

Systematic Review

Systematic Review of Lumbar Provocation Discography in Asymptomatic Subjects with a Meta-analysis of False-positive Rates

Lee R. Wolfer, MD, MS¹, Richard Derby, MD^{1,2}, Jeong-Eun Lee, PT^{1,3},
and Sang-Heon Lee, MD, PhD^{1,3,4}

From: ¹Spinal Diagnostics and Treatment Center, Daly City, CA; ²Division of Physical Medicine and Rehabilitation, Stanford University Medical Center, Stanford, CA; ³Graduate School of Medicine, College of Medicine, Korea University, Seoul, Korea; ⁴Department of Physical Medicine and Rehabilitation, Korea University Medical Center, Seoul, Korea

Dr. Wolfer is with the Spinal Diagnostics and Treatment Center, Daly City, CA. Dr. Derby is Medical Director of the Spinal Diagnostics and Treatment Center, Daly City, CA; and Associate Clinical Professor, Division of Physical Medicine and Rehabilitation, Stanford University Medical Center, Stanford, CA.

Ms. Lee is a physical therapist with the Spinal Diagnostics and Treatment Center, Daly City, CA; and the Graduate School of Medicine, College of Medicine, Korea University, Seoul, Korea

Dr. Lee is with the Spinal Diagnostics and Treatment Center, Daly City, CA, and a professor at the Graduate School of Medicine, College of Medicine, Korea University, Seoul, Korea, and with the Department of Physical Medicine and Rehabilitation, Korea University Medical Center, Seoul Korea.

Address correspondence:
Lee Wolfer, MD, MS
Spinal Diagnostics and Treatment Center
901 Campus Drive, Suite 312
Daly City, CA 94015
E-mail: drwolfer@spinaldiagnostics.com

Disclaimer: There was no external funding in the preparation of this manuscript.

Conflict of interest: None.

Manuscript received: 04/30/2008
Revised manuscript received:
06/13/2008
Accepted for publication: 06/30/2008

Free full manuscript:
www.painphysicianjournal.com

Background: Lumbar provocation discography is a controversial diagnostic test. Currently, there is a concern that the test has an unacceptably high false-positive rate.

Study Design: Systematic review and meta-analysis.

Objective: To perform a systematic review of lumbar discography studies in asymptomatic subjects and discs with a meta-analysis of the specificity and false-positive rate of lumbar discography.

Methods: A systematic review of the literature was conducted via a PUBMED search. Studies were included/excluded according to modern discography practices. Study quality was scored using the Agency for Healthcare Research and Quality (AHRQ) instrument for diagnostic accuracy. Specific data was extracted from studies and tabulated per published criteria and standards to determine the false-positive rates. A meta-analysis of specificity was performed. Strength of evidence was rated according to the AHRQ U.S. Preventive Services Task Force (USPSTF) criteria.

Results: Eleven studies were identified. Combining all extractable data, a false-positive rate of 9.3% per patient and 6.0% per disc is obtained. Data pooled from asymptomatic subjects without low back pain or confounding factors, shows a false-positive rate of 3.0% per patient and 2.1% per disc. In data pooled from chronic pain patients, asymptomatic of low back pain, the false-positive rate is 5.6% per patient and 3.85% per disc. Chronic pain does not appear to be a confounding factor in a chronic low back pain patient's ability to distinguish between positive (pathologic) and negative (non-pathologic) discs. Among additional asymptomatic patient subgroups analyzed, the false-positive rate per patient and per disc is as follows: iliac crest pain 12.5% and 7.1%; chronic neck pain 0%; somatization disorder 50% and 22.2%, and, post-discectomy 15% and 9.1%, respectively. In patients with chronic backache, no false-positive rate can be calculated. Low-pressure positive criteria (≤ 15 psi a.o.) can obtain a low false-positive rate. Based on meta-analysis of the data, using the ISIS standard, discography has a specificity of 0.94 (95% CI 0.88 – 0.98) and a false-positive rate of 0.06.

Conclusions: Strength of evidence is level II-2 based on the Agency for Healthcare Research Quality (USPSTF) for the diagnostic accuracy of discography. Contrary to recently published studies, discography has a low false-positive rate for the diagnosis of discogenic pain.

Key words: Meta-analysis, lumbar discography, false-positive, asymptomatic subjects

Pain Physician 2008; 11:4:513-538

In 1929, Dandy (1) utilized oil-contrast myelography to describe "loose cartilage simulating a tumor of the spinal cord" (herniated disc) as a cause of radicular pain. Myelography was the standard diagnostic test for disc protrusions or herniations until the introduction of discography by Lindblom in the 1940s (2). During the "herniated disc" era, both axial and referred radicular pain were thought to be due to a herniated disc compressing neural elements. Initially, discography was used as an imaging test to demonstrate the structural morphology of disc protrusions or herniations; however, it also revealed anular disruption as a common topography of lumbar discs. More importantly, some of these degenerated discs with anular disruption were painful when injected with contrast, thus giving rise to the term provocation discography (3-6) (for brevity, the term discography is often used in this text, but the test is understood to be more than an imaging test). These observations led surgeons to use provocation discography not only to reveal structural abnormalities, but also to identify and treat painful discs.

Since discography was introduced, computed axial tomography (CT) and magnetic resonance imaging (MRI) scanning have also added to our knowledge of the lumbar disc, however, because structural abnormalities such as degenerative disc changes, herniations, and anular tears, occur in patients asymptomatic of low back pain (7,8), discography is our only direct method to assess if a disc is painful. Discography has also been shown to reveal abnormalities in symptomatic patients with normal MRI scans (9,10). Discography has, therefore, remained the criterion standard (11,12) to determine whether or not a particular disc is painful. Provocation discography is considered to be an extension of the physical examination. Since most structural disc abnormalities are not life threatening and the treatment of discogenic pain often involves an interventional or surgical procedure, the false-positive rate of a provocative test that relies on a subjective response of a patient with chronic pain is the primary contentious issue in an ongoing controversy regarding the true value of this diagnostic test.

Discography in asymptomatic subjects has been studied over the last 40 years (13-21). Concerns have been raised in regard to the reported high false-positive rate, the lack of concordance, potential confounding factors, and safety (16,17,22,23). To our knowledge, no prior publications have systematically reviewed and critically synthesized all the available

data, as well as reported confidence intervals, to arrive at a current evidence-based estimate of the false-positive rate as indirectly studied by performing discography on asymptomatic volunteer subjects. Additionally, no prior publication has investigated the diagnostic accuracy (specificity) of lumbar discography through meta-analysis of all published studies.

We use the standard definition defining a false-positive test as an erroneously positive test when the test is in fact negative. Statistically, this is considered a Type 1 error or an alpha error, whereby the null hypothesis is erroneously rejected. Ideally, a test has a reference standard (gold standard) to confirm the presence or absence of a disease. However, tissue confirmation for the presence of a painful degenerated disc is inaccurate due to similarities between the normally aging and painfully degenerating disc.

The goal with a diagnostic test is to set a decision threshold which strikes a balance between acceptable levels of false-positive and false-negative results. If the threshold for false-positives is set too high, there will be an unacceptable number of persons with a negative test who in fact have the index disease. Over the last 2 to 3 decades, discography techniques and criteria have been refined to meet this requirement. In this analysis, we evaluate and combine data from available published experimental studies investigating the false-positive rate of lumbar discography and test accepted criteria and standards. Walsh et al (15) introduced thresholds for pain intensity, pain behaviors, concordancy, and pressure limits combined with abnormal morphology to define a "positive discogram". Carragee et al (17) used criteria similar to Walsh, except for a higher pressure limit to 100 psi a.o. Derby (24) recommended pressure and speed-controlled manometry with a pressure limit of 50 psi a.o. based on studies of intra-discal pressure and pain in 150 patients with chronic lumbar pain. As discography is a provocative test, it inherits the major liability of all provocative tests which is that pain response is related to the intensity of stimulus. Derby (24) reported that opening disc pressure in side-lying with a normal nucleogram was 27 psi versus 17 psi in a disc with greater than 50% degeneration. Disc pressure was found to decrease with increasing degeneration. In degenerated discs, concordant pain provocation occurred within a 50% increase above the opening pressure (ratio 1:1.5). Overall, pain response was usually maximal at pressures only 10 to 30 psi above the opening pressure. Increasing pressure

increased pain intensity in most degenerated discs, including non-pathologic discs. Based on this research and other studies, the International Association for the Study of Pain (IASP) and the International Spine Intervention Society (ISIS) adopted a pressure limit of < 50 psi a.o. (25). Reasonable pressure limits must be set otherwise non-pathologic discs can be rendered painful with excessive pressurization.

Our analysis seeks to review a complex and contradictory body of literature and perhaps resolve contrary findings across studies. Evaluating the data based on various accepted criteria and standards will also help improve the diagnostic accuracy and set an appropriate decision threshold for provocation discography. Pooling data from individual studies with meta-analysis improves the precision of statistical conclusions. Ideally, knowing the percent false-positive rate of lumbar discography, based on the best standards for pain response and intensity of provocation, would allow the physician to give greater or lesser importance to the patient's response to disc stimulation when he or she is weighing the evidence to confirm or refute the hypothesis that a particular disc is the probable source of a patient's pain.

METHODS

Inclusion Criteria

The types of studies included were clinical studies of asymptomatic subjects or asymptomatic subject discs. One study of discography in subjects without significant low back pain illness was also included. Subjects may or may not have had a history of spine surgery. We searched for studies using modern discographic techniques which reported numerical ratings of pain intensity, concordancy, pain behaviors, pressure, degree of anular disruption, and a control disc. There were relatively few studies meeting this criteria. No randomized controlled trials have been performed on asymptomatic subjects to date.

Exclusion Criteria

Studies using older discographic techniques, including noxious dyes, were excluded from data analysis and synthesis. However, a description of the older studies is included for historical perspective. The following types of articles were also excluded: descriptive studies, expert opinion, review articles, technical papers, and non-clinical studies.

Search Strategy

Clinical research studies satisfying the inclusion criteria for the review were identified by a database search of PUBMED from January 1, 1960 to March 30, 2008. Key search terms were intervertebral disc, discography, discogram, false-positive, asymptomatic, normal, and intervertebral disc injection. The search was refined with Boolean operators (AND/OR). Limits applied were English language only, human studies, and adults. The references of each article were reviewed by hand to identify additional studies.

Method of Review

After the literature review, abstracts were obtained and examined for inclusion criteria. Full journal articles were obtained if the inclusion criteria were satisfied. Three physicians reviewed the articles. Data extraction was performed by 3 researchers (LW, RD, and SL). The primary data from the experimental studies was extracted as published per individual disc injection.

Methodological Quality

The quality of each article was scored according to the Agency for Healthcare Research and Quality (AHRQ) (Table 1) rating scale from 0 – 100 (26). Three physician reviewers scored the articles separately. Any disagreement was discussed until a consensus was reached. For inclusion in data analyses, the study had to score at least 45/100 on the AHRQ scale. Studies which scored below this threshold are described and critiqued separately.

Strength of Evidence

Quality of evidence was evaluated by the criteria developed by U.S. Preventive Services Task Force (USPSTF) in Table 2 (27).

Data and Statistical Analysis

Data from each study was reviewed according to various reported discographic criteria and standards (Table 3). There are 3 criteria reported by Walsh et al, Carragee et al and Derby et al for lumbar discography (15,17,28). The Carragee criteria differ from Walsh with a pressure limit of 100 psi a.o. (pounds per square inch above opening) versus Walsh's 400 – 500 kiloPascals or 58 to 72 psi a.o. The Derby criteria uses a pain response $\geq 6/10$, pressure limit of ≤ 50 psi a.o., grade 3 anular tear, and, a control disc with pain $< 6/10$. There are 2 published low pressure criterias: < 22 psi a.o. (21) and

Table 1. *Diagnostic interventions evaluation form per Agency for Healthcare Research Quality (AHRQ).*

Criterion	Weighted Score
1. Study Population	30
• <i>Subjects similar to populations in which the test would be used and with a similar spectrum of disease</i>	
2. Adequate Description of Test	15
• <i>Details of test and its administration sufficient to allow for replication of study.</i>	
3. Appropriate Reference Standard	20
• <i>Appropriate reference standard ("gold standard") used for comparison</i>	10
• <i>Reference standard reproducible</i>	10
4. Blinded Comparison of Test	20
• <i>Evaluation of test without knowledge of disease status, if possible</i>	10
• <i>Independent, blind interpretation of test and reference</i>	10
5. Avoidance of Verification Bias	15
• <i>Decision to perform reference standard not dependent on results of test under study</i>	
TOTAL SCORE	100

Adapted and modified from West S et al. Systems to Rate the Strength of Scientific Evidence, Evidence Report, Technology Assessment No. 47. AHRQ Publication No. 02-E016 (47).

Table 2. *Quality of evidence developed by USPSTF* .*

Level	STRENGTH OF EVIDENCE GRADING SYSTEM
I	Evidence from at least one properly controlled randomized trial
II-1	Well-designed controlled trials without randomization
II-2	Well-designed cohort or case-control analytic studies, preferably from more than one center or research group
II-3	Multiple time series with or without the intervention (also includes dramatic results in uncontrolled experiments)
III	Opinions of respected authorities, descriptive studies, and case reports, reports of expert committees.

*Adapted from the Agency for Healthcare Research and Quality U.S. Preventive Services Task Force (USPSTF) (27).

Table 3. *Discographic criteria and standards for a positive discogram.**

	Pain response NRS	Pressure (psi a.o.)	Pain behaviors	Grade 3 anular tear	Control disc NRS
Walsh/Carragee (15,17)	≥6/10	≤100	≥ 2/5	-	-
Derby (28)	≥6/10	≤50	-	Y	< 6/10
ISIS/IASP(a) (25)	≥7/10	<50	-	Y	-
ISIS/IASP (b) (25)	≥7/10	<50	-	Y	< 6/10
ISIS/IASP(c) (25)	≥7/10	<50	-	Y	0/10
Low pressure < 22 psi (Carragee) (21)	≥6/10	<22	≥ 2/5	-	-
Low pressure ≤ 15 psi (Derby) (28)	≥6/10	≤15	-	Y	<6/10

* Criteria: Walsh et al (15)/Carragee et al (17)(pain ≥ 6/10; pressure ≤ 100 psi a.o.; concordant pain; pain behaviors ≥ 2/5). Derby et al (28) (pain ≥ 6/10; pressure ≤ 50 psi a.o.; concordant pain; grade 3 anular tear; control disc with pain < 6/10). ISIS/IASP (pain ≥ 7/10; pressure < 50 psi a.o.; concordant pain; grade 3 anular tear; (a) = no control disc; (b) = control disc pain response < 6/10; (c) = painless control disc, 0/10). Low pressure < 22 psi a.o (Carragee et al [21]); pain ≥ 6/10; ≥ 2/5 pain behaviors. Low pressure ≤ 15 psi a.o. (Derby et al [28]); pain ≥ 6/10; grade 3 anular tear; control disc with pain < 6/10).

ISIS=International Spine Intervention Society; IASP=International Study for the Association of Pain; psi=pounds per square inch; a.o.=above

≤ 15 psi a.o. (28). There is one published standard for lumbar discography, per ISIS/IASP: concordant pain ≥ 7/10, pressure < 50 psi a.o., grade 3 annular tear, volume limit 3.5 mL, and a painless control disc (25). A control disc is considered a critical element for defining a positive discogram, because it serves as an internal patient control disc (particularly if it has a grade 3 annular tear) and because it serves as a possible indicator of central sensitization. To reveal the effect that including a control disc has a positive discogram, we included tiers of analysis: (a) no control disc; (b) control disc < 6/10; (c) painless control disc 0/10. All tables, except somatization disorder patients, include only discogrammed discs with grades 2 and 3 annular tears. The false-positive rate per patient and per disc are calculated with confidence intervals according to the various criteria and standards for comparison. Statistical analysis was performed with SPSS 12.0 software (SPSS, Chicago, IL).

A meta-analysis solely of the false-positive data was performed to determine the performance characteristics of discography in terms of specificity. There are 4 basic types of Cochrane meta-analyses: intervention reviews, diagnostic test accuracy reviews, methodology reviews, and overviews of reviews. We used the diagnostic test accuracy review sub-type of meta-analysis. Specificity is defined as the ability to correctly identify patients who are known not to have the disease. Specificity is the percentage of true-negative results. The remaining percentage represents the false-positive results. Only studies with AHRQ quality scores of ≥ 45/100 were included. As no reference standard exists for discography, only a meta-analysis of the specificity can be calculated; sensitivity cannot be calculated from available data in asymptomatic subjects. A forest plot can be drawn for specificity, but not sensitivity. As the data are limited to false-positives, odds ratio and effect size cannot be calculated; therefore, the Galbraith plot for heterogeneity, the funnel plot for publication bias, and the risk difference are not calculated. Data was weighted equally using a random effects model. Meta-analysis was performed with Review Manager Version 5.0 (RevMan, Copenhagen, Denmark).

RESULTS

From January 1, 1968, to March 30, 2008, the database yielded 11 papers. Studies were scored based on the AHRQ criteria based on 5 elements: study population, adequate description of test, appropri-

ate reference standard, blinding comparison of test, and avoidance of verification bias (Table 1). Salient features of the 11 articles are summarized in Table 4. Eight papers pertained to the false-positive rate in asymptomatic subjects (13-18,20,21); 2 papers to asymptomatic discs in subjects with chronic low back pain (29,30); and, one paper to subjects with mild persistent or clinically insignificant low back pain (19). Types of experimental studies included: one case control study; one case series; 2 prospective case series; 3 prospective, single-blinded case series; and 4 prospective, single-blinded case control studies. The demographics of the study populations are summarized (Table 4). Holt (14) and Walsh (15) studied populations with lower average ages than other studies, ages 26 years and 23 years, respectively versus populations ages 40 – 45 years amongst later studies. The earliest 4 studies were male only, later studies were performed with more balanced gender representation, except for the somatization disorder study which had more women.

Early Studies

The earliest 2 studies (13,14), received low AHRQ scores (Table 4). Discography criteria, technique, and quality assessment of research have changed markedly since the 1960s. Massie and Stevens (13) submitted a brief discussion of discography in 52 asymptomatic subjects with 570 symptomatic subjects. Their paper contained no quantitative data on pain and pressure; hence, no data could be abstracted for data synthesis or meta-analysis. All volunteers were males, age range 20 – 52 years of age. No pain scale was used, however, the authors stated that injection of Hypaque only “occasionally produced symptoms” in the asymptomatic group. Whereas, in the symptomatic group, they stated, “usually only one interspace was the source of symptoms; this space could be easily pointed out by the patients” (13). Among asymptomatic subjects, 60% (31 patients) had normal appearing discograms for the lower 3 lumbar segments versus 10% (57 patients) among symptomatic subjects. Morphologically abnormal discs were more common in patients as opposed to asymptomatic subjects; and the authors reported that discography clearly distinguished between negative asymptomatic discs versus positive symptomatic patient discs.

Holt (14) performed discograms on 30 volunteer inmates via the midline transthecal technique. The average age of the subjects was 26 years (range 21

Table 4. Summary of experimental studies of discography on asymptomatic patients and asymptomatic patient discs with evidence rating per AHRQ criteria (26,27).

Study	AHRQ score	Type of study	Patient group	Patients/ number of discs	Pain scale y/n	FP rate/ patient %	FP rate/ disc %
Massie & Stevens (13) (1967)	30	CC	Asymptomatic volunteers	52/156	N	NR	NR
Holt (14) (1968)	20	CS	Volunteer inmates	30/70 (20 failed injections)	N	NR	37%
Walsh et al (15) (1990)	70	P, CC, SB	Asymptomatic volunteers	10/30	Y	0%	0%
Carragee et al (16) (1999)	70	P, CS, SB	Iliac crest pain	8/24	Y	50%	29%
Carragee et al (17) (2000)	55	P, CS, SB	Asymptomatic volunteers	26/78 Pain free (10) Chronic pain (10) Somatization disorder (6)	Y	0% pain-free; 40% chronic pain; 75% SD	0% pain-free; 58% chronic pain; 50% SD
Carragee et al (18) (2000)	45	P, CC, SB	Post-discectomy	20/61	Y	40%	30.3%
Carragee et al (19) (2002)	40	P, CC, SB	Mild persistent back pain	25/33	Y	36%	37.5%
Derby et al (20) (2005)	80	P, CS, SB	Asymptomatic volunteers	13/43	Y	0%	0%
Carragee et al (21) (2006)	45	P, CS, SB	Asymptomatic patients (< 22 psi a.o.)	69/32	Y	25%	37.5%
Derby et al (20) (2005)	80	P, CC, SB	Asymptomatic discs in patients with LBP vs. controls	16 control patients/55 discs & 90 CLBP patients/282 discs	Y	NR	NR
Shin et al (30) (2006)	80	P, CS	Asymptomatic discs in patients with LBP	21/51 CLBP patients	Y	NR	NR

AHRQ = Agency for Healthcare Research and Quality; P = Prospective; CC = case control; C = Control; CS = Case series; FP = false-positive; SB = single blind; E = extrapedicular; NR = not reported; psi = pounds per square inch; a.o. = above opening; CLBP = chronic low back pain.

– 41 years). A volume of one to 2 cubic centimeters of 50% sodium diatrizoate (Hypaque) was injected into the disc. He reported a false-positive rate of 37%. Seventy-two accurate discograms were reported: “little discomfort” resulted from injections into 45 discs with a normal disc pattern; “severe pain” was reported in 11 discs (15%) when the dye pattern irregularly extended to the anulus; and, in 16 discs (22%), with a dye leak, “back pain or leg aching or both always resulted. It was often quite severe and took several minutes to abate”(14). Holt (14) opined that the pain from discography was caused by contact of the noxious Hypaque with sensory nerves in the outer anulus or with structures outside the disc.

Holt: Data Analysis

Reviewing Holt’s (14) tabular data directly, a total of 70/90 injections were successful from the L2-3 to L5-S1 levels. Twenty-two percent of disc injections were unsuccessful. No mention is made of sedation. Pain response, pressure criteria, concordance or a control disc were not utilized to define a positive discogram. Based on this limited data, if a positive discogram is interpreted as a degenerated discogram pattern with any pain response, the false-positive rate is 26% (18/70). Holt’s (14) data is not combined with more recent studies, as there is a concern that the dye used in the study was irritating to innervated structures and may have instead been a significant pain generator as op-

posed to the usual pressure stimulus of water-soluble, non-ionic contrast. In his paper, Holt (14) opined that the pain response during discography was due to the contrast contacting "any tissue having sensory innervation." Sixteen discs were reported as ruptured, indicating that pain may have been elicited by contrast leaking outside the disc and stimulating innervated structures such as the posterior longitudinal ligament, dural tissue, outer disc margin, ventral rami, the sino-vertebral nerve, or the spinal nerve. Theoretically, a noxious substance could also cause pain in normal nerve fibers in the outer 1/3 of the annulus fibrosus. If the 16 ruptured discs are removed from the analysis, the false-positive rate is reduced to 3.7% (2/54).

Asymptomatic Subjects without Confounding Factors

Three studies evaluated subjects asymptomatic of low back pain without confounding factors. In 1990, Walsh et al (15) compared 10 asymptomatic male volunteers to 7 patients with low back pain. The asymptomatic volunteers were male, average age 22 years old, with a prevalence of 17% (5/35) abnormal discograms. Walsh et al used a 0 – 5 pain scale. Patients

received 5 – 10 mg of versed. Water soluble, minimally irritating, non-ionic contrast was utilized (Iopamidol). Contrast "was slowly injected through a 3 mL plastic Luer-lock syringe either until maximal pressure (400 to 500 kilopascals = 58-72 psi a.o.) had been applied with the thumb or until severe pain was elicited" (15). A positive discogram was defined as pain \geq 3/5, \geq 2 pain behaviors, concordant pain, and abnormal disc morphology. Walsh et al (15) described concordant pain as a "pattern of pain...typical of (concordant with) the participant's usual pain." Until the Walsh et al (15) study, a standardized assessment of pain response was not part of the criteria for a positive discogram. In efforts to refine the criteria for provocation discography, Walsh et al (15) stated that "replication, during injection of the patient's typical pattern of pain should be considered an external corroborative test for a positive interpretation of a discogram." They reported a false-positive rate of 0% per patient and 0% per disc (Tables 6 and 7).

Carragee et al (17) studied a group of pain-free (no low back pain, no chronic pain) patients with an average age of 45 years old (Table 8). These 10 subjects were recruited from a registry of patients with

Table 5. Demographic characteristics of experimental study groups*

	Massie & Stevens (13) (1967)	Holt (14) (1968)	Walsh et al (15) (1990)	Carragee et al (16) Iliac crest (1999)	Carragee et al (17) Pain-free (cs-good) (2000)	Carragee et al (18) Chronic pain (cs-failed) (2000)	Carragee et al (22) Som (2000)	Carragee et al (18) Post disc (2000)	Carragee et al (19) Backache (2002)	Derby et al (20) Asymp volunteers (2005)	Derby et al (29) Asymp control discs vs. asymp discs in CLBP patients (2005)	Shin et al (30) Asymp discs in CLBP patients (2006)
Number of subjects	52	30	10	8	10	10	6	20	25	13	16 vs. 90	21
Age yrs (Range)	36 (20-52)	26 (21-41)	22.6 (18-32)	39.88 \pm 7.04	44.50 \pm 11.45	42.40 \pm 6.58	41.33 \pm 8.43	35.4 \pm 7.7	42	46.5 (33-61)	47 (32-61) vs. 44.7 (23-81)	52 (23-81)
Sex (M/F)	52/0	69/0	10/0	8/0	8/2	8/2	2/4	18/2	20/5	9/4	11/5 59/31	12/9

*cs-good = cervical spine, good outcome (no pain, no low back pain); cs-failed = cervical spine poorest outcome (chronic pain); Som = somatization disorder; asymp = asymptomatic; CLBP = chronic low back pain; vs = versus; pts=patients; no. = number; yrs = years; M = male; F = female

Table 6. Pooled table of volunteers asymptomatic of low back pain.* †

		Pain response NRS 1-10											
		0	1	2	3	4	5	6	7	8	9	10	
Pressure psi a.o.	100	○ † ○ † ○ † ○ †	○										
	90	○ †											
	80	○	○	■		■	○ †						
	70	○○○		○									
	60	○○○○	○			○	○						
	50	○○○■ ▲▲ ▲	○	■▲	○ † ○	○▲	○						
	40	■				○							
	30	○■■	○			○							
	20		○			■						□	
	10			■									
	0												

*○: 33 discs per Derby et al (20); ▲: 5 discs per Walsh et al (15) study (pressure range 58-72 psi a.o.); ■: 9 discs reported as negative per Carragee et al (17) (no pain, no low back pain group); □: case reported as positive per Carragee et al (17). Light and dark gray: Derby et al (28) criteria; dark gray: ISIS/IASP standard (25).

†Grade 2 annular tear (all other patients with grade 3 annular tears)

ISIS = International Spine Intervention Society; IASP = International Association for the Study of Pain; NRS = Numerical rating scale

excellent cervical spine surgery outcomes. Carragee et al (17) used the Walsh et al (15) criteria, but raised the pressure limit to ≤ 100 psi a.o. They reported 10% (1/10) of patients (95% CI 0 – 33%) and 10% (1/10) of discs (95% CI 0-33%) to be positive. Applying the Derby et al (20) criteria or ISIS/IASP standard, the false-positive rate is also 10% per patient and per disc (Table 7).

Using pressure and speed-controlled manometry, as opposed to manual pressurization, Derby et al (20) studied 13 asymptomatic volunteers (Table 6) (20). Six of 13 asymptomatic volunteers did have occasional low back pain (< 3 episodes/year). This study included 9 physicians and 4 lay persons, average age 47 years. Dynamic and static pressures were recorded. The dis-

cography injection rate was controlled to 0.05 mL per second (Intellisystem, Merit Medical Systems, South Jordan, UT). A negative discogram required no pain response ≥ 6/10 at pressures ≤ 50 psi a.o. A positive discogram was defined as pain ≥ 6/10, pressure ≤ 50 psi a.o. with a grade 3 annular tear. The total volume of contrast injected into the disc was limited to 3.5 mL. Derby et al (20) reported a false-positive rate of 0% per patient and per disc (Table 7).

Asymptomatic Subjects without Confounding Factors: Data Analysis

Discography data can be pooled from these 3 patient populations asymptomatic of low back pain (Table 6). Demographically, the Walsh et al (15) popu-

Systematic Review of Lumbar Provocation Discography in Asymptomatic Subjects

Table 7. Summary of false-positive rates (%) per patient and per disc for experimental studies in subjects asymptomatic of low back pain.* †

STUDY	Walsh et al (15)/ Carragee et al (17)		Derby et al (28)		ISIS/IASP (25)						Low pressure < 22 psi a.o (Carragee)		Low pressure ≤ 15 psi a.o. (Derby)		
	%FP /pt	%FP /disc	%FP /pt	%FP /disc	a		b		c		%FP /pt	%FP /disc	%FP /pt	%FP /disc	
					%FP /pt	%FP /disc	%FP /pt	%FP /disc	%FP /pt	%FP /disc					
Walsh et al (15): Asymptomatic volunteers (95% CI)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)
Carragee et al (16) Iliac crest (95% CI)	50 (5 - 95%)	28.6 (2 - 56%)	37.5 (0 - 81%)	21.4 (0 - 46%)	12.5 (0 - 42%)	7.1 (0 - 23%)	12.5 (0 - 42%)	7.1 (0 - 23%)	12.5 (0 - 42%)	7.1 (0 - 23%)	25 (0 - 64%)	14.3 (0 - 35%)	12.5 (0 - 42%)	7.1 (0 - 23%)	
Carragee et al (17): pain-free (cs-good) (95% CI)	10 (0 - 33%)	10 (0 - 33%)	10 (0 - 33%)	10 (0 - 33%)	10 (0 - 33%)	10 (0 - 33%)	10 (0 - 33%)	10 (0 - 33%)	10 (0 - 33%)	10 (0 - 33%)	0 (-)	0 (-)	0 (-)	0 (-)	
Carragee et al (17): chronic pain (cs-failed) (95% CI)	40 (3 - 77%)	58.3 (26 - 91%)	30 (0 - 65%)	33.3 (2 - 65%)	20 (0 - 50%)	16.7 (0 - 41%)	10 (0 - 33%)	8.3 (0 - 27%)	0 (-)	0 (-)	30 (0 - 65%)	25 (0 - 54%)	10 (0 - 33%)	8.3 (0 - 27%)	
Carragee et al (22): Somatization disorder (95% CI)	75 (0 - 100%)	44.4 (4 - 85%)	50 (0 - 100%)	22.2 (0 - 56%)	50 (0 - 100%)	22.2 (0 - 56%)	50 (0 - 100%)	22.2 (0 - 56%)	50 (0 - 100%)	22.2 (0 - 56%)	50 (0 - 100%)	22.2 (0 - 56%)	25 (0 - 100%)	11.1 (0 - 37%)	
Derby et al (20): Asymptomatic volunteers (95% CI)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	
Carragee et al (19) mild backache (95% CI)	36 (16 - 56%)	37.5 (20 - 55%)	36 (16 - 56%)	31.3 (14 - 48%)	20 (3 - 37%)	15.6 (2 - 29%)	20 (3 - 37%)	15.6 (2 - 29%)	16 (1 - 31%)	12.5 (0.4 - 25%)	28 (9 - 47%)	21.9 (7 - 37%)	28 (9 - 47%)	21.9 (7 - 37%)	
Carragee et al (18): Post-discectomy (95% CI)	35 (12 - 58%)	24.2 (9 - 40%)	35 (12 - 58%)	24.2 (9 - 40%)	25 (4 - 46%)	15.2 (2 - 28%)	25 (4 - 46%)	15.2 (2 - 28%)	15 (0 - 32%)	9.1 (0 - 19%)	25 (4 - 46%)	18.2 (4 - 32%)	25 (4 - 46%)	15.2 (2 - 28%)	

*ISIS = International Spine Intervention Society; IASP = International Association for the Study of Pain; a = no control disc; b = control disc < 6/10; c = painless control disc; FP = false positive; pt = patient; cs-good: cervical spine surgery, good outcome; cs-failed: cervical spine surgery, poorest outcome; CI : Confidence Intervals

† Holt (14) and Massie & Stevens (13) studies are not included as pain and pressure were not reported in the published study.

lution had subjects with a significantly lower age of 23 years. However, the negative effects of combining a clinically heterogeneous population were minimal because only 5 abnormal discs were contributed to a total of 33 discs. The Carragee et al (17) and Derby et al (20) populations had similar average ages of 45 years and 47 years, respectively. Only the asymptomatic groups were included in this combined analysis. The 2 patient populations with chronic pain (failed cervical spine surgery; iliac crest pain) are analyzed separately. The somatization disorder patients (n = 6) were not in-

cluded due to the nature and severity of their primary psychological disorder, small sample size, incomplete data from 2 of 6 patients, and lack of abnormal disc morphology (2/4 positive discs had a grade 1 anular tear). The study on asymptomatic post-discectomy patients (18) is analyzed separately due to the subject's history of significant low back pain requiring surgery and surgical distortion of disc anatomy. There were normal discs without prior surgery in the post-discectomy study, with 11 reported as negative and 2 as positive; however, partitioning these discs out of the

study as “normal” discs would not change the false-positive rate per patient and would in fact lower the per disc rate. The final data set includes 33 patients and 48 discs. The false-positive rate is 3.0% (1/33) per patient (95% CI 0 – 9%) and 2.1% (1/48) per disc (95% CI 0 – 6%) (Table 9), with all reported criteria and standards in agreement. With a larger sample size, confidence intervals shorten considerably as compared to Carragee et al’s (17) 95% confidence interval of 0-33% in their 10 pain-free subjects (Table 7).

Iliac Crest Pain Subjects

Carragee et al (16) performed lumbar discography in patients with non-spinal pain in the low back and gluteal regions, post-iliac crest bone harvest, to investigate concordance. They reported that 50% (4/8) of patients or 28.6% (4/14) of discs had pain exactly like or similar to their residual iliac crest donor site pain during the provocative discography of one or more of their intervertebral discs (16) (Tables 7 and 10). Subjects included 8 patients who had residual pain 2 – 4 months after undergoing iliac crest bone graft harvesting for non-spinal procedures. The pain was located in the low back and gluteal regions which are common

discogenic pain referral areas for low back pain. One of the positive subjects (case number 3) had more extensive procedures with 2 iliac crest bone harvesting procedures on the same side. This subject reported disc injections which were moderately painful (4/10) and “very bad” (8/10). Carragee et al (16) used manual pressurization with a pop-off valve set at 100 psi a.o. Pressure measurements were made during injection at each 0.5 mL of the injection (Hewlett Packard, Palo Alto, CA). They did not report whether discography was performed on the asymptomatic side. Subjects reported pain responses as none, dissimilar, similar, or exact. Concordancy was defined as exact or similar pain versus discordant pain or no pain. Using the ISIS/IASP standard (Table 3), the false-positive rate drops to 12.5% (1/8) as calculated per patient and 7.1% (1/14) as calculated per disc (Table 7).

Chronic Cervical Pain Subjects

Carragee et al (18) performed discography in 10 patients with chronic neck pain who had the “poorest” (direct quote) outcomes after cervical spine surgery (i.e. cervical spine surgery failed, “cs-f”). Demographic and psychometric characteristics of the 4 patients with

Table 8. Pain-free volunteer subjects by case number (good outcome after cervical spine surgery). * † ‡

		Pain response NRS 1 – 10										
		0	1	2	3	4	5	6	7	8	9	10
Pressure psi a.o.	100											
	90											
	80			14		13						
	70											
	60											
	50	13		17								
	40	12										
	30	15 20										
	20					13					11‡	
	10			15								
	0											

*Bold Italic: cases reported as positive per Carragee et al (17). Light and dark gray: Derby et al (28) criteria; dark gray: ISIS/IASP (25) standard
 †Grade 2 anular tear (all other patients with grade 3 anular tears)
 ‡Pressure: case no. 11: 25 psi a.o.
 ISIS = International Spine Intervention Society; IASP = International Association for the Study of Pain; NRS = Numerical rating scale; psi = pounds per square inch; a.o. = above opening; no. = number

Systematic Review of Lumbar Provocation Discography in Asymptomatic Subjects

Table 9. Summary of false-positive rates (%) per patient and per disc for combined analyses of experimental studies in subjects asymptomatic of low back pain. *†‡♦

COMBINED ANALYSIS GROUP	Walsh/Carragee		Derby		ISIS/IASP						Low pressure < 22 psi a.o. (Carragee)		Low pressure ≤ 15 psi a.o. (Derby)	
	%FP/pt	%FP/disc	%FP/pt	%FP/disc	a		b		c		%FP/pt	%FP/disc	%FP/pt	%FP/disc
					%FP/pt	%FP/disc	%FP/pt	%FP/disc	%FP/pt	%FP/disc				
Asymptomatic subjects† (95% CI)	3 (0~9%)	2.1 (0~6%)	3 (0~9%)	2.1 (0~6%)	3 (0~9%)	2.1 (0~6%)	3 (0~9%)	2.1 (0~6%)	3 (0~9%)	2.1 (0~6%)	0 (-)	0 (-)	0 (-)	0 (-)
Chronic pain (IC + cs-failed) ‡ (95% CI)	44.4 (19~70%)	42.3 (22~63%)	33.3 (9~58%)	26.9 (9~45%)	16.7 (0~36%)	11.5 (0~25%)	11.1 (0~27%)	7.7 (0~19%)	5.6 (0~17%)	3.85 (0~12%)	27.8 (5~51%)	19.2 (3~36%)	11.1 (0~27%)	7.7 (0~19%)
All subjects♦ (95% CI)	25.3 (15~35%)	20.7 (13~28%)	21.3 (12~31%)	15.5 (9~22%)	14.7 (7~23%)	9.5 (4~15%)	13.3 (6~21%)	8.6 (3~14%)	9.3 (3~16%)	6.0 (2~10%)	16.0 (8~25%)	11.2 (5~17%)	10.7 (4~18%)	6.9 (2~12%)

*ISIS = International Spine Intervention Society; IASP = International Association for the Study of Pain; a = no control disc; b = control disc < 6/10; c = painless control disc; FP = false positive; pt = patient; cs-good: cervical spine surgery, good outcome; cs-failed: cervical spine surgery, poorest outcome; IC = Iliac Crest; CI = Confidence Intervals

†Asymptomatic subjects: Walsh et al (15), Derby et al (20) and Carragee et al (18) (pain free/cervical spine good outcome) studies.

‡Chronic pain subjects: Carragee et al (22) studies (iliac crest pain and chronic pain/cervical spine failed outcome)

♦All subjects: excludes mild persistent backache patients and 2 somatization disorder patients, case numbers 25 and 26 with incomplete discogram data set.

Table 10. Iliac crest bone harvesting patients by case number (Carragee et al (16) study) * †‡

		Pain response NRS 1 – 10											
		0	1	2	3	4	5	6	7	8	9	10	
Pressure psi a.o.	100												
	90												
	80			3†				1†					
	70												
	60	8											
	50	1†		8†						3‡			
	40			4		6							
	30												
	20			4		1 2 3		7‡					
	10									8‡			
	0												

*Bold, italic: cases reported as positive per Carragee et al (16). Light and dark gray: Derby et al (28) criteria; dark gray: ISIS/IASP (25) standard

†Grade 2 annular tear (all other patients with grade 3 annular tears)

‡Pressures: case no. 3: 50 psi; case no. 7: 20 psi a.o.; case no. 8: 15 psi a.o.

ISIS = International Spine Intervention Society; IASP = International Association for the Study of Pain; NRS = Numerical rating scale; psi = pounds per square inch; a.o. = above opening; no.=number

Table 11. Demographic characteristics of the 4 patients with positive discograms in Carragee et al (18) asymptomatic volunteer patients with chronic pain.

Case# Operation	Age	Sex	Work	Disabled Y/N	AWC v. WC*	Meds†	MSPQ‡	ZUNG‡	DRAM §
3 C5/6f	47	F	Computer assembly	Y	AWC	3	29	33	DS (3)
4 C6/7f	42	M	Line assembly	Y	AWC	4	56	30	DS (3)
8 C5/6f	40	M	Warehouse worker	Y	AWC	2	16	50	DD (4)
10 C5/6d	55	M	Physician	N	n/a	3	5	12	N (1)

*AWC: active worker's compensation litigation in California; WC: worker's compensation claim with original injury

†Meds: 0 none; 1 occasional non-narcotic; 2 daily non-narcotic; 3 occasional narcotic; 4 daily narcotic

‡MSPQ = Modified Somatic Perception Questionnaire; ZUNG = Zung Self-rated Depression Scale

§DRAM = Distress and Risk Assessment Method; normal N = 1; at risk R = 2; distressed somatic DS = 3; distressed depressive DD = 4

f = fusion; d = discectomy; n/a = n of applicable

positive responses are shown (Table 11). Three of the 4 patients (75%) shared the following characteristics: status post failed cervical fusions, disabled, with active worker's compensation claims in litigation, and with psychometric testing indicating they were either distressed somatics or distressed depressives (Distress and Risk Assessment Method scores 3 and 4, respectively) (31). The fourth patient was a physician status post a failed cervical discectomy with normal psychometrics. Discography results are shown in Table 12. Among chronic pain patients, Carragee et al (18) reported a false-positive rate of 40% (4/10) per patient (95% CI 13 – 100%) and 58% (7/12) per disc (95% CI 26 – 91%) (Table 7).

Chronic Non-Spinal Pain Subjects: Data Analysis

The data was tabulated and analyzed according to various reported criteria (Table 7). Per ISIS/IASP (25) criteria, the false-positive rate is 0% (0/10) per patient and per disc. A false-positive rate of 0% was determined as follows. Cases 3 and 8 were not included because they did not have a painless control disc and the data set was incomplete. Of note, regarding cases 3 and 8. Carragee et al (17) stated that they attempted to find control discs cephalad to the abnormal disc injections, however, the video malfunctioned and therefore pain behaviors could not be evaluated and therefore this data was not included in the tabular results. Partial data on the L2-L3 disc was included in the narrative of the results section. Both discs showed grade 2 or greater anular disruption and pain response was 6/10 for case 3 and 2/10 for case 8. However, they did not report corresponding pressures or comment

on whether the pressure gauge also malfunctioned. Based on the ISIS/IASP (25) standard, case numbers 10 and 23 were also excluded per manometric criteria requiring pressures < 50 psi a.o. Case number 4 did not meet the pain standard.

Data from 2 chronic pain populations (n = 18 patients) can be combined from Carragee et al's (16,17) study of iliac crest bone graft harvest patients and failed cervical spinal surgery (Table 13). The Derby et al study (32) of negative discs in patients with chronic low back pain could not be included because individual pressure and pain responses were not reported. Per the Walsh/Carragee et al (15,17) criteria, the false-positive rate is 44% (8/18) per patient and 42% (11/26) per disc. With the Derby et al (28) criteria (with pressure ≤ 50 psi a.o.), the false-positive rates are lowered to 33% (6/18) per patient and 27% (7/26) per disc, due to the number of the cases clustered at exactly 50 psi a.o. If ISIS/IASP (25) criteria are applied, the false-positive rate is 5.6% (1/18) per patient (95% CI 0 – 17%) and 3.9% (1/26) per disc (95% CI 0 – 12%) with only a single disc from the iliac crest study considered positive (Table 9).

Chronic Low Back Pain Subjects: Asymptomatic Discs

Derby et al (32) studied the effect of chronic pain, specifically chronic low back pain, on discography results. They compared discographic responses between the control discs of asymptomatic patients versus the negative and positive discs of subjects with chronic low back pain. The discographic characteristics of negative discs in patients with chronic low back pain had not

Table 12. Chronic pain patients (cervical spine surgery-failed: cs-f), by case number. * †‡

		Pain response NRS 1-10											
Pressure psi a.o.		0	1	2	3	4	5	6	7	8	9	10	
	100		5†										
	90												
	80	10				3						8	
	70												
	60												
	50							3‡		10‡		8‡	
	40	4											
	30												
	20	9						4‡		8‡			
	10										3‡		
	0												

*Bold, italic: cases reported as positive by Carragee et al (18); Light + dark gray: Derby et al (28) criteria. Dark gray: ISIS/IASP (25) standard.
 †Grade 2 anular tear (all other patients with grade 3 anular tears)
 ‡Pressures: case no. 3: pain 9/10, 12 psi a.o. and pain 6/10, 50 psi a.o.; case no. 4: 20 psi a.o.; case no. 8: pain 8/10, 20 psi a.o and pain 10/10, 50 psi a.o.; case no. 10: 50 psi a.o.
 ISIS = International Spine Intervention Society; IASP = International Association for the Study of Pain; NRS = Numerical rating scale; psi = pounds per square inch; a.o. = above opening; no. = number

Table 13. Chronic pain patients (cervical spine surgery-failed and iliac crest pain).* †

		Pain response NRS 1 – 10											
Pressure psi a.o.		0	1	2	3	4	5	6	7	8	9	10	
	100		■†										
	90												
	80	■		○†		■		●†				□	
	70												
	60	○											
	50	○†		○†				□		□ ●		□	
	40	■		○		○							
	30												
	20	■		○		○○○		□ ●		□			
	10									●	□		
	0												

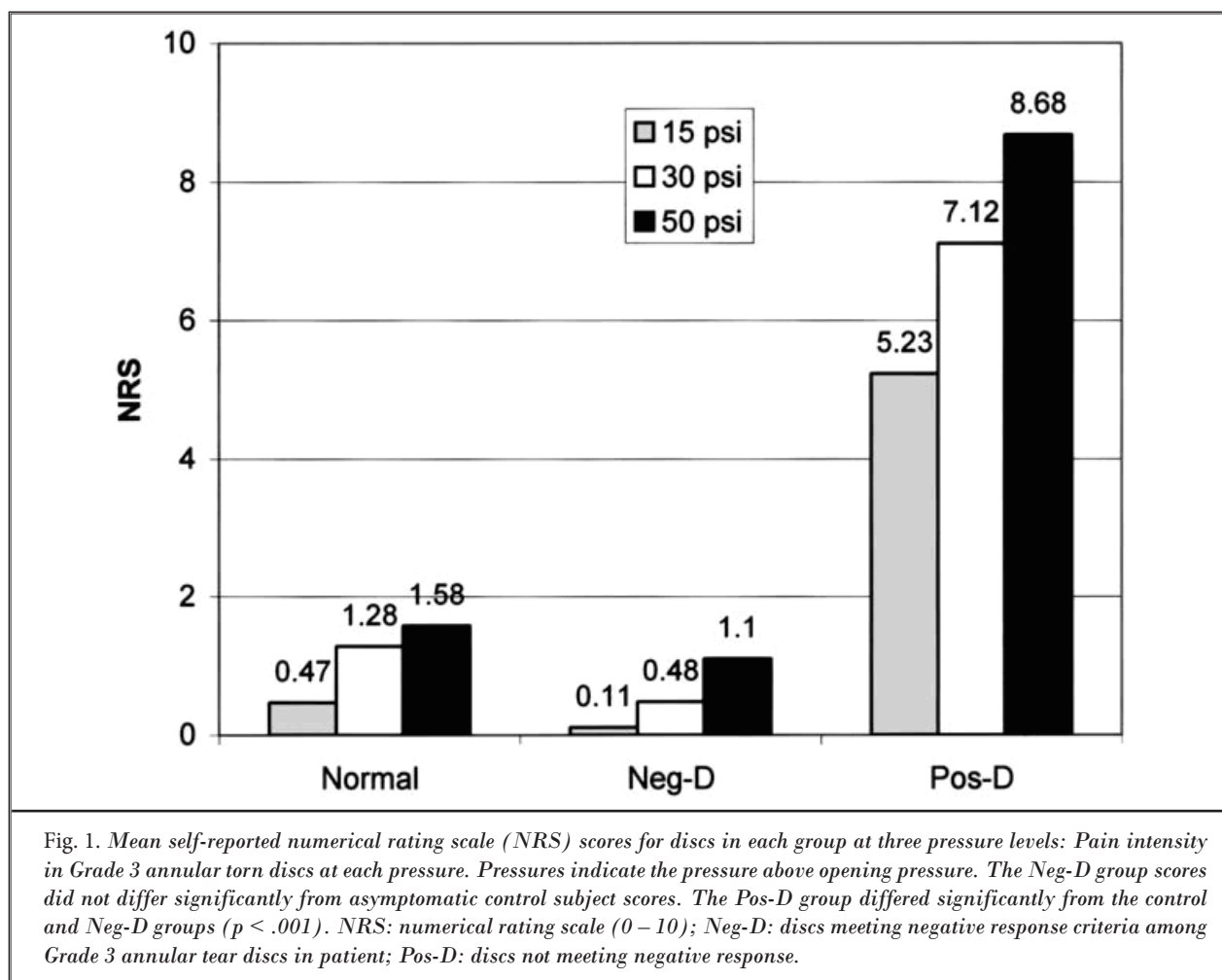
* ■: Caragee chronic pain negative discs (cervical spine surgery failed, poorest outcome); □ chronic pain, positive discs (17) ○: Carragee iliac crest, negative discs; ●: iliac crest pain, positive discs (16). Light and dark gray: Derby et al (28) criteria; dark gray: ISIS/IASP (25) standard
 †Grade 2 anular tear (all other patients with grade 3 anular tears)
 ISIS = International Spine Intervention Society; IASP = International Association for the Study of Pain; NRS = Numerical rating scale; psi = pounds per square inch; a.o. = above opening

been reported to date. The purpose of the study was to determine if discography could distinguish asymptomatic (negative) discs from symptomatic (positive) discs in patients with suspected chronic discogenic low back pain. Only discs with grade 3 annular tears (Dallas Discogram Scale) (25,33) were included in the analysis. The characteristics of the asymptomatic (negative) patient discs were compared to control discs of asymptomatic volunteers and to positive subject discs.

The patient sample included 55 discs from a control group (11 men, 5 women, 32 – 61 years of age, mean age: 47 years) of volunteers without current low back pain and 282 discs from a patient group of 90 low back pain patients (59 men, 31 women, 20-70 years of age, mean age: 44.7 years) (29). Discography was performed with speed and pressure-controlled manometry as described in an earlier study(20). Pressures and NRS pain responses were recorded at 15, 30, and 50

psi a.o. A negative discogram was defined as no pain reported as "familiar," no pain response $\geq 6/10$ at pressures < 50 psi a.o., grade 3 annular tear and ≤ 3.5 mL total injected volume. Patient discs were divided into 2 sub-groups based on discographic findings: negative discs (Neg-D) and positive discs (Pos-D).

Among 55 asymptomatic control group discs, 32 (58.2%) had grade 3 annular tears and all discs had negative discograms (29). Of 282 patient group discs, 199 (70.6%) demonstrated grade 3 annular tears. Of these discs, 104 (52.3%) met negative response criteria and were labeled as the negative discogram (Neg-D) group. The other 95 discs (47.7%) met positive response criteria and were categorized as the positive discogram (Pos-D) group (29). The control and negative disc groups showed similar pressures and volumes at which pain was initially evoked. The mean NRS score at 50 psi above opening pressure were $1.58 \pm$



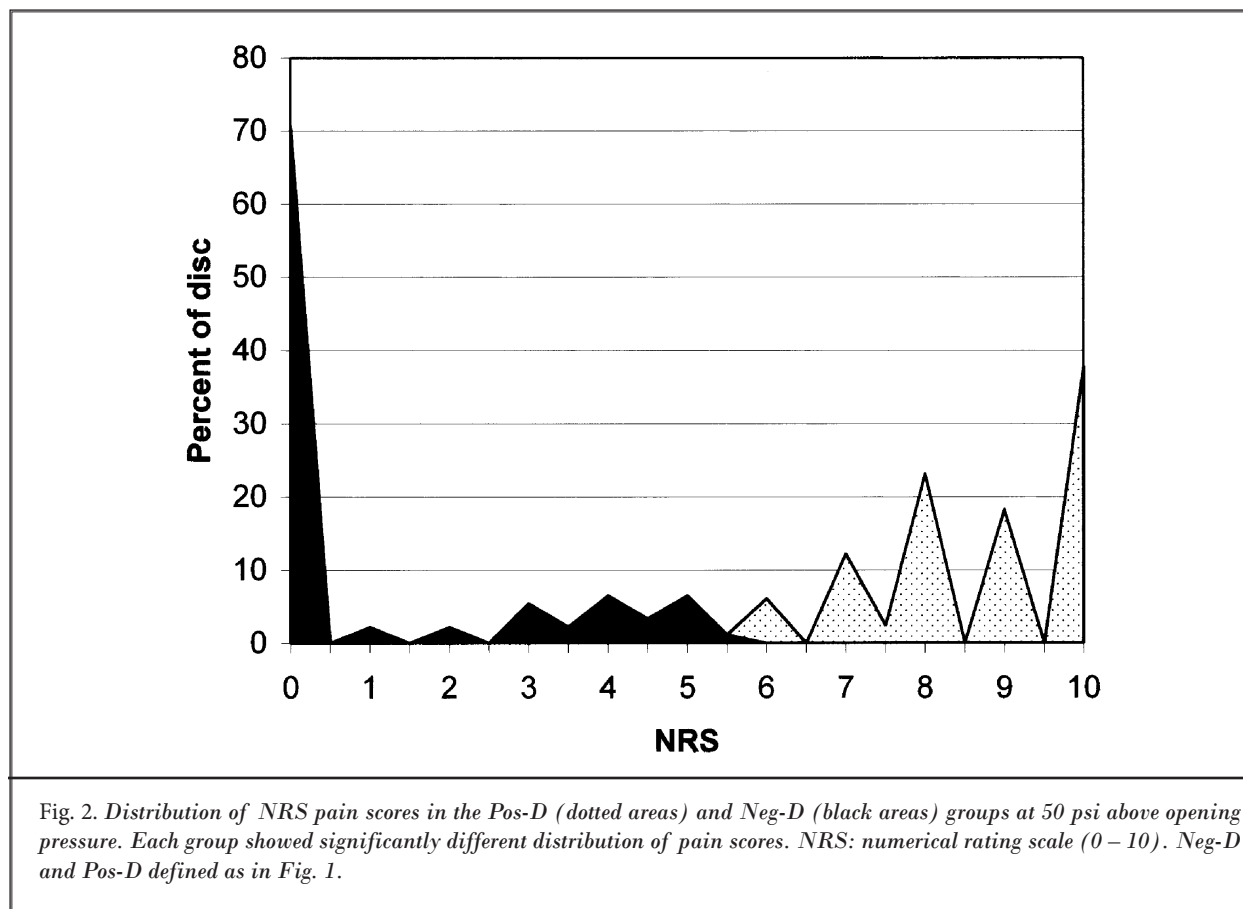
1.89 for the control, 1.10 ± 1.83 for the Neg-D, and 8.68 ± 1.27 for the Pos-D group (Fig. 1). At 15, 30, and 50 psi a.o., there were no statistically significant differences between the control group and negative discogram group. However, the positive discogram group NRS scores were statistically significantly higher, averaging, 8.68 ± 1.27 ($P < 0.001$) (29). Pain responses of the Pos-D and Neg-D groups differed significantly. In the Pos-D group, 78.9% (75/95) of discs demonstrated pain responses $\geq 8/10$ NRS, whereas, under the same pressure stimulation, 70.2% (73/104) of the Neg-D group did not cause any pain (0/10 NRS) ≤ 50 psi a.o. (Fig. 2).

Shin et al (30) performed a study similar to Derby et al (29) using a manometric syringe (Atrion QL 1015, Atrion Medical Products, AL, USA) in patients with chronic low back pain to compare the characteristics of negative and positive response discs. A total of 21 patients with 51 discs were evaluated (12 men,

9 women; mean age 52 years, range 23 – 81 years). Positive and negative responses were defined similar to the Derby study. Among the 51 discs, 82% (42/51) had grade 3 or greater anular tears. Sixty-nine percent (29/42 discs) had a negative response and 31% (13/42 discs) had a positive response. Under similar pressure stimulation of 50 psi, the intensity of pain responses were significantly different between the negative and positive discs, with 87% of negative discs having pain $< 6/10$ vs. 100% of positive discs with pain $> 6/10$.

Somatization Disorder Subjects

Carragee et al (22) studied discography in somatization disorder patients, asymptomatic of low back pain (Table 14). A total of 6 patients participated in the study. Four of 6 patients were able to complete all disc injections. Psychometric testing in these patients revealed that they were all distressed somatics or depressives. He reported a false-positive rate of



75% (3/4) per patient (95% CI 0 – 100%) for subjects completing at least 3 injections (Table 7) (22). They also reported that 83% (5/6) patients who completed at least one disc injection had a positive discogram. Two of the 6 (33%) of the patients could not complete the discogram: case number 25 stopped after 2 injections and case number 26 stopped after one injection. Case numbers 21 and 23 had grade 1 annular tears; case number 23 also had pain with high pressure (= 100 psi a.o.) provocation. With the Derby et al (28) criteria or ISIS/IASP (25) standard applied, the false-positive rate can be reduced to 50% (2/4) per patient (95% CI 0 – 100%) or 22.2% (2/9) per disc (95% CI 0 – 56%) (Table 7).

Post-Discectomy Subjects

Carragee et al (18) performed discography in 20 asymptomatic post-discectomy patients (Table 15) (18). All false-positive responses, except for one disc, occurred in discs with prior surgery. All positive responses occurred at ≤ 25 psi a.o. Carragee et al reported a 40% (8/20) false-positive rate per patient. Closer inspection of the data suggests that there were actually 7 patients which were positive, reducing the

false-positive rate to 35% (7/20) per patient and 24% (8/33) per disc. If ISIS/IASP (25) standards are applied, the false-positive rate falls to 15% (3/20) per patient (95% CI 0 – 32%) and 9.1% (3/33) per disc (95%CI 0 – 19%) (Table 7).

Chronic Persistent Low Back Pain Subjects

Carragee et al (18) studied 25 subjects (25) with what he termed “mild persistent backache, “benign” backache and/or “minimal low back pain.” These patients were compared to 52 symptomatic patients with a presumed diagnosis of discogenic pain. Thirteen patients had a history of a good outcome after cervical spine surgery, case numbers 1 – 13; 12 patients had the poorest outcome after cervical spine surgery, case numbers 14 – 25. Provocative discography was performed in these subjects and compared to the results of subjects with chronic low back pain illness. To be included in the study, the patients had to answer yes to the following question: “I have low back pain every day” OR “I have back pain almost everyday.” Visual Analog Scores (VAS) for subjects averaged 2.2 to 4.1. The patients did not restrict their activity or seek medical care for their low back pain. The patients also had to answer yes to

Table 14. Somatization disorder patients, by case number. *†‡±

		Pain response NRS 1 – 10											
		0	1	2	3	4	5	6	7	8	9	10	
Pressure psi a.o.	100							23s					
	90												
	80												
	70												
	60												
	50											25ss± 26s±	
	40												
	30												
	20	22S	22S						21s‡		24S‡		
	10	22S 22S 23S											21S‡
	0												

*Bold, italic: cases reported as positive by Carragee et al (22)(e.g. 75% false-positive rate per patient with cases no. 21, 23, and 24; 83% false-positive rate if case nos. 25 and 26 included). Light + dark gray: Derby et al (28) criteria; dark gray: ISIS/IASP (25) standard.

†Anular tears: s = grade 1; ss = grade 2; S = grade 3.

‡Pressures: case no. 21s: 25 psi a.o.; case no. 21S: 12 psi a.o.; case no. 24S: 20 psi a.o.

±Patients excluded from current analysis: case no. 25 completed 2/4 disc injections; case no. 26 completed 1/4 disc injections.

ISIS = International Spine Intervention Society; IASP = International Association for the Study of Pain; NRS = Numerical rating scale; psi = pounds per square inch; a.o. = above opening; no. = number.

Systematic Review of Lumbar Provocation Discography in Asymptomatic Subjects

Table 15. *Post-discectomy patients by case number*†‡§*

		Pain response NRS 1 – 10											
Pressure psi a.o.		0	1	2	3	4	5	6	7	8	9	10	
	100			2†§ 5†			18†§						
	90												
	80	5§		1†§ 14		3 8 13		7†§ 10†					
	70												
	60												
	50	13		17*						10‡§			
	40	4 12§											
	30	15 16§		20§		1							
	20	11 14§		9§		13§		8‡		11‡§			
	10			15§		3		19‡§		4‡§		6‡§	
	0									3‡§		8‡§	

*Bold, italic: cases reported as positive per Carragee et al (19). Light + dark gray: Derby et al (28) criteria. Dark gray: ISIS/IASP (25) standard. †Grade 2 anular tears (all other patients with grade 3 anular tears) ‡Pressures: case no. 3: 5 psi a.o.; case no. 4: 10 psi a.o.; case no. 6: 12 psi a.o.; case no. 8: pain 6/10, 20 psi a.o. and pain 10/10, 5 psi a.o.; case no. 10: 50 psi; case no. 11: 25 psi a.o.; case no. 19: 10 psi a.o. §Disc with prior surgery. ISIS = International Spine Intervention Society; IASP = International Association for the Study of Pain; NRS = Numerical rating scale; psi = pounds per square inch; a.o. = above opening; no. = number.

Table 16. *Volunteer subjects with mild persistent low back pain; case nos. 1 – 13: good cervical spine surgery outcomes; case nos. 14 – 25: poorest cervical spine surgery outcomes. *†‡*

		Pain response NRS 1 – 10											
Pressure psi a.o.		0	1	2	3	4	5	6	7	8	9	10	
	100												
	90												
	80	5† 15		10† 13 12†22†						11		23	
	70												
	60							4					
	50			7 21		10				15‡ 24‡			
	40			14									
	30	15†		8						24‡		17‡	
	20			20 25									
	10	3		8		16		3‡ 4‡ 13‡ 23‡		1‡ 14‡ 21‡		16†‡	
	0												

*Bold, italic: cases reported as positive per Carragee et al (19). Carragee et al excluded cases: case no. 3, one pain behavior; case no. 11 non-concordant pain; case no. 15: non-concordant pain; included case no. 23: 10/10 pain, pressure 80 psi a.o. concordant pain, 3 pain behaviors, grade 3 anular tear, control disc. Light + dark gray: Derby et al (28) criteria. Dark gray: ISIS/IASP (25) standard. †Grade 2 anular tears (all other patients with grade 3 anular tears) ‡Pressures: case no. 1: 10 psi a.o.; case no. 3: 15 psi a.o.; case no. 4: 12 psi a.o.; case no. 13: 15 psi a.o.; case no. 14: 15 psi a.o.; case no. 15: 50 psi a.o.; case no. 16: 10 psi a.o.; case 21: 10 psi a.o.; case 23: 15 psi a.o.; case no. 24: pain 8/10, 30 psi a.o.; case no. 24: pain 8/10, 50 psi a.o. ISIS = International Spine Intervention Society; IASP = International Association for the Study of Pain; NRS = Numerical rating scale; psi = pounds per square inch; a.o. = above opening; no. = number.

the following question: "I take no regular medications for my low back pain." According to the demographic table, 72% (18) of patients were on regular non-opiates or opiates for their other pain conditions. The results of the study are shown (Table 16). In their paper, Carragee et al (19) reported false-positives rates of 36% (9/25) per patient and 37.5% (12/32) per disc, respectively. Of their comparison group with chronic low back pain, 38/52 (73%) had positive discograms. If ISIS/IASP (25) criteria are applied, the false-positive rate falls to 16% (4/25) per patient and 12.5% (4/32) per disc (Table 7). This subject group was not combined in additional data analyses because subjects were symptomatic, with chronic low back pain.

Low Pressure Positives in Discography

Carragee et al (18) also addressed the issue of low-pressure positive discography in subjects asymptomatic of significant low back pain illness. Using a post-hoc, exploratory analysis, data was combined from 5 patient subgroups in prior studies (n = 69): no low back pain, no chronic pain (n = 10); no low back pain, chronic pain (n = 10); somatization disorder (n = 4); no low back pain, previous lumbar discectomy (n = 20); and minor "benign" backache (n = 25). They did not include the iliac crest study patients. A low pressure positive disc was defined as pain at less than 22 psi a.o. (21). They reported the following number of false-positive responses by group: pain-free 0/10, chronic pain 3/10, somatization disorder 2/4, post-discectomy 5/20, and, minor benign backache 7/25. In the experimental group they found at least one positive disc in 17/69 (25%) patients versus the clinical low back pain group 14/52 (27%), and reported no statistically significant difference between the asymptomatic low-pressure positive patients and symptomatic patients with chronic low back pain.

Low Pressure Positives: Data Analysis

Reviewing all the study populations individually and in combined analyses, a few comments can be made about using low pressure criteria for a positive discogram (Table 9). Setting the threshold at < 22 psi a.o. does not significantly improve the test if individual studies are analyzed, but if applied to all subjects, the false-positive rate is 16% per patient and 11.2% per disc. Setting the threshold at ≤ 15 psi a.o. obtains false-positive rates of approximately 10% for the iliac crest pain and chronic pain subjects, however, confi-

dence intervals are sizeable. Reviewing the combined analyses, a ≤ 15 psi a.o. threshold obtains a false-positive rate of 0% for the asymptomatic subjects; 11.1% per patient and 7.7% per disc, for chronic pain patients; and 10.7% per patient and 6.9% per disc, for all subjects combined. For all subjects combined (Table 9) the low pressure criteria of ≤ 15 psi a.o., performs slightly better than the < 22 psi threshold, with 10.7% per patient, 6.9% per disc and 16.0% per patient and 11.2% per disc, respectively.

All Subjects Asymptomatic of Low Back Pain

One can combine all the data into one table (Table 17) and arrive at various false-positive rates according to published criteria and standards that have been used to determine whether the provoked pain constitutes a false-positive response (Table 9). This combined analysis excluded 2 patient groups: symptomatic backache patients and 2 somatization disorder patients with an incomplete data set. The potential false-positive rate is presented per patient and per disc. In the per patient analysis, any one positive disc (according to criteria being used) places the patient in the false-positive category. In a per disc analysis, all the discogrammed discs (with \geq Grade 2 annular tears) in all patients are combined (Table 17). Using the ISIS/IASP (25) standard, the combined analysis of 75 patients and 116 discs obtains the most acceptable potential false-positive rate of 9.3% (95 CI 3 – 16%) per patient and 6.0% (95 CI 2 – 10%) per disc (Table 9). Based on review of the combined analysis, reasonable false-positive rates (< 15%) can also be obtained if one does not use a control disc or if the control disc is used, must have pain < 6/10.

Meta-Analysis: Diagnostic Test Accuracy Review

The false-positive data from all high quality studies of discs in subjects asymptomatic of low back pain was analyzed with the resulting forest plot of specificity (Fig. 3). The forest plot provides a visual representation of individual studies and all studies combined with their confidence intervals. The Holt (14) and Massie and Stevens(13) studies were excluded because they reported descriptive pain responses without pressures. The Carragee et al (19) study of low back pain patients was excluded because the subjects were symptomatic. Based on meta-analysis of the data, using the ISIS/IASP (25) standard, discography has a high specificity of 0.94 (95% CI 0.89 – 0.98) with a false-

Table 17. Pooled table of all studies of volunteers asymptomatic of low back pain. *±

		Pain response NRS 1 – 10										
		0	1	2	3	4	5	6	7	8	9	10
Pressure psi a.o.	100	○ ○ ○ ○	○ ■	■ ■		■		□				
	90	○										
	80	○ ■ ■	○	■ ■ ■ ■		■ ■ ■ ■ ■ ■	○	□ □ □				□
	70	○ ○ ○		○								
	60	○ ○ ○ ○ ■	○			○	○					
	50	○ ○ ○ ▲ ▲ ▲ ■ ■ ■	○	▲ ■ ■ ■	○ ○	○ ▲	○	□		□ □ □		□ □± □±
	40	■ ■ ■ ■		■		○ ■						
	30	○ ■ ■ ■ ■ ■	○	■		○ ■						
	20	■ ■ ■ ■	○ ■	■ ■		■ ■ ■ ■ ■ ■		□ □ □ □		□ □ □	□	
	10	■ ■ ■		■ ■		■		□		□ □	□	□ □
	0									□		□

*○ : Derby et al (28) study, ▲:Walsh et al (15) study, ■: Carragee et al (16-19,22) studies, discs reported as negative; □ Carragee et al (16-19,22), discs reported as positive. Light and dark gray: Derby et al (28) criteria; dark gray: ISIS/IASP (25) standard
±Symptomatic backache patient discs excluded; 2 patients with somatization disorder with incomplete dataset excluded.
ISIS = International Spine Intervention Society; IASP = International Association for the Study of Pain; NRS = Numerical rating scale; psi = pounds per square inch; a.o. = above opening

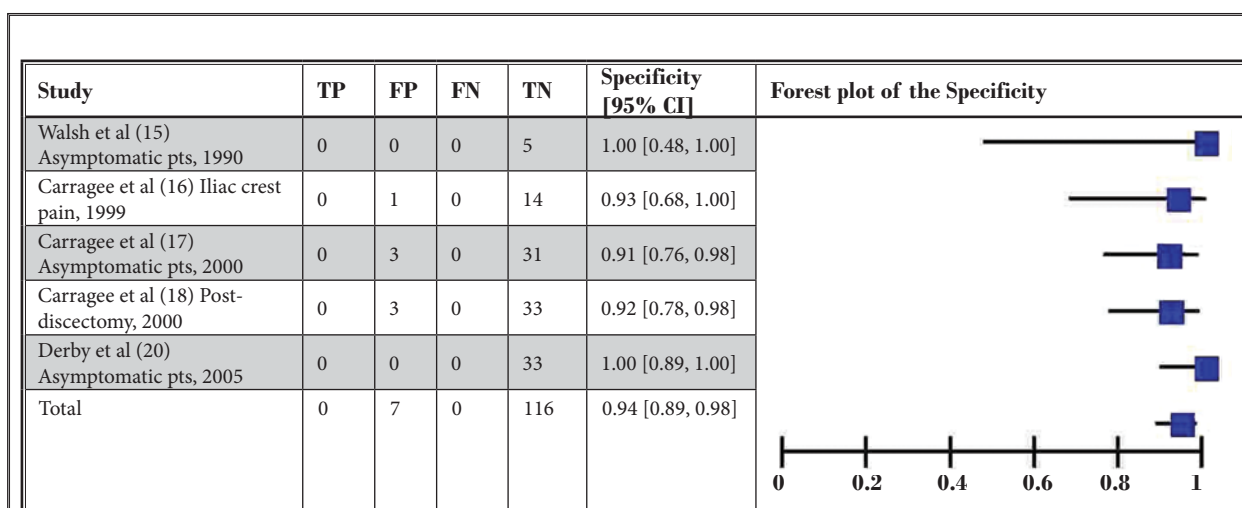


Fig. 3. Forest plot of specificity. All individual studies of asymptomatic subjects included in analysis. TP = true positives; FP = false-positives; FN = false-negatives; TN = true negatives; CI = confidence intervals; pts=patients.

positive rate of 6%.

Discussion

We analyzed the experimental data of all published lumbar discography studies investigating the potential false-positive rate among asymptomatic patients and discs. Except for 2 early studies, all studies reported pressures (either obtained from manual static readings, or using validated pressure-controlled manometry) during provocation of maximal pain of individual volunteer discs. Since stimulating any innervated structure with graduated intensities of stimulus will provoke pain in an asymptomatic individual, provocation criteria are required (20,25,28). The goal is to utilize criteria which provide the best compromise between limiting false-positive responses and false-negative responses. Since pressure is the primary stimulus for pain, in so far as we understand in our attempts to replicate discogenic pain, pressure limit criteria best determines the graduated order of criteria.

We have combined all data into one group (Tables 9 and 17) which obtains false-positive rates less than 10% (9.3% per patient and 6.0% per disc). The authors also analyzed the data according to the particular confounding factor of the individuals studied. The only like group of volunteers included by the 3 separate investigator groups were volunteers that were asymptomatic of any type of persistent pain. One investigator (17) studied 2 subject populations with chronic pain ($n = 18$) unrelated to their lumbar spine; one group with somatization disorder; and one post-discectomy group

Asymptomatic Subjects without Confounding Factors

In patients without chronic pain and asymptomatic of low back pain, discography has a low potential for false-positive results if one adheres to the standardized criteria for a positive response (25). In spite of the now obvious shortcomings of the study performed by prominent spine surgeon Earl Holt, MD (14), a 37% per disc false-positive for discography stood as the standard for over 2 decades, until publication of the 1990 Walsh et al study (15). Interestingly, if the Holt (14) data is re-examined based on any pain response with an abnormal discogram (minus the ruptured discs leaking irritating Hypaque), the false-positive rate is actually 3.7% per disc. Walsh et al (15) and Derby et al (20) both reported a false-positive rate of 0%. Car-

ragee et al (17) found a false-positive rate of 10% in pain-free post-cervical spine surgery patients asymptomatic of low back pain. If the data from these studies is combined, the false-positive rate is 3.0% (95% CI 0 – 9%) per patient and 2.1% (95% CI 0 – 6%) (Table 9). Based on all available data, the evidence shows discography to be a reliable test with a low false-positive rate in asymptomatic volunteers without confounding factors.

Iliac Crest Pain Study

The only other studied group of volunteers having chronic pain unrelated to the lumbar spine were 8 volunteers with normal psychometrics who underwent iliac crest bone grafting and reported persistent post-procedure pain (16). Carragee et al (16) reported a 50% false-positive rate per patient (28.5% per disc) (Table 7). These results challenged the reliability and utility of concordancy as a component of discography. However, when the ISIS/IASP (25) criteria are applied, the rate falls to 12.5% (95% CI 0 – 42%) per patient and 7.1% (95% CI 0 – 23%) per disc (Table 7). Although this was a small sample size with large confidence intervals, the potential false-positive rate is slightly elevated as compared to other subject groups. This study of subjects with persistent pain after bone graft harvesting raises an important caveat for discographers. It suggests that there may be cases where segmental innervation and nociceptive receptive fields overlap. The iliac crest and cutaneous superior gluteal region are innervated by the superior cluneal nerves, derived from the T12 and L1-L3 dorsal rami (34). The nerves cross over the iliac crest 7 cm lateral to the midline. Animal studies on disc sensory innervation show that cell bodies for the L4-5 disc are present at both the segmental levels as well as in the upper lumbar DRGs (primarily L2) traveling via the paravertebral sympathetic trunks (35-37).

We agree that concordancy is not sufficient to stand alone to determine a positive discogram; however, it is a necessary component of the test. To lessen false-positive responses, discographers need to be wary of other possible sources of pain in patients undergoing discography. The current recommendation is that other significant sources of spinal pain are evaluated prior to interpreting provocative discography (38,39). Patients should undergo appropriate diagnostic imaging to rule-out other sources of pain including neoplasm, cysts, infection, fracture, non-spinal sources,

etc. Patients should also undergo appropriate diagnostic blocks of zygapophyseal joints and the sacro-iliac joints. It is well known that pain may be reproduced at various locations along the neural axis. For this reason, many expert discographers place the needles from the asymptomatic or least symptomatic side to lessen any false-positive response due to pain from the needle itself or due to the needle stimulating other structures besides the disc (e.g. ventral ramus or DRG). In the Carragee et al (16 study, no mention is made of the side of needle insertion.) The rate of concordant pain may be theoretically even lower if needles are placed contralateral to the side of the pain. Discography should be performed on the asymptomatic or least symptomatic side to limit false-positive responses.

Chronic Pain Subjects: Is Chronic Pain a Significant Confounding Factor?

In patients with chronic cervical pain, Carragee et al (18) reported a false-positive rate of 40% (95% CI 13-100%) If the ISIS/IASP (25) standard is utilized the false-positive rate is 0% per patient and per disc. If the data is pooled from the 2 chronic non-spinal pain studies, new false-positive rates with markedly narrowed confidence intervals can be calculated. Per the ISIS/IASP criteria (25) the false-positive response is 5.6% (95% CI 0 – 7%) per patient and 3.85% (95% CI 0 – 12%) per disc (Table 9). The marked differences in false-positive rates between Carragee et al's (18) report and ISIS/IASP (25), however, rest on small distinctions (Table 11): one-point on the NRS scale, whether or not pressure was less than or equal to 50 psi a.o. and whether or not a control disc was obtainable and painless. Carragee et al (18) did attempt to find control discs. It may be that the false-positive rate in this patient population with the worst of the worst in terms of psychological and psychosocial risk factors, is indeed higher. However, rather than using the data obtained from Carragee et al 's (18) 10 chronic pain patients and arguing about the appropriate criteria for a false-positive response, one could argue that the pooled data from the larger asymptomatic study population should be used to determine the false-positive rate in a majority of patients.

The argument against using this population is that volunteers without chronic pain will react differently to applied stimulus compared to the typical patient undergoing lumbar discography. However appealing the argument, it is largely conjecture and in fact the evidence indicates that patients with chronic or chron-

ic intermittent low back pain respond the same to disc stimulation as an asymptomatic group without chronic pain. Derby et al (29) studied the effect of chronic low back pain on negative and positive patient discs versus asymptomatic controls. Comparing grade 3 discs of asymptomatic controls and negative patient discs, there was no statistical difference in the pain intensity at incremental distending pressures. That is, the stimulus (as measured by the distending dynamic pressure) required to provoke pain was statistically the same at various distending pressures. For example, when grade 3 discs at 50 psi a.o. were compared (Fig. 1), asymptomatic volunteers reported 1.6/10 pain and chronic low back pain patients reported 1.1/10 pain. In contrast, the difference between a negative and positive disc in patients was substantial, and patients could easily distinguish between the two. For example, symptomatic grade 3 patient discs exhibited a statistically significant and clinically meaningful pain response as compared to both the control and negative patient groups, reporting 8.7/10 pain at 50 psi a.o. ($P < 0.001$) (Fig. 1). One would expect patients with chronic pain to overreact to pain stimulus and one would expect that a significant and perhaps clear majority of these abnormal discs when stimulated would provoke a response of $> 6/10$ pain at pressures less than 50 psi above opening. In fact, 70% of the negative discs did not evoke any pain (0 NRS) at < 50 psi a.o. stimulation (29) (Fig. 2). Shin et al (30) also studied pain responses in negative versus positive discs in chronic low back pain patients and found that at 50 psi a.o. stimulating pressure, among grade 4 or greater discs, patients could distinguish from positive and negative discs by pain response. One cannot, therefore, arbitrarily accept the argument that a majority of patients with chronic low back pain presenting for discography react differently to volunteers without chronic pain and in fact the evidence to date indicates that these patients do not over-report pain.

Chronic Pain Subjects: Is Abnormal Psychological Testing a Confounding Factor?

Another argument against using asymptomatic volunteers without chronic pain as the comparison cohort for discography is based on the conjecture that most patients with chronic pain have abnormal psychological profiles that influence their reporting of pain during discograms. In Carragee et al's 10 patients with chronic pain 3 of 4 (75%) patients had abnormal DRAM (Distress and Risk Assessment Method) scores

(17). The DRAM is a commonly utilized psychometric test for somatization and depression. However, in a study of 81 patients, Derby et al (40) found a much lower percentage of patients presenting with abnormal psychometrics, 33% as opposed to Carragee et al's 75% or 3/4 patients. Derby et al (40) reported the following DRAM scores: 15% (12/81) normal; 52% (42/81) at risk, and 33% (27/81) abnormal (distressed somatic and distressed depressive). The positive rates of discography were not statistically significant by subgroup: normal 75% (9/12), at risk 59.5% (25/42), and 70.4% (19/27) positive discograms respectively, ($P > 0.05$) (40). In patients with chronic low back pain, there was no correlation between presenting psychometric DRAM score and discography results. These studies do not suggest that elevated psychometric scores will generate higher than normal false-positive results. The available evidence therefore indicates that in a majority of patients, the pooled data from the discs of asymptomatic volunteers without chronic pain are an appropriate comparison cohort.

Somatization Disorder Subjects

Carragee et al (17) reported a high false-positive rate of 75% per patient (95% CI 0 – 100%) with somatization disorder. A 4-patient sample size with large confidence intervals, limits the generalizability of these findings to all patients with somatization disorder. A randomized-controlled trial comparing 25 back pain patients with and without somatization disorder (41) found no significant difference in positive discogram responses between groups (as well between patients with and without depression and general anxiety disorder). The current recommendation for patients with somatization disorder is that discographers and surgeons should consider invasive testing and surgery only for the best surgical candidates, recognizing that patients with somatization disorder commonly complain of recurrent pain, conversion (pseudoneurologic) symptoms; and commonly self-medicate and are at risk for iatrogenic illness (42). Somatization disorder-patients are also hospitalized or undergo surgery 3 times as often as depressed patients (43).

Post-Discectomy Subjects

The current data (18), based on 20 volunteers undergoing discography in previously operated discs, shows a false-positive rate ranging from 35% per patient and 24% per disc by the Walsh et al (15)/Carragee et al (19 criteria) to 15% and 9.1% by the ISIS/ISAP

(25) criteria. In this sub-group, statistically, the false-positive rate is expected to be higher because the subjects have known pre-existing, albeit quiescent, discogenic disease. When the pre-test probability of disease (prevalence) is high, the positive predictive value (the likelihood that a patient with a positive discogram will have the disease) is also high. Given our limited knowledge of discography in post-discectomy patients and the possibility that provocation may open previously healed granulation tissue along surgical planes, discographers may consider the use of pressure and speed-controlled manometry to limit false-positive responses in this sub-group (44).

Chronic Persistent Low Back Pain Subjects

The inclusion of volunteers with chronic daily low back pain in a study (19) investigating the false-positive potential of discography is inappropriate. The authors assert that subjects did not restrict their activity or seek medical care for their low back pain therefore this was not significant low back pain. However, this level of pain does not represent "benign" low back pain; instead, it arguably represents chronic mild to moderate low back pain. The subject's VAS scores also represent chronic mild to moderate chronic low back pain. Seventy-two percent (18 patients) were on non-opiates or opiates for their other pain conditions, which may have also masked the severity of their low back pain. Carragee et al (19) reported 9/25 (36%) false-positives, however, these results are arguably true-positives. However, one could argue that these chronic low back pain volunteers are no different from patients undergoing discography who often have varying degrees and duration of pain "flare-ups." Arguably, during discography, many of these volunteers had pain caused by internally disrupted lumbar intervertebral discs. The finding that 36% had a positive discogram is consistent with the reported prevalence of low back pain in a patient commonly undergoing discography, ranging from 26% to 39% (45,46). The argument that these positive responses represent "false-positive responses" is not supportable. Discography was not developed and should not be used to determine the clinical significance of a patient's perceived suffering and disability related to chronic low back pain.

Low Pressure Positive Lumbar Discography

In response to criticisms of use of high pressure provocation, Carragee et al (21) performed a post-hoc exploratory analysis of prior data to evaluate the

false-positive rate among patients with low pressure positive discograms (defined as < 22 psi a.o.). Similar percentages of positive discs were found in asymptomatic subjects as patients with chronic discogenic low back pain, 17/69 (25%) and 14/53 (27%) low pressure positive responses, respectively. Carragee et al (21) concluded that there was no statistically significant difference between the asymptomatic or minimally symptomatic low-pressure positive subjects and symptomatic chronic low back pain patients. There are significant problems with the combination of several heterogeneous populations in this study. Two populations should arguably be removed from the analysis: patients with somatization disorder and patients with chronic low back pain ("backache"). Somatization disorder is a very rare, severe condition with a prevalence of 0.2% in males and 0.2 – 2.0% in females (47). Also, the data set of 4 patients is very small with a false-positive rate of 50% per patient and a confidence interval from 0 – 100%. Next, patients with symptomatic low back pain cannot be combined with 4 asymptomatic patient groups. Based on ISIS/IASP (25) criteria, the remaining pain-free, chronic pain, and post-discectomy groups have acceptably low false-positive rates whether analyzed individually or in combination analyses. For the chronic pain group, the ISIS/IASP (25) standards obtain a false-positive rate of 0% (Table 7); or a rate of 5.6% or 3.85% per patient and per disc respectively, based on combined analysis (Table 9). Amongst the post-discectomy patients, the false-positive rate is 15% per patient and 9.1% per disc (ISIS/IASP). Lastly, an arbitrary definition of low pressure positive of < 22 psi a.o. was utilized, which is not in keeping with generally reported criteria and standards of ≤ 15 psi a.o. (25,28). Combining these arbitrary groups to conclude that there is no difference between low pressure positive discography in subjects asymptomatic of significant low back pain illness and chronic discogenic pain patients is not supportable.

Carragee et al's post-hoc exploratory analyses of their prior studies concluded that low pressure criteria were essentially of no value. This result was obtained because the inclusion of symptomatic patients with chronic low back pain (as well as somatization disorder patients) obscured the merit of low pressure criteria. Our analysis shows that the low pressure criteria is reliable, and in fact, it performs similarly to the IASP/ISIS (25) standard. Reviewing the combined analyses, with pressure criteria set at ≤ 15 psi a.o., we obtain a false-positive rate of 0% for the asymptomatic subjects;

11.1% per patient and 7.7% per disc, for chronic pain patients; and 10.7% per patient and 6.9% per disc, for all subjects combined. The low pressure criteria of ≤ 15 psi a.o., performs slightly better than the < 22 psi threshold, with 10.7% per patient, 6.9% per disc and 16.0% per patient and 11.2% per disc, respectively.

If low pressure criteria are to be adopted in the future, discographers will most certainly have to refine their technique with pressure and speed-controlled manometry to be confident of the applied pressure stimulus. Future research will be needed to understand the significance of low pressure positive discs, as to whether there are indeed a subset of "chemically-sensitive" discs as suggested by previous research (28).

Meta-Analysis: Diagnostic Test Accuracy Review

Meta-analysis of all studies of asymptomatic patients undergoing discography (excluding symptomatic chronic low back pain patients and 2 somatization disorder patients with an incomplete dataset), obtains a specificity of 0.94 (95% CI 0.89 – 0.98) and a false-positive rate of 6%. Short-comings of this meta-analysis include lack of randomized-controlled trials and lack of a reference standard. Outcomes in diagnostic test accuracy meta-analyses are subject to bias introduced by heterogeneity, specifically in terms of patient population, use of different techniques, and the various positivity thresholds.

SUMMARY

Looking back over the last 40 years, provocation discography continues to incite controversy, with formidable opponents and proponents on both sides of the debate. Based on both a systematic review of lumbar discography in asymptomatic subjects (with a critical synthesis of the data), and a meta-analysis of the specificity of the test, discography cannot be dismissed by the results of several recent studies with alarming statistical calculations and clinical conclusions. If all subjects are combined and the ISIS/IASP (25) standard is applied, a false-positive rate of 9.3% per patient and 6.0% per disc can be obtained. Meta-analysis of all the studies obtains a test specificity of 0.94, on a per disc basis. Strength of evidence is rated at level II-2 per USPSTF criteria. The available data from asymptomatic discs of patients with chronic low back pain suggest that discography is a useful diagnostic test to confirm or refute the hypothesis that a particular disc

is or is not painful. Utilizing validated pressure instruments and subjects with a low pre-test prevalence of discogenic disease, false-positive responses are low and can be reduced to an acceptable level by using established standardized operational criteria for pain response, pressure stimulus, concordance, degree of annular disruption, and use of a control disc (25). The recommendation is that discography should be performed when the diagnosis of discogenic pain is highly suspected, i.e. after the history, physical exam, advanced imaging, negative diagnostic blocks of other common confounding pain generators (zygapophysial joints, sacroiliac joints etc.). The findings in this paper are also supported by 2 systematic reviews of discography as a diagnostic test for spinal pain, wherein strong evidence was reported for the ability of discography to evoke pain; for the role of discography in identification of patients with discogenic pain, and for the diagnostic accuracy of discography as an imaging tool (48,49).

There are several reasons for why a high false-positive rate may have been obtained in recent studies. First, the studies reported false-positive rates on a per patient instead of per disc basis. In 3 studies (Table 7) populations (iliac crest pain, somatization disorder, and post-discectomy) this leads to a significantly higher absolute number than if the data was presented per disc. It can be argued that the false-positive rate is best presented per disc, as provocation discography is designed as a per disc test to confirm or refute the hypothesis that the disc is the probable source of pain. As well, for treatment purposes, surgeons are interested in the number of pathologic disc levels. Next, the recent negative studies used uncontrolled manual pressurization to 100 psi a.o. with recording of static pressures only. Although this method is commonly used, uncontrolled pressurization can produce high intradiscal dynamic which can evoke significant pain in an otherwise non-pathologic disc (44).

Additionally, Carragee et al's studies (15) may have been biased towards a higher false-positive rate because of the subject population. All subjects, except the iliac crest pain and somatization disorder patients, had known symptomatic degenerative disc disease severe enough to require surgery. The study population of Walsh et al (15) was asymptomatic of any low back pain. In Derby et al's study (20), 6 of 13 subjects had occasional low back pain (less than 3 episodes of back pain per year). When the pre-test probability of dis-

ease (prevalence) is high, the positive predictive value (the likelihood that a patient with a positive discogram will have the disease) is also high. For example, when the pre-test probability of a disease (prevalence) is 90% the positive predictive value (the likelihood that a positive test will be a true positive test) is 98.6% (50). Carragee et al's subjects may have been asymptomatic with subclinical or not yet symptomatic disease which was provokable with high pressurization and high dynamic intradiscal pressures. The literature suggests that co-existence of cervical and lumbar disc disease is common. Researchers posit a common genetic influence on disc degeneration. In MRI studies of twins, heritability for "severe disease" was 79% and 64% in the cervical and lumbar spine, respectively (51). In another study of 200 patients with severe cervical disc disease requiring surgery (mean follow-up of 14 years), 100% of subjects reported significant episodes of back pain (suggestive of disc herniation) and/or underwent back surgery or had significant myelographic abnormalities (52).

As our knowledge accrues, we will continue to refine lumbar provocation discography as a diagnostic test. As stated by Jaeschke et al (53): "The ultimate criterion for the usefulness of a diagnostic test is whether it adds information beyond that otherwise available, and whether this information leads to a change in management that is ultimately beneficial to the patient." Confidence in the ability of provocation discography to determine the presence or absence of disease is a critical step towards appropriate therapeutic interventions. We also must continue to research the test's predictive value to select the best conservative, interventional, or surgical treatments for patients.

CONCLUSION

In this systematic review of lumbar provocation discography in asymptomatic subjects with a meta-analysis of false-positive rates, the experimental data for all published lumbar discography studies was analyzed. Meta-analysis of all studies of asymptomatic patients undergoing discography (excluding symptomatic chronic low back pain patterns and 2 somatization disorder patients with an incomplete data set) obtained specificity of 0.94 (95% CI 0.89 – 0.98) and a false-positive rate of 6%.

ACKNOWLEDGMENTS

The authors thank the Editorial Board of *Pain Physician* for suggestions to improve the manuscript and assistance in preparation of the final manuscript.

REFERENCES

- Dandy W. Loose cartilage from intervertebral disc simulating tumor of the spinal cord. *Arch Surg* 1929; 19:660.
- Lindblom K. Diagnostic disc puncture of intervertebral discs in sciatica. *Acta Orthop Scandinav* 1948; 17:231-239.
- Cloward R, Buzaid L. Technique, indications and evaluation of the normal and abnormal intervertebral disc. *A J R* 1952;552-564.
- Lindblom K. Technique and results of .diagnostic disc puncture and injection (discography) in the lumbar region. *Acta Orthopaedica Scandinav* 1951; 20:315-326.
- Gardner W, Wise R, Hughes C, FB OC, Wieford E. X-ray visualization of the intervertebral disk: With a consideration of the morbidity of disc puncture. *Arch Surg* 1952; 64:355-364.
- Wise R, Gardner W, Hosier R. X-ray visualization of the intervertebral disc. *New England J Med* 1957; 257:6-10.
- Boden SD, Davis DO, Dina TS, Patronas NJ, Wiesel SW. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg A* 1990; 72:403-408.
- Jensen MC, Brant-Zawadzki MN, Obuchowski N, Modic MT, Malkasian D, Ross JS. Magnetic resonance imaging of the lumbar spine in people without back pain. *N Engl J Med* 1994; 331:69-73.
- Horton WC, Daftari TK. Which disc as visualized by magnetic resonance imaging is actually a source of pain? A correlation between magnetic resonance imaging and discography. *Spine* 1992; 17:S164-S171.
- Zucherman J, Derby R, Hsu K, Picetti G, Kaiser J, Schofferman J, Goldthwaite N, White A. Normal magnetic resonance imaging with abnormal discography. *Spine* 1988; 13:1355-1359.
- Fraser RD. The North American Spine Society (NASS) on lumbar discography. *Spine* 1996; 21:1274-1276.
- Guyer RD, Ohnmeiss DD. Lumbar discography. Position statement from the North American Spine Society diagnostic and therapeutic committee. *Spine* 1995; 20:2048-2059.
- Massie WK, Stevens DB. A critical evaluation of discography. *J Bone Joint Surg Am* 1967; 49A:1243-1244.
- Holt EP. The question of lumbar discography. *Jl Bone Joint Surg Am* 1968; 50:720-726.
- Walsh TR, Weinstein JN, Spratt KF, Lehmann TR, Aprill C, Sayre H. Lumbar discography in normal subjects. A controlled prospective study. *J Bone Joint Surg Am* 1990; 72:1081-1088.
- Carragee EJ, Tanner CM, Yang B, Brito JL, Truong T. False-positive findings on lumbar discography. Reliability of subjective concordance assessment during provocative disc injection. *Spine* 1999; 24:2542-2547.
- Carragee EJ, Tanner CM, Khurana S, Hayward C, Welsh J, Date E, Truong T, Rossi M, Hagle C. The rates of false-positive lumbar discography in select patients without low back symptoms. *Spine* 2000; 25:1373-1380; discussion 1381.
- Carragee EJ, Chen Y, Tanner CM, Truong T, Lau E, Brito JL. Provocative discography in patients after limited lumbar discectomy: A controlled, randomized study of pain response in symptomatic and asymptomatic subjects. *Spine* 2000; 25:3065-3071.
- Carragee EJ, Alamin TF, Miller J, Grafe M. Provocative discography in volunteer subjects with mild persistent low back pain. *Spine J* 2002; 2:25-34.
- Derby R, Lee SH, Kim BJ, Chen Y, Aprill C, Bogduk N. Pressure-controlled lumbar discography in volunteers without low back symptoms. *Pain Med* 2005; 6:213-221; discussion 222-214.
- Carragee EJ, Alamin TF, Carragee JM. Low-pressure positive discography in subjects asymptomatic of significant low back pain illness. *Spine* 2006; 31:505-509.
- Carragee EJ, Chen Y, Tanner CM, Hayward C, Rossi M, Hagle C. Can discography cause long-term back symptoms in previously asymptomatic subjects? *Spine* 2000; 25:1803-1808.
- Carragee EJ, Barcohana B, Alamin T, van den Haak E. Prospective controlled study of the development of lower back pain in previously asymptomatic subjects undergoing experimental discography. *Spine* 2004; 29:1112-1117.
- Derby R. Lumbar discometry. *Newsletter of International Spine Injection Society* 1993; 1:8-17.
- Bogduk N. Lumbar disc stimulation; in Bogduk N (ed) *Practice guidelines for spinal diagnostic and treatment procedures*. International Spine Intervention Society, San Francisco, 2004, pp 20-46.
- West S, King V, Carey TS, Lohr KN, McKay N, Sutton SF, Lux L. Systems to rate the strength of scientific evidence. *Evidence report/technology assessment No.47. AHRQ Publication No. 02-E016*. Rockville, MD: Agency for Healthcare Research and Quality, 2002.
- Berg AO, JD A. Introducing the third United States Preventive Task Force. *Am J Prev Med* 2001; 20 (suppl 3):21-35.
- Derby R, Howard MW, Grant JM, Lettice JJ, Van Peteghem PK, Ryan DP. The ability of pressure-controlled discography to predict surgical and nonsurgical outcomes. *Spine* 1999; 24:364-371; discussion 371-362.
- Derby R, Kim BJ, Lee SH, Chen Y, Seo KS, Aprill C. Comparison of discographic findings in asymptomatic subject discs and the negative discs of chronic lbp patients: Can discography distinguish asymptomatic discs among morphologically abnormal discs? *The Spine J* 2005; 5:389-394.
- Shin D, Kim H, Jung J, Sin D, Lee J. Diagnostic relevance of pressure-controlled discography. *J Korean Med Sci* 2006; 21:911-916.
- Main CJ, Wood PL, Hollis S, Spanswick CC, Waddell G. The distress and risk assessment method. A simple patient classification to identify distress and evaluate the risk of poor outcome. *Spine* 1992; 17:42-52.
- Derby R, Eek B, Lee SH, Seo KS, Kim BJ. Comparison of intradiscal restorative injections and intradiscal electrothermal treatment (IDET) in the treatment of low back pain. *Pain Physician* 2004; 7:63-66.
- Sachs BL, Vanharanta H, Spivey MA, Guyer RD, Videman T, Rashbaum RF, Johnson RG, Hochschuler SH, Mooney V. Dallas discogram description. A new classification of CT/discography in low-back disorders. *Spine* 1987; 12:287-294.
- Maigne JY, Doursounian L. Entrapment neuropathy of the medial superior cluneal nerve. Nineteen cases surgically treated, with a minimum of 2 years' follow-up. *Spine* 1997; 22:1156-1159.
- Aoki Y, Takahashi Y, Takahashi K, Chiba T, Kurokawa M, Ozawa T, Moriya H. Sensory innervation of the lateral portion of the lumbar intervertebral disc in

- rats. *The Spine J* 2004; 4:275-280.
36. Morinaga T, Takahashi K, Yamagata M, Chiba T, Tanaka K, Takahashi Y, Nakamura S, Suseki K, Moriya H. Sensory innervation to the anterior portion of the lumbar intervertebral disc. *Spine* 1996; 21:1848-1851.
 37. Ohtori S, Takahashi K, Chiba T. Sensory innervation of the dorsal portion of the lumbar intervertebral disc in rats. *Spine* 2001; 26:946-950.
 38. Derby R, Lee S, Kim B. Discography; in Slipman C, Derby R, Simeone F, Mayer T (eds). *Interventional spine: An algorithmic approach*. , Saunders Philadelphia, PA, 2007, pp 291-302.
 39. Bogduk N, Aprill C. On the nature of neck pain, discography and cervical zygapophysial joint blocks. *Pain* 1993; 54:213-217.
 40. Derby R, Seo K, Chen Y, Kim B, Lee C, Hong Y, Lee J, Lee S. The influence of psychologic factors on diskography in patients with chronic axial low back pain. *Arch Phys Med Rehab* 2008 (in press).
 41. Manchikanti L, Singh V, Pampati V, Fel-lows B, Beyer C, Damron K, Cash KA. Provocative discography in low back pain patients with or without somatiza-tion disorder: A randomized prospec-tive evaluation. *Pain Physician* 2001; 4:227-239.
 42. Ketterer MW, Buckholtz CD. Somatiza-tion disorder. *J Am Osteo Assoc* 1989; 89:489-490, 495.
 43. Zoccolillo MS, Cloninger CR. Excess medical care of women with somati-zation disorder. *South Medl J* 1986; 79:532-535.
 44. Seo KS, Derby R, Date ES, Lee SH, Kim, BJ, Lee CH. In vitro measurement of pressure differences using manometry at various injection speeds during dis-cography. *The Spine J* 2007; 7:68-73.
 45. Schwarzer AC, Aprill CN, Derby R, Fortin J, Kine G, Bogduk N. The relative contri-butions of the disc and zygapophyseal joint in chronic low back pain. *Spine* 1994; 19:801-806.
 46. Manchikanti L, Singh V, Pampati V, Damron KS, Barnhill RC, Beyer C, Cash KA. Evaluation of the relative contri-butions of various structures in chron-ic low back pain. *Pain Physician* 2001; 4:308-316.
 47. APA: Diagnostic and statistical manual of mental disorders, DSM-IV (4th ed). Washington, DC, 1994.
 48. Shah RV, Everett CR, McKenzie-Brown AM, Sehgal N. Discography as a diag-nostic test for spinal pain: A systemat-ic and narrative review. *Pain Physician* 2005; 8:187-209.
 49. Buenaventura RM, Shah RV, Patel V, Be-nyamin R, Singh V. Systematic review of discography as a diagnostic test for spinal pain: An update. *Pain Physician* 2007; 10:147-164.
 50. Lang TA, Secic M. *How To Report Statis-tics in Medicine. Determining the Pres-ence or Absence of Disease*, ed 2nd edition. Philadelphia, PA, American College of Physicians, 2006.
 51. Sambrook PN, MacGregor AJ, Spector TD. Genetic influences on cervical and lumbar disc degeneration: A magnetic resonance imaging study in twins. *Ar-thr Rheum* 1999; 42:366-372.
 52. Jacobs B, Ghelman B, Marchisello P. Co-existence of cervical and lumbar disc disease. *Spine* 1990; 15:1261-1264.
 53. Jaeschke R, Guyatt GH, Sackett DL. Us-er's guide to the medical literature. Iii. How to use an article about a diagnos-tic test. B. What are the results and will they help me in caring for my patients? The evidence-based medicine working group. *JAMA* 1994; 271:703-707.