Provocative diskography: safety and predictive value in the outcome of spinal fusion or pain intervention for chronic low-back pain

Paul C Willems
Department of Orthopedic Surgery, CAPHRI Research School, Maastricht University Medical Center, Maastricht, the Netherlands

Abstract: There is still no clear definition of diskogenic low-back pain and no consensus on a generally agreed test, such as provocative diskography (PD), to diagnose painful disk degeneration, and probably more importantly, to predict the outcome of therapy intended to reduce pain that is presumed to be diskogenic in nature. Nevertheless, PD is the most specific procedure to diagnose diskogenic low-back pain. Its accuracy, however, is rather low or at best unknown. Although rare, the most prevalent complication, postdiskography diskitis, can be devastating for the individual patient, so all measures, like strict sterile conditions and antibiotic prophylaxis, should be taken to avoid this complication. It is advised to perform the procedure in a pressure-controlled way with a constant low flow, and optionally computed tomography imaging. PD should not be performed in morphologically normal disks. A standardized execution of the test should be established in order to perform high-quality studies to determine its accuracy to lead to meaningful interventions, and find best practices for diagnosis and treatment of diskogenic back pain. Possibly, PD may have detrimental effects on the disk, causing early degeneration, although it is unknown whether this will be related to clinical symptoms. Especially with these possible adverse side effects in mind, the risk–benefit ratio with the lack of clear benefits from treatments provided, and possible complications of disk puncture, the rationale for PD is questionable, which should be stressed to patients in the process of shared decision making. Diskography as a stand-alone test is not recommended in clinical decision making for patients with chronic low-back pain.

Keywords: provocative diskography, chronic low-back pain, prognostic accuracy, spinal fusion, pain intervention

Introduction

Chronic low-back pain (CLBP) is a major health problem in modern society, with lifetime prevalence up to 84%.1 The economic burden of low-back pain (LBP) is huge, and consists of direct costs of health care utilization (hospitalization, medication, tests, and therapies) and indirect costs of lost productivity due to work absenteeism and early retirement. In a small Western European country, such as the Netherlands, the total annual costs of back pain were estimated at €4.4 billion2 (mainly employment-related costs).3 The total annual costs in the US have exceeded $100 billion.4 Roughly, LBP can be categorized as pain caused by spinal pathology, such as tumors, infection, trauma, nerve root or radicular pain, and a large heterogeneous group of patients (about 85% of total cases6) suffering from LBP in whom imaging reveals signs of degeneration of one or more intervertebral disks (disk-space narrowing, vertebral end-plate changes,7,8 annular disruption,9,10 and/or facet joint arthropathy). These degenerative findings can also be observed in asymptomatic subjects,11 and thus
this last category of patients is often mentioned as suffering from CLBP. CLBP has a substantial impact on the lifestyle of those affected, and several socioeconomic risk factors have been identified.\textsuperscript{12,13} Job dissatisfaction, low education, strenuous work, and workers’ compensation insurance are all associated with CLBP and resulting disability.\textsuperscript{14,15} In addition, factors such as smoking, obesity, psychological distress, depressive mood, and to a lesser extent somatization, have been reported to result in increased risk of chronicity.\textsuperscript{16}

**Rationale for invasive treatment of CLBP**

Intervertebral disk degeneration is considered to be the main cause of CLBP. The etiology and pathophysiology of disk degeneration, however, is unknown. Genetic predisposition may play an important role.\textsuperscript{17,18} Histologically, an abnormal ingrowth of sensory nerve fibers in the end plates and the nucleus pulposus has been observed.\textsuperscript{20} Radial fissures in the annulus fibrosus have been linked to the ingrowth of free nerve endings (nociceptors) and blood vessels with a granulation zone.\textsuperscript{21}

Additionally, degenerative herniated disks contain high levels of proinflammatory mediators and cytokines,\textsuperscript{22-23} eg, interleukin 6, nitric oxide, prostaglandin E\textsubscript{2},\textsuperscript{24} and phospholipase A\textsubscript{2}.\textsuperscript{25}

These findings have led to the idea that disk degeneration, with or without secondary osteoarthritis of the synovial facet joints, is a major source of pain in CLBP, which has led to minimally invasive interventional treatment of the disk to relieve pain. As the pain appears to be largely associated with continued motion and loading of the affected disk level, the idea has emerged of treating a painful degenerative-motion segment by spinal fixation and fusion.

First described in 1889 for stabilization of vertebral segments in spinal tuberculosis,\textsuperscript{26} spinal fusion is a successful treatment for fractures and deformity of the spine. Spinal fusion for CLBP, however, remains controversial.\textsuperscript{27,28} There is enormous variation in the rate of spinal fusion operations between countries and even regions within countries.\textsuperscript{29-31} Apparently, there is no true consensus on its indications, as was shown in a national survey among Dutch spine surgeons, which revealed a complete lack of uniformity in surgical decision making for CLBP.\textsuperscript{32}

In a Cochrane review in 1999,\textsuperscript{31} there was no evidence for the effectiveness of spinal fusion for CLBP. Nevertheless, there was a 220% increase in spinal fusion surgery from 1990 to 2001 in the US,\textsuperscript{33} and Medicare expenses for lumbar fusion increased in 10 years from $75 million to $482 million, which was about 50% of the total costs for spine surgery in the US.\textsuperscript{31}

The first randomized controlled trial\textsuperscript{34} showed a better clinical outcome for patients treated who underwent spinal fusion compared to patients who received standard conservative care. At longer follow-up, however, this beneficial effect was attenuated.\textsuperscript{35} Two more recent randomized controlled trials compared the outcome of spinal fusion to cognitive behavioral-based exercise therapy\textsuperscript{36} or an intensive rehabilitation program.\textsuperscript{37} The clinical results for surgery and the rehab programs at 2 years follow-up were equal. However, in the last trial, fusion had a higher complication rate and was less cost-effective than the structured rehab programs.\textsuperscript{38}

Most spine surgeons do not consider complications or bad surgery, but bad patient selection, wrong diagnosis, and disproportionate expectations of patients to be the major factors for poor outcome in spinal surgery.\textsuperscript{39}

The identification of subgroups of CLBP patients who truly benefit from spinal fusion would be helpful to improve the results of surgery. Unfortunately, the current routine diagnostic tools to select the right patient for lumbar fusion appear to be inadequate. History taking and physical examination can reveal psychosocial factors that lead to chronicity of pain,\textsuperscript{40,41} but there are no specific physical findings to predict the outcome of lumbar fusion.\textsuperscript{42} The association of CLBP with findings on imaging is weak.\textsuperscript{43} Degenerative signs on plain radiographs are not correlated with persisting symptoms of CLBP,\textsuperscript{44} and magnetic resonance imaging (the imaging study of choice for the initial evaluation of patients with CLBP)\textsuperscript{45} shows high false-positive rates of degenerative findings in asymptomatic people, especially in the elderly.\textsuperscript{11,46}

**Provocative diskography as an aid in clinical decision making for pain intervention**

To identify those patients who will benefit from invasive treatment aimed at reducing presumed diskogenic pain, physicians may use provocative diskography (PD), which is assumed to predict the outcome of such treatment.

PD can be performed in an outpatient operating room with monitored anesthetic care under strict sterile conditions. The patients should receive instructions to fast for at least 6 hours before the procedure. Optimally, all diskographies are performed by a certified pain-interventional physician who performs the procedure frequently (preferably more than 50 annually). Thirty minutes before the intervention, prophylactic antibiotics (2 g cefazolin) can be administered intravenously.

Patients can be placed in the lateral decubitus or in the prone position on a radiolucent table (Figure 1). The skin in the low back and gluteal region is thoroughly disinfected with
Provocative diskography is performed under sterile conditions and with biplanar fluoroscopic control. The intended disk space is injected with contrast agent to provoke usual back pain. In addition, the amount of degeneration can be visualized on radiographs or computerized tomography (CT).

Chlorhexidine or iodine and draped. In the prone position, fluoroscopy is performed with the direction of the radiation beam positioned parallel to the subchondral plate of the lower vertebral end plate of the disk. In the disks above L5–S1, the fluoroscopy tube is rotated ipsilaterally until the lateral aspect of the superior articular process projects over the axial middle of the disk to be punctured. The target area for injection is the lateromedial side of the disk, just lateral to the edge of the superior articular process. At the L5–S1 level, the fluoroscopy tube is rotated until the lateral edge of the superior articular process of S1 projects approximately 25% over the posteroanterior distance of the vertebral body. The double-needle technique is advocated, in which an outer 18 g x 3.5” (8.89 cm) introducer needle is advanced under fluoroscopic control to the margin of the disk (Figure 2). Next, an inner 22 g x 6.0” needle is led through the outer needle and placed in the center of the disk. The double-needle technique is assumed to help in reducing the incidence of postdiskography diskitis, and allows the entering of the disk with needles of a small diameter.

When contrast dye is injected manually, generally high-pressure peaks are achieved. It has been shown in asymptomatic subjects that in the case of provocation with high pressures above 50 psi, eventually all disks can be painful. Therefore, in order to avoid false-positive findings, PD should be carried out with a constant low flow and controlled continuous manometry by an automated system. For this, the diskography needle is connected to a pressure-controlled diskography device, which is calibrated for needle diameter.

Figure 2 (A) A schematic axial view of the intervertebral disk with preferred location of the needle tip in the center. Needle-approach 3 is recommended as approaches 1 and 2 have a higher risk of leakage of cerebrospinal fluid with subsequent postprocedural headache. (B) A radiographic image of the lumbar spine after multilevel diskography showing in the upper disk an intact annulus fibrosus (contained contrast agent in the nucleus pulposus). The lower 3 disks show degeneration with annular tears as evidenced by leakage of contrast agent to the outer disk.
needle length, contrast-dye viscosity, and fluid velocity. Preferred controlled-injection velocities are 0.05 mL/second and 0.02 mL/second.\textsuperscript{49} During injection, there should be continuous communication with the patient, and the patient’s pain should be noted as rated on a numeric rating scale before and during the procedure. The moment the contrast dye first enters the nucleus pulposus, pressure is recorded as the opening pressure. The maximum accepted pressure above the opening pressure has been recommended to be 50 psi,\textsuperscript{60} and the maximum injection volume 3 mL. Pain pressure level and volume level are recorded at the moment when the patient experiences concordant pain.

The degree of disk degeneration (Figure 2) can be classified according to the Dallas diskogram scale,\textsuperscript{51} in which grade 0 means no disruption, grades 1, 2, and 3 represent leakage of the contrast dye into the inner, middle, and outer layers of the annulus fibrosus, respectively, and in grade 4 there may be leakage outside the annulus (complete rupture). Vanharanta et al\textsuperscript{52} showed that grades 0–2 are rarely painful and over 75% of grade 3 diskographies are painful.

Nevertheless, PD remains a controversial test, because it requires the patient’s subjective pain response, and its reported accuracy for identifying diskogenic LBP has been reported as moderate to low.\textsuperscript{42} Of course, precise measurement of contrast volume and intradiskal pressure during injection will improve standardization and reproducibility of diskography, as the patient’s pain response is related to the intensity of the provocative stimulus.\textsuperscript{49} Whether it will increase the diagnostic accuracy of PD, however, remains to be seen. The International Spine Intervention Society\textsuperscript{53} recommends four criteria at diskography in order to define diskogenic back pain: 1) provocation of the suspect disk induces concordant pain, 2) the pain is at least 7/10 on a numeric rating scale, 3) the pain is provoked by less than 50 psi above opening pressure, and 4) provocation of at least one adjacent disk does not induce concordant pain. In a meta-analysis, pressure-controlled diskography with low-pressure criteria appeared to reduce the rate of false-positive disks to 0.06 in asymptomatic subjects.\textsuperscript{54} However, the review suffered from serious methodological limitations, and as there are no data on the sensitivity of PD for pain-interventional treatment, the clinical utility of these findings is low.

In a recent systematic review, diagnostic criteria of presumed diskogenic back pain were studied.\textsuperscript{55} In 36 studies, PD was used as the predominant technique. Additionally, analgesic diskography or “diskoblock” was used in one study.\textsuperscript{56} Of all included PD studies, only eight used pain scales to quantify the intensity of pain during disk injection. Six studies had measured intradiskal pressure during the procedure, and the absence of concordant pain at an adjacent level has been required in only 17 studies. As treatment for the suspected diskogenic back pain included spinal fusion, total disk replacement, disk denervation, and intradiskal injection techniques were used. Remarkably, provocation diskography was used in virtually all studies (95%) that used minimally invasive pain intervention as treatment, whereas in surgical studies PD was used preoperatively in less than 60%. As terminology and definitions used depended mainly on the treating physician’s specialty, there appeared to be no clear definition of a painful disk and no reliable test to predict treatment outcome.

### The accuracy of PD in diagnosing diskogenic pain as a prognostic tool for successful spinal fusion

PD is based on the assumption that pressurizing a symptomatic degenerated disk will incite nociceptive nerve endings, thus producing the pain that the patient recognizes as his or her daily-life back pain. Ideally, a gold standard confirming the presence or absence of diskogenic back pain should be available to determine diagnostic accuracy. However, surgical inspection of a degenerated disk cannot determine whether this disk is causing the patient’s pain or not. The greatest problem for PD is the absence of a reliable reference standard.

Spinal fusion results have been used as a reference standard to determine the prognostic accuracy of PD. However, it should be stressed that neither spinal fusion nor lumbar disk replacement is a reliable and predictable effective treatment for diskogenic LBP. Although PD is used widely as a selection tool for operative treatment, very few studies have reported its prognostic accuracy for surgical outcome. In a recent systematic literature review,\textsuperscript{57} only four studies could be identified in which sensitivity, specificity, and predictive values of PD for spinal fusion outcome could actually be calculated, as in these studies negatively tested patients were also operated on and the risk of verification bias was thus avoided (Table 1). In particular, it appeared that the specificity of PD was very low (below 0.5), meaning that unnecessary surgery in patients who would not benefit from fusion could not be avoided, as the test did not identify these patients. It was therefore concluded that PD should not be used in clinical practice as a preoperative prognostic tool for the outcome of spinal fusion for LBP.

### Safety and long-term effects of PD

Postdiskography diskitis, mainly caused by \textit{Staphylococcus aureus}, is the most feared complication of PD. Although rare, diskitis can have serious consequences. The risk of
postdiskography diskitis can be reduced by the use of intravenous or intradiskal antibiotic prophylaxis. Fraser et al. showed that the use of a double-needle technique reduced the risk of postdiskographic diskitis from 2.7% to 0.7%. In a systematic literature review, Willems et al. reported that in PD, without the use of prophylactic antibiotics, an overall diskitis incidence of 0.25% by patient and 0.094% by disk was found. In one of these selected papers, an unusually high incidence (1.90% by patient and 0.98% by disk) was reported. This may have been caused by the fact that PD had been executed by less experienced diskographers. In the literature search, only one study of lumbar diskography with the use of prophylactic antibiotics, in which diskitis was specifically looked for, was selected. In this series of 127 patients (337 disks), no case of diskitis was registered. However, given the aforementioned incidence of 0.25% by patient, to actually prove that additive antibiotics really prevent postdiskographic diskitis, a randomized trial of 9,000 patients would be needed. It was concluded that the routine use of prophylactic antibiotics in PD is not indicated. PD should be carried out by experienced physicians who are meticulous with the two-needle technique and aseptic conditions. Nevertheless, international guidelines have recommended the routine use of antibiotic prophylaxis for PD, in order to minimize the risk of postdiskography diskitis.

An alternative or addition to traditional PD is the technique known as diskoblock, which involves injecting the disk with an anesthetic agent instead of a contrast agent in an effort to eliminate instead of reproduce a patient’s pain. This diskoblock technique was compared with PD by Ohtori et al. Forty-two patients planned for noninstrumented anterior lumbar interbody fusion with iliac crest autograft were randomized for a preoperative workup with either PD or diskoblock. Twelve patients were eliminated because of a lack of response with both techniques. No patient showed signs of pseudarthrosis at 2 years follow-up, based on plain radiographs and computed tomography scans. Those fused patients with positive diskoblock findings had a significantly better outcome in terms of pain and disability compared to those fused patients who had positive PD. This study provided level II evidence that the diskoblock technique might be a better predictor of clinical outcome after fusion surgery than traditional PD. A small sample size, lack of a power analysis, and the presence of verification bias (negatively tested patients did not receive fusion) limited this conclusion.

Recently, there has been concern that diagnostic disk injections may lead to iatrogenic injury of the disk, thus accelerating disk degeneration. Degeneration of an intervertebral disk due to a needle puncture has been demonstrated in animals. Carragee et al. followed the progression of disk degeneration in magnetic resonance imaging studies in patients with and without a history of diskography. There was significantly more progression of disk degeneration and herniations in patients who had undergone PD.

### Conclusion

There is still no clear definition of diskogenic LBP and no consensus on a generally agreed test, such as PD, to diagnose painful disk degeneration, and probably more importantly, to predict the outcome of therapy intended to reduce pain that is presumed to be diskogenic in nature.

PD is used in clinical practice to diagnose diskogenic LBP. It is advised to perform the procedure in a pressure-controlled way with a constant low flow, and optionally computed tomography imaging. The accuracy of PD in predicting the outcome of spinal fusion is low, with specificity less than 0.5, meaning that most patients who will not benefit from fusion cannot be identified by the test. For pain-interventional treatment, there are not enough reliable data on the sensitivity of PD to predict treatment outcome in diskogenic back pain patients at all. Therefore, the test should not be recommended as a standalone test for treatment strategy in clinical practice.

Possibly, PD may have detrimental effects on the disk, causing progression of degeneration, although it is unknown whether this will be related to clinical symptoms. With regard to these possible adverse side effects, the risk–benefit ratio is uncertain, which should be stressed in the process of shared decision making with patients.

### Table I Summary prognostic accuracy of provocative diskography for spinal fusion outcome

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample size</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Positive predictive value</th>
<th>Negative predictive value</th>
<th>Positive LR (95% CI)</th>
<th>Negative LR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colhoun et al</td>
<td>168</td>
<td>0.88</td>
<td>0.48</td>
<td>0.88</td>
<td>0.48</td>
<td>1.71 (1.21–2.41)</td>
<td>0.24 (0.13–0.43)</td>
</tr>
<tr>
<td>Esses et al</td>
<td>22</td>
<td>0.40</td>
<td>0.43</td>
<td>0.60</td>
<td>0.25</td>
<td>0.70 (0.29–1.70)</td>
<td>1.40 (0.54–3.62)</td>
</tr>
<tr>
<td>Gill and Blumenthal</td>
<td>53</td>
<td>0.81</td>
<td>0.41</td>
<td>0.74</td>
<td>0.50</td>
<td>1.37 (0.89–2.10)</td>
<td>0.47 (0.20–1.13)</td>
</tr>
<tr>
<td>Willems et al</td>
<td>82</td>
<td>0.73</td>
<td>0.27</td>
<td>0.45</td>
<td>0.55</td>
<td>0.99 (0.76–1.30)</td>
<td>1.01 (0.49–2.08)</td>
</tr>
</tbody>
</table>

**Abbreviations:** CI, confidence interval; LR, likelihood ratio.
Disclosure

The author reports no conflicts of interest in this work.

References