who are treated practice?</td>
</tr>
<tr>
<td>Are the interventions and treatment settings described in sufficient detail to apply its use to patients in interventional pain management settings?</td>
</tr>
<tr>
<td>Were clinically relevant outcomes measured and reported?</td>
</tr>
<tr>
<td>Is the size of the effect clinically meaningful?</td>
</tr>
<tr>
<td>Do the likely treatment benefits outweigh the potential harms?</td>
</tr>
</tbody>
</table>

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 (64-66,68-70,136,137).

Fig. 1. The flow diagram illustrating published literature evaluating thoracic provocation discography.

2.0 Results

Figure 1 shows a flow diagram of study selection. There were 2 studies considered for inclusion (88,90).

2.1 Clinical Relevance

Of the 2 studies assessed for clinical relevance, both studies met criteria scoring of 5 out of 5. Table 4 illustrates the assessment of clinical relevance.
2.2 Methodological Quality Assessment

A methodological quality assessment of diagnostic accuracy studies meeting inclusion criteria was carried out utilizing QAREL criteria as shown in Table 5. Studies achieving 50% or higher scores were included. Scores of 67% or higher were considered to be high quality, ≥50% were considered to be moderate quality, and studies scoring less than 50% were considered to be of poor quality and excluded.

There were 2 studies evaluating provocation discography (88, 90).

2.3 Study Characteristics

In 1994, Schellhas et al (88) published their experience with thoracic discograms performed on 100 outpatients by a retrospective analysis. After MRI, clinically suspect, morphologically abnormal thoracic discs and at least one nearby controlled level disc were injected with either non-ionic contrast or saline, filmed, and individually described by the patient as concordant versus non-concordant relative to clinical pain, and rated in pain intensity on a scale of 0 to 10. The results illustrated that discs with annular tears, intrinsic degeneration,
and vertebral body endplate infarctions were painful approximately 75% of the time. They demonstrated a clinical concordance of 50% with painless control levels. In this series, clinically concordant extraspinal pain such as chest wall, intrathoracic, and upper abdominal pain were frequently provoked with thoracic disc injections. They described non-protruding disc derangements such as may be seen either in active or old juvenile discogenic disease (Scheuermann’s disease). Internal disc derangements may be painful and clinically significant with more than 50% of the painful discs that they studied falling into this category. The authors concluded that thoracic discography can be performed safely by experienced individuals as a reliable tertiary diagnostic procedure to determine if degenerated discs on MRI studies are related to clinical complaints. The shortcomings of this evaluation include it being a retrospective evaluation. They described the technical aspects extensively, even though characteristics of patients’ pain patterns were not provided at baseline. Further, a consistent reference standard was not applied. There was no blinded comparison of the test.

Wood et al (90) performed a prospective evaluation. They sought to determine the responses to thoracic discography by asymptomatic and symptomatic individuals. They evaluated 10 adult lifelong asymptomatic volunteers, ages 23 to 45 years, who underwent MRI of the thoracic spine, by a 4-level discography. Provocation responses were graded on a scale of 0 (no sensation) to 10 (extreme pain or pressure), and filmed discs were graded using a modified Dallas scheme. Concomitantly, 10 non-litigious adults, ages 31 to 55 years, experiencing chronic thoracic pain were similarly studied. The results showed the mean pain responses in the asymptomatic volunteers to be 2.4/10. Three discs in the asymptomatic group were intensely painful with scores of 7/10, 8/10, and 10/10, with all 3 exhibiting prominent endplate irregularities and annular tears typical of thoracolumbar Scheuermann’s disease. On discography, 27 of 40 discs were abnormal, with endplate irregularities, annular tears, and/or herniations. They also reported that the 10 discs read as normal on MRI showed annular pathology on discography. In the group with chronic thoracic pain, the average pain response was 6.3/10 (P < 0.05). Of the 48 discs studied, 50% or 24 were concordantly painful, with a response of 8.5/10 (P < 0.05). Seventeen discs had non-concordant pain or pressure, with an average pain score of 4.8/10 (P < 0.05) and 5 had no response. On MRI, 21 of the 48 discs appeared normal, whereas on discography, only 10 were judged as normal. They concluded that on discography, thoracic discs with prominent Schmorl’s nodes may be intensely painful, even in lifelong asymptomatic individuals, but the pain is unfamiliar or non-concordant. They also concluded that thoracic discography may demonstrate disc pathology not seen on MRI.

Evidence was also provided for the relative lack of reliability of MRI at identifying painful deranged discs (90). They reported a high incidence of relatively painless disc pathology, including annular tears and frank herniations, with discography in both the symptomatic and asymptomatic patients that was missed on MRI. Further, they noted that a general trend toward more painful responses was being observed with greater degrees of pathology, especially with endplate pathology such as Scheuermann’s disease. Variability was reported in perceived pain or pressure, even though typically it was on the same side as the disc pathology, whether it was a tear or herniation.

This first ever controlled prospective study in asymptomatic and symptomatic individuals had some deficiencies (90). There were only 10 lifelong asymptomatic volunteers. While they concluded that thoracic discography in the truly asymptomatic individual is not painful, regardless of the degree of pathology observed, they reported 3 of the 40 discs (7.5%) as intensely painful with pain of 7, 8, and 10 on a scale of 0 to 10. However, the 3 of them exhibited prominent endplate changes typical of thoracolumbar Scheuermann’s pathology. Two of these painful responses were in one volunteer. Consequently, 20% of the asymptomatic volunteers reported pain when they had severe Scheuermann’s pathology. Once the 3 painful discs or 2 painful patients were removed, the average pain response was less than 2/10. Only one volunteer reported aching muscle-like pain for 48 hours, which resolved quickly at that point with no sequelae. The authors have not provided detailed results with regards to negative contiguous discs, one above and one below, thus, the criteria was limited to only elicitation of concordant pain. Twenty-seven of 49 or 55% of the discs studied in the symptomatic group were concordant.

2.4 Validity

Wood et al (90) evaluated the validity of the concordant pain and the role of false-positive responses. They reported the mean pain response in the asymptomatic volunteers as 2.4/10 even though 3 discs exhibiting prominent endplate irregularities and annular tears typical of thoracolumbar Scheuermann’s disease
were intensely painful. Furthermore, of the 48 discs studied, only 21 appeared normal on MRI and only 10 were judged as normal after provocation discography. The discs which exhibited concordant pain (24 of 48 or 50%) exhibited a pain response of 8.5/10, statistically higher pain levels than the 17 discs that exhibited non-concordant pain pressure with an average pain of 4.8/10, and 5 discs with no pain response at all.

Schellhas et al (88) evaluated concordant pain and also at least one nearby controlled level disc. They demonstrated clinical concordance in approximately 50% of the discs, with controlled levels being painless.

2.5 Prevalence
The prevalence of thoracic discogenic pain has not been determined.

2.6 False-positive Rates
Utilizing the data by Wood et al (90), it appears that the false-positive rate with thoracic discograms is 0 if a pain response of 7 or above is considered as positive with concordant pain with negative contiguous discs. When endplate irregularities and annular tears are taken into consideration as shown in the asymptomatic patients, even though the mean response in volunteers was 2.4/10, 3 discs in 2 patients were intensely painful with scores ranging from 7 to 10 of 10. Consequently, in patients with severe pathology, pain may be produced in 20% of the patients. Considering the clinical realities which dictate provocation thoracic discography to be performed only in symptomatic patients, utilizing the IASP criteria (112), and that these positive patients may have been dormant and fallen within the range of the prevalence of discogenic pain, it is considered that the false-positive rate with thoracic provocation discography is low.

Schellhas et al (88) evaluated concordant pain with a controlled disc at least at one level. They demonstrated a clinical concordance of 50% with painless control levels. They also concluded that they were able to determine whether observed disc pathology related to clinical pain complaints in every patient. In isolated cases in which patient uncertainty existed after thorough questioning about pain-pressure concordance at individual disc levels, the authors interpreted the response at that level as to be either indeterminate or non-concordant. They showed that in these cases, at least one other disc was clearly concordant; hence, the total examination was considered to be conclusive. In this study, they evaluated a total of 306 discs. Custom-arily they studied consecutive discs including more than one normal-appearing control level disc if necessary (138). Morphologically deranged thoracic discs produced more painful responses compared with normal-appearing control levels, even though these responses were not necessarily concordant relative to the pain being investigated (97,139,140). Based on the results of this study, it appears that false-positive rates are low when discography is performed appropriately using a concordance of pain and negative control discs.

Fluke (141) criticized the report on its definition of reliable, high degree of accuracy, and Schellhas et al’s conclusion that they were able to determine whether observed disc pathology related to clinical pain in every patient. Fluke contended that Schellhas et al failed to provide the data necessary to determine whether their techniques were accurate or not, because true-positive, true-negative, and false-negative rates were not reported.

In reply, Schellhas (142) referred to the formal prospective investigations of false-positive rates in the lumbar spine (140,143). Schellhas also pointed out that discography results do not provide an “excuse to operate.”

2.7 Analysis of Evidence
Based on the USPSTF criteria, the evidence is considered at 3 levels – good, fair, and limited or poor. The evidence based on this analysis is limited due to only 2 moderate quality studies with no recent literature available.

3.0 Discussion
This systematic review provides limited evidence for using provocation discography to identify patients with chronic thoracic discogenic pain. There are no prevalence or false-positive data available. Considering that thoracic facet joint pain is present in 34% to 48% of patients with chronic non-specific function-limiting mid back and upper back pain with false-positive rates of 42% to 58% with a single block (32,33,40), it appears that thoracic discogenic pain may be present in at least an equal proportion of patients. Chronic lumbar discogenic pain has been reported in 26% to 39% of patients regardless of internal disc disruption (80,81). The prevalence of chronic discogenic neck pain has been reported as 16% and possibly 41% of the patients (82). Singh et al (41) in a 2008 systematic review determined the accuracy of thoracic discography in an evaluation of chronic thoracic pain utilizing 2 studies (88,90) with
Thoracic Discography in Chronic Mid-Back Pain

evidence which was shown as limited or Level II-3 with a weak recommendation derived from low quality or very low quality studies, indicating that other alternatives may be equally reasonable. They also noted that the drawbacks of the evaluation on thoracic discography were that only 2 studies were available, both from the same group of authors, with the last study being published in 1999, and there have been no subsequent attempts by the same authors or others to replicate or confirm previously published results. There also was not any literature showing the effectiveness of therapeutic modalities based on the results of discography.

The same problems from the previous systematic review (41) were experienced in the present systematic review. There is continued paucity of the literature and methodological challenges in assessing the accuracy of thoracic provocation discography also continue. Furthermore, there have only been 2 studies evaluating the value of discography (88,90) and there have not been any new studies since 1999. Since the previous publication (41), however, 3 reports have been published showing the management techniques of thoracic discogenic pain with nucleoplasty, intradiscal biaculoplasty, and laser disc decompression (54,55,57). For percutaneous laser disc decompression for thoracic disc disease, patients underwent discography prior to percutaneous laser disc decompression. These are all case reports, with the largest one being the laser disc decompression (57) with 10 patients. These do not provide any substantial support to provocation discography; however, such emerging studies may in the future provide appropriate evidence.

The issues related to criterion or gold standard and methodologic challenges continue in this evaluation. For this evaluation we utilized new methodologic quality assessment criteria as developed by Lucas et al (126) that may be similar or superior to QUADAS and others.

There is no gold standard for discogenic pain. A concordantly painful disc with contiguous negative discs was considered as the gold standard for this systematic review. Both the studies (88,90) considered for inclusion in this systematic review were able to describe, in detail, concordant pain, but only one study (88) described a negative disc. The assessment of pain may be considered as a soft measure and has been challenged repeatedly (71,140,144-155). The patients’ ability to consistently report pain accurately during discography along with multiple confounding factors has been extensively discussed (71,156,157). There are no studies evaluating the outcomes based on surgical treatments derived from opinions of appropriately performed thoracic provocation discography. A gold standard of tissue biopsy, which is not reliable in the case of degenerative discs, may also be inappropriate as a criterion standard for provocation discography.

The evidence for thoracic discography consists of only 2 studies from the same group of authors. Fluke (141) in response to the study by Schellhas et al (88) wrote that the data presented in this article are not sufficient to support the conclusion that thoracic discography is a “reliable tertiary diagnostic procedure to determine if degenerated discs on MR studies are related to clinical complaints.” Fluke criticized that authors failed to define “reliable” as it relates to this study. Moreover, the study was criticized for not providing the data necessary to determine whether their techniques are accurate or not, because true-positive, true-negative, and false negative rates were not reported in this study. However, Fluke (141) felt that thoracic discography is technically feasible and probably safe without providing any significant evidence.

The basic principles for thoracic provocation discography are to determine whether or not a thoracic disc(s) is the source of a patient’s thoracic, chest wall, or upper abdominal pain. Thus, thoracic discography is performed in an attempt to provoke pain with each injection at each designated level. A local anesthetic can be used to anesthetize painful discs to further refine the identification of a concordantly painful disc. If thoracic pain is reproduced during thoracic discography, the disc or discs are responsible for at least part or all of the pain (112,115). If thoracic pain is not reproduced during thoracic discography, then the discs are excluded as potential pain generators. Disc stimulation is analogous to palpation for tenderness (112,145). The rationale for thoracic provocation discography is based on the fact that thoracic discs are innervated and therefore can elicit pain (97,115,158-165). Anatomical studies have demonstrated that intervertebral discs receive an innervation posteriorly from the sinuvertebral nerves, laterally from the vertebral nerves, and anteriorly from the sympathetic trunks (43,99,112,119,158-168). In addition, thoracic discs have been shown to cause chronic upper back and mid back pain (88,90). Discogenic pain has been described to be dull and aching in quality, whereas, neurogenic pain has been described to be lancinating in quality (86). Imaging studies such as radiographs, myelography, CT, CT-myelography, and MRI are inaccurate in determining whether a thoracic disc is responsible for a patient’s pain complaints or the
presence or absence of disc pathology (91). In addition, the patterns for thoracic discogenic pain are expected to be indistinguishable from those of thoracic facet joint pain, as in the lumbar and cervical regions (115-118). Even though these pain patterns can be used to indicate the most likely segmental source of pain and, therefore, the levels at which investigations should be focused, these patterns reflect the innervation of the source of the pain, and they do not implicate a particular structure as the source. Thus, it is essential to rule out thoracic facet joint pain prior to embarking on provocation discography. Physical examination will distinguish the source of pain because shear stress applied to the thoracic spine will simultaneously stress not only the disc, but also facet joints.

The criteria developed by IASP (112) have recommended that in order to be valid, thoracic provocation discography must be subjected to anatomical controls. Consequently, the 2 diagnostic criteria for discogenic pain must be met in each and every case. The first criterion is that the provocation of the target disc reproduces the patient’s pain. The second criterion is that provocation of adjacent discs does not reproduce the pain. Validity may also be enhanced by appropriate assessment of the pain including location, quality, intensity, and concordance.

The role of placebo and nocebo is also a factor in diagnostic accuracy and treatment effect and has been extensively discussed (169-187). Similarly, role of placebo design, and the effect of injecting various solutions into active and inactive structures may influence patient responses (188-197). Discography may also be associated with multiple adverse effects (198-219).

In summary, extensive research is not currently available regarding the various causes of thoracic pain and the diagnosis of those causes. Some studies (32,34,69) have focused on the prevalence of thoracic facet joint pain, but there are no studies evaluating the prevalence of thoracic discogenic pain. This systematic review provides limited (poor) evidence for thoracic discography as a diagnostic tool.

4.0 Conclusion

Based on the present systematic review the strength of evidence is limited (poor) based on the AHRQ USP-STF criteria for the diagnostic accuracy for discography, with a recommendation that the procedure is reserved for rare occasions.
Thoracic Discography in Chronic Mid-Back Pain

References


www.painphysicianjournal.com
Thoracic Discography in Chronic Mid-Back Pain

Pain Physician 2012; In Press.


101. Cloward RB. Anterior herniation of a ruptured lumbar intervertebral disk; comments on the diagnostic value of the diskogram. AMA Arch Surg 1952; 64:457-463.


191. Chu CR, Izzo NJ, Papas NE, Fu FH. In vitro exposure to 0.5% bupivacaine is cytotoxic to bovine articular chondrocytes. *Arthroscopy* 2006; 22:893-999.


